

Microphone array positioning technique with Euclidean distance geometry

Simon Bouley, Charles Vanwynsberghe, Thibaut Le Magueresse, Jérôme Antoni, Allan Outrequin

▶ **To cite this version:**

Simon Bouley, Charles Vanwynsberghe, Thibaut Le Magueresse, Jérôme Antoni, Allan Outrequin. Microphone array positioning technique with Euclidean distance geometry. Applied Acoustics, 2020, 167, pp.107377. 10.1016/j.apacoust.2020.107377 . hal-02560747

HAL Id: hal-02560747
https://hal.science/hal-02560747

Submitted on 3 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Applied Acoustics 167 (2020) 107377

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Microphone array positioning technique with Euclidean distance geometry

Simon Bouley [a,b,*], Charles Vanwynsberghe [c,b], Thibaut Le Magueresse [a], Jérôme Antoni [b], Allan Outrequin [a]

[a] MicrodB, F-69134 Écully, France
[b] Univ Lyon, INSA-Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France
[c] Lab-STICC UMR 6285, CNRS, ENSTA Bretagne, 2 rue François Verny, 29200 Brest, France

ARTICLE INFO

Article history:
Received 8 October 2019
Received in revised form 11 February 2020
Accepted 9 April 2020

Keywords:
Array processing
Source localization
Self-calibration

ABSTRACT

Source localization and quantification by an acoustic array of microphones depend to a great extent on an accurate knowledge of the antenna position towards the radiating device. The present work details a methodology to determine the location of the microphones in relation to an object of study, starting from its geometric shape and that of the array, in order to reproduce an experimental configuration in any retro-propagating method. A set of reference sources are placed on several prominent locations of the device to estimate the times of flight (ToF) (and distances) between them and the microphones, connecting the array and the object together. The overall geometric configuration is thus defined by an Euclidean Distance Matrix (EDM), which is basically the matrix of squared distances between points. First, MultiDimensional Unfolding (MDU) technique is used to reconstruct the point set from distances. Second, this point set is then aligned with the device, using reference sources as anchor nodes. This orthogonal Procustes problem is solved by the Kabsch algorithm to obtain the optimal rotation and translation matrices between the coordinate system of the array and that of the object of study. The methodology is detailed, validated first by a numerical simulation of a typical experimental set-up. An experimental campaign is finally carried out to assess the robustness of the method in a typical test case.
© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The localization and quantification of acoustic sources radiated by a device depend on numerous physical parameters as well as the microphone array geometry or the chosen retro-propagating method. As highlighted recently by Gilquin et al. [1] by means of sentitivity analysis, deviations and uncertainties of the antenna position and orientation in its experimental environment greatly influence the sound source reconstruction, both with a classical beamforming technique or an inverse matrix formulation.

Furthermore, the recent development of wireless acoustic sensor on devices such as smartphones led to a great deal of efforts to determine the true position of each sensor in a network, in order to enhance speech processing and speaker localization. The self-calibration of a microphone network has then become an active research field, as shown by the review articles of Plinge et al. [2] and Wang et al. [3]. Valente et al. [4] proposed a method to estimate the position of one array with respect to another one.

The antenna positioning problem tackled in this paper is rather different: one array of microphones is facing towards a radiating device to obtain a sound source map. The shapes of the antenna and the device are assumed to be completely known in their own coordinate system. A mesh of the surface of interest is built to retro-propagate the acoustic field measured by the array on the device. Finally, this study proposes a method to determine the antenna position in relation to the device during an experimental campaign and collect all microphone coordinates in the mesh coordinate system to perform the retro-propagation.

This application-oriented paper gathers different techniques based on Euclidean distance geometry to develop a practical tool. This paper is organized as follows. The theoretical elements are first developped step by step. Section 2.1 details how to define the overall geometric configuration, composed of the antenna and a set of acoustic sources placed on the device, with an Euclidean distance matrix (EDM). Taking the uncertainties of the sources position into account, the EDM is then completed and its related point set reconstructed from the distances with a MultiDimensional Unfolding (MDU) technique, presented in Section 2.2. The set is then aligned with the collection of sources, defined as anchors, by a rigid transformation between the coordinate system of the array and the device, provided by the Kabsch algorithm (Section 2.3). The robustness of the method is illustrated with numerical studies, whose configuration is representative of a

* Corresponding author at:MicrodB, F-69134 Écully, France.
E-mail address: simon.bouley@microdb.fr (S. Bouley).

https://doi.org/10.1016/j.apacoust.2020.107377
0003-682X/© 2020 Elsevier Ltd. All rights reserved.

typical experimental set-up (Section 3). Finally, an experimental evaluation is carried out to assess the robustness of the method in an industrial application (Section 4).

2. Methodology

2.1. Euclidean distance matrix

2.1.1. Properties

The principle of the proposed method can be summarized as follows: there is a mathematical relationship between the position of points in an Euclidean space and the distances between them. This connection is performed by a mathematical tool called Euclidean Distance Matrix (EDM).

Consider a d-dimensional Euclidean space, where a collection of n points, representative of a network or the geometry of an object, are set and described by the columns of the matrix $\mathbf{X} \in \mathbb{R}^{d\times n}, \mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n], \mathbf{x}_i \in \mathbb{R}^d$. The terms d_{ij} of an Euclidean distance matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ are the squared distances between points $\mathbf{x}_i$ and $\mathbf{x}_j$:

$$d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|_2^2, \tag{1}$$

where $\|\cdot\|_2$ is the Euclidean norm. The main use of this tool is to reconstruct the initial configuration corresponding to this matrix, which defines an inverse problem. An EDM fulfills the following properties:

- Non-negativity ($d_{ij} \geqslant 0, i \neq j$)
- Hollow matrix ($d_{ij} = 0 \Longleftrightarrow i = j$)
- Symmetry: $d_{ij} = d_{ji}$

Furthermore, as shown by Gower [5], the rank of an EDM $\mathbf{D}$ related to the set of points $\mathbf{X}$ satisfies the inequality:

$$\text{rank}(\mathbf{D}) \geqslant d + 2. \tag{2}$$

The EDM is also invariant under orthogonal and rigid transformations (i.e. any rotation, reflection or translation). As a consequence, the absolute position and orientation of a point set cannot be reconstructed from its associated EDM. Each result is then a rigid transformation of another one. For more details, Parhizkar [6] provides a complete description of the EDM properties.

2.1.2. Constituting the EDM

Fig. 1 illustrates the ground truth of the antenna positioning problem, as met in an experimental campaign. The coordinates of each microphone (dots) in the array reference system is known, as the distances between them. Acoustic sources (squares) are placed at some prominent locations of the device. They are gathered in a group of four by a structural support (Fig. 2), whose fulcrum is sequentially in contact with each prominent location of the device (at least four, also, see Section 2.3). In that way, the position of the fulcrum in the reference frame of the device is determined. The geometry of this structural support and the position of the sources in its coordinate system are set. Because of the experimental difficulties to align the support with the device, its true orientation (azimuthal and elevation angles) is considered unknown.

At the initializing step of the method, the EDM of the antenna ($\mathbf{D}_a$) and the one formed by the acoustic sources set on one support $\mathbf{D}_{s_i}$ are already known, contrary to the distances between the sources and the microphones, gathered in the off-diagonal submatrice $\mathbf{D}_{as}$. The lack of knowledge due to the orientation of the support can be integrated in the proposed method. The EDM $\mathbf{D}_s$ of all the sources placed on the device is composed of diagonal block submatrices related to the sources of one support, sequentially moved to several locations. If the geometry of the support is unchanged between two measurement points, those submatrices are identical. Off-diagonal submatrices of all zeros represent the unknown distances between sources put on supports located at different places of the device.

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_a & \mathbf{D}_{as} \\ \mathbf{D}_{as}^T & \mathbf{D}_s \end{pmatrix}, \quad \mathbf{D}_s = \begin{pmatrix} \mathbf{D}_{s_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{s_i} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{D}_{s_n} \end{pmatrix}. \tag{3}$$

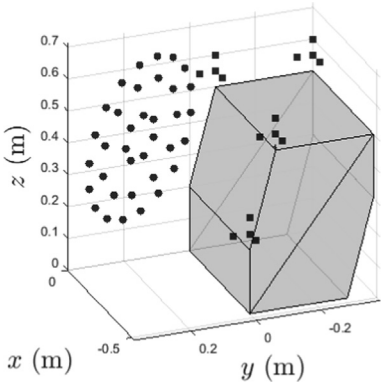


Fig. 1. Ground truth of the antenna positioning problem. Microphones are represented by dots while squares depict acoustic sources.

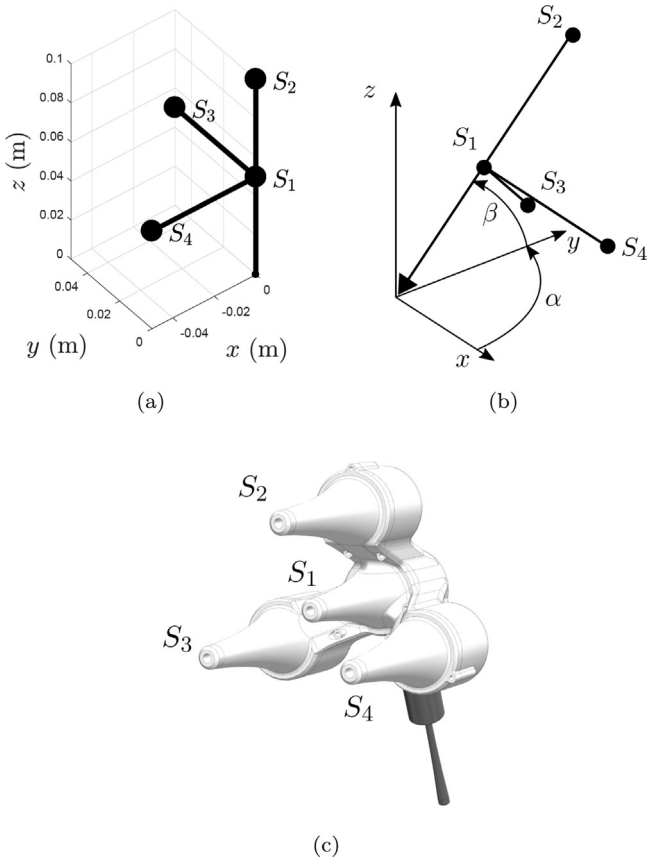


Fig. 2. (a): Structural support of the four acoustic sources (dots) in its coordinate system. At the origin is a fulcrum (tip down triangle), in contact with a prominent location of the device. (b): Unknown angular orientation of the sources. α and β are the azimuthal and elevation angles, respectively. (c): CAD illustration of the experimental structural support.

Denote by n the number of measurement points on the device and s the number of acoustic sources on a support, the number of missing entries is equal to $(n^2 - n) \times s^2$. For instance, for a group of four sources set on a support and four measurement points (theoretical minimum to solve the problem, see Section 2.3), the number of unknowns due to the orientation of sources is 192. The known and unknown data in the EDM related to sources (black and white, respectively) are displayed within a logical matrix in Fig. 3-a.

Considering the antenna, a logical matrix of the global configuration can be defined, represented in Fig. 3-b, composed of the diagonal block submatrices related to microphones (top-left block) and to sources (bottom-right block). The size of each block is defined by the number of microphones or sources, respectively. The distances between microphones and sources, represented by the blank off-diagonal submatrices remain unknown.

The distance between a source and microphone is determined by a Time of Flight (ToF) measurement. It is obtained by cross-correlating the microphone signal with the sound emitted by the synchronized reference source. The benefit of the synchronization is the possibility to select the first peak of cross-correlation related to the straight path, avoiding reflection issues due to the proximity between the array and the device. An evaluation of the speed of sound is finally needed to calculate the distance from the time of flight. The sound of reference sources is chosen according to the quality of the cross-correlation measurements. Each source placed on one support can simultaneously emits uncorrelated white noise or modulated sweep signals. As the peak detection depends on the length of the time sample, the sampling frequency is an essential parameter. At a fixed sampling frequency f_s, the maximum error on distance Δd induced by the time sampling is $\Delta d = c_0/f_s$. At

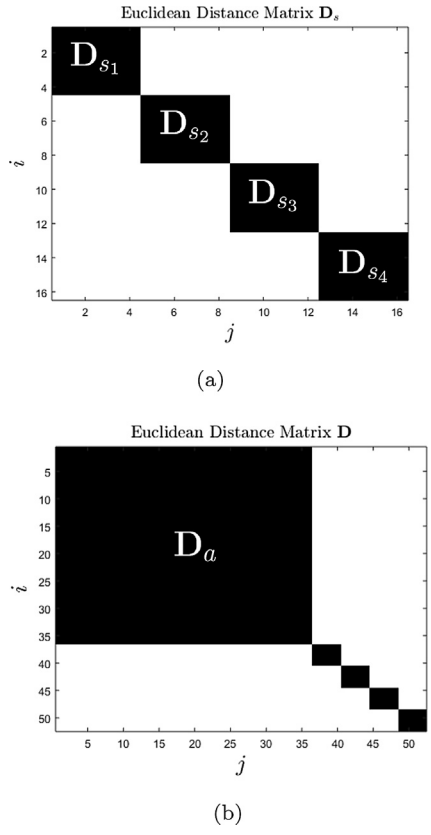


Fig. 3. Logical matrices of the sources (a) and the global configuration (b). The black diagonal submatrices represent the observed entries while blank off-diagonal submatrices illustrate the unknowns.

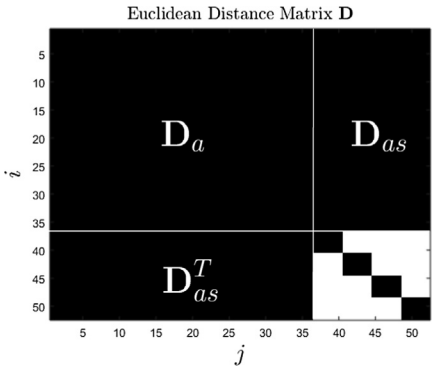


Fig. 4. Logical matrix of the global configuration after the determination of the source – microphone distances. Black and white submatrices represent the observed and missing entries, respectively.

$f_s = 50$ kHz, Δd is theoretically smaller than 7 mm ($\pm$3.5 mm). Fig. 4 represents the obtained EDM at the end of this step, where only the distances between sources on structural supports placed at different locations are missing.

2.2. Multidimensional Unfolding

Once the EDM is set, the second step is to reconstruct the overall geometry, gathering the antenna and the acoustic sources from the distances between points. MultiDimensional Scaling (MDS) refers to a collection of techniques for the analysis of similarity or dissimilarity in a dataset. Initially developped in psychometrics [7], MDS allows to model a wide range of data as distances and visualize them as points in a geometric space. The classical MDS, also known as Torgerson-Gower scaling [8], finds a coordinate matrix $\widehat{\mathbf{X}}$ starting from an EDM $\mathbf{D}$ and the embedded dimension d. The mathematical developments are detailed by Borg and Groenen [8] and summarized by Dokmanić et al. [9].

The method is based on an eigendecomposition of the Gram matrix ($\mathbf{G} = \mathbf{X}^T\mathbf{X} = \mathbf{U}\Lambda\mathbf{U}^T$, where $(\bullet)^T$ denotes the transpose operator). According to Gower [10], this matrix can be computed with a double centering, using the geometric centering matrix $\mathbf{J}$:

$$\mathbf{G} = -(1/2)\mathbf{JDJ}, \quad \mathbf{J}(n) = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T, \tag{4}$$

where $\mathbf{1}$ is a column vector filled with ones. The eigenvalues λ_i are sorted in order of decreasing amplitude and only the first d values are selected. Thus, the point set $\widehat{\mathbf{X}} = \left[\mathrm{diag}\left(\sqrt{\lambda_1}, \cdots \sqrt{\lambda_d}, \mathbf{0}_{d\times(n-d)}\right)\right]\mathbf{U}^T$ is embedded in a d-dimensional space.

This algorithm is well-suited when the EDM is complete, which is not the case of this study. However, the EDM is a highly structured matrix, whose properties are largely used in completion algorithms from the literature. The particular configuration of this study reduces to a simplified case of completion by MultiDimensional Unfolding (MDU). Introduced by Schönemann [11], MDU was developped to solve algebraically the problem of locating two sets of points in a joint space, given the Euclidean distances between elements from distincts sets. In this classical exemple, only the distances between microphones and sources are known, while the true geometries of the antenna and the sources are missing. The EDM reduces to off-diagonal block submatrices. In our particular configuration, the EDM of the antenna and the sources placed on a support are considered as input data and the distances between microphones and sources are experimentally determined.

After Schönemann [11], several methods have been recently proposed to tackle this completion problem (see e.g. Crocco et al.

[12]). In their review article, Dokmanić et al. [9] detailed some algorithms and compared their performance on MDU. One of them, developped by Parhizkar [6] is thereafter used in the methodology to perform this step.

2.2.1. Reconstructing the geometry with incomplete EDM

Consider a set of n points $\mathbf{X}$ embedded in a d-dimensional space and its corresponding EDM $\mathbf{D}$. As seen in the previous section, the observed EDM $\tilde{\mathbf{D}}$ can be noisy (sampling issue (see subSection 3.3.1), measurement uncertainty) and incomplete:

$$\tilde{\mathbf{D}} = \mathbf{W} \circ \mathbf{D} + \mathbf{N}, \tag{5}$$

where $\circ$ is the Hadamard product, $\mathbf{N}$ is a symmetric noise matrix and $\mathbf{W}$ a logical (or mask) matrix:

$$w_{i,j} \triangleq \begin{cases} 1, & (i,j) \in E \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

The set E gathers the indices of observed distances of the matrix $\tilde{\mathbf{D}}$. The aim of the method here is to find an estimated point set $\widehat{\mathbf{X}}$ so that its corresponding EDM $\widehat{\mathbf{D}}$ is close to the observed EDM $\tilde{\mathbf{D}}$. The algorithm is based on the minimization of the dissimilarity metric called *s-stress criterion*, introduced by Takane et al. [13]:

$$\underset{\widehat{\mathbf{X}} \in \mathbb{R}^{d \times n}}{\text{minimise}} \sum_{(i,j) \in E} \left(\text{edm}\left(\widehat{\mathbf{X}}_{ij}\right) - \tilde{d}_{ij} \right)^2. \tag{7}$$

This cost function is the Frobenius norm of the differences between the estimated and the observed EDM: $\| \left(\widehat{\mathbf{D}} - \tilde{\mathbf{D}} \right) \|_F^2$. Since the *s-stress criterion* is separable across points and coordinates, Parhizkar [6] proposed a distributed algorithm called *alternating coordinate descent method*. For each point, the derivative of the cost function with respect to each coordinate consists in a polynomial whose roots can be analytically calculated. Their coordinates are updated at a time and the process is repeated for each point of the set until convergence.

2.3. Optimal orientation and position of the array

As explained in Section 2.1.1, the EDM is invariant under orthogonal and rigid transformations. Therefore, the absolute orientation and position of the point set cannot be derived from the multidimensional unfolding. A new step is needed to find the optimal rotation/reflection and translation matrices which align the point set in the reference frame of the device. This is performed with a selection of n_s points, denoted as anchors, whose positions $\mathbf{X}_s$ in this particular coordinate system are known. This step is usually called orthogonal Procustes analysis [14]. One solution, stemmed from crystallography, is the Kabsch algorithm [15,16], which computes the optimal rotation matrix $\mathbf{R}$ between two sets of points by minimizing the root mean square deviation (least RMSD, Eq. (8)).

$$\text{lRMSD} = \underset{\mathbf{R} \in \mathbb{R}^{d \times d}}{\text{argmin}} \sqrt{\frac{1}{n_s} \sum_{i=1}^{n_s} | \mathbf{R}\widehat{\mathbf{x}}_{s,i} - \mathbf{x}_{s,i} |^2}. \tag{8}$$

First of all, the two sets of anchors in both reference frames ($\widehat{\mathbf{X}}_s$ obtained by MDU and $\mathbf{X}_s$, related to the device) must be translated to align their centroid ($\widehat{\mathbf{X}}_{s,c}$ and $\mathbf{X}_{s,c}$) with the origin of the coordinate system, providing the sets $\widehat{\mathbf{X}}_{s,0}$ and $\mathbf{X}_{s,0}$.

$$\widehat{\mathbf{X}}_{s,c} = \frac{1}{n_s} \sum_{i=1}^{n_s} \widehat{\mathbf{x}}_{s,i}, \quad \mathbf{X}_{s,c} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{x}_{s,i}, \tag{9}$$

$$\widehat{\mathbf{X}}_{s,0} = \widehat{\mathbf{X}}_s - \widehat{\mathbf{X}}_{s,c}\mathbf{1}^T = \widehat{\mathbf{X}}_s \mathbf{J}(n_s), \quad \text{and} \quad \mathbf{X}_{s,0} = \mathbf{X}_s \mathbf{J}(n_s). \tag{10}$$

The Kabsch algorithm is then based on a singular-value decomposition (SVD) of the cross-covariance matrix $\widehat{\mathbf{X}}_{s,0}\mathbf{X}_{s,0}^T$:

$$\mathbf{U}\lceil \mathbf{S} \rfloor \mathbf{V}^T = \widehat{\mathbf{X}}_{s,0}\mathbf{X}_{s,0}^T, \tag{11}$$

where $\lceil \bullet \rfloor$ denotes a diagonal matrix. The optimal rotation matrix reads:

$$\mathbf{R} = \mathbf{V} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d_c \end{pmatrix} \mathbf{U}^T, \quad d_c \in \{-1, 1\}. \tag{12}$$

The translation vector is then $\mathbf{T} = \mathbf{X}_{s,c} - \mathbf{R}\widehat{\mathbf{X}}_{s,c}$. If $d_c = 1$, the two matrices $\widehat{\mathbf{X}}_s\mathbf{R} + \mathbf{T}$ and $\mathbf{X}_s$ are identical whereas one of the matrices is the reflection of the other if $d_c = -1$.

The Kabsch algorithm is carried out to find the optimal rotation matrix between the position of anchors obtained by the MDU and their exact position in the device coordinate system. As explained in Section 2.1.2, four sources are set on a structural support, whose fulcrum is sequentially in contact with at least four prominent locations of the device, so that the position of the fulcrum in the reference frame of the device is determined. A minimum number of contact points is required to solve the symmetry issues involved in the EDM. Indeed, in order to represent the absolute orientation of an object in a unique three-dimensional space, this one must be at least formed by four points. Consider a three-dimensional space and a collection of points randomly set in an arbitrary coordinate system. If the position of one of these points (defined as anchors) is known in another reference frame, the solutions of the Kabsch algorithm consist in an infinite number of rotation matrices. In this case, the possible point sets form a sphere whose center is the anchor. If two points are defined as anchors, the possible points sets form a cylinder whose symmetric axis is defined by the segment between the two anchors. Three anchors then constitute a symmetric plane and each one of the two possible points set is the reflection of the other one. Four distinct anchors allow the determination of the true orientation without ambiguity.

The MDU allows the reconstruction of the point set from distances. However, instead of the sources, the fulcrums are not considered in the EDM. Thus, their positions are not included in the reconstructed point set. The first step of the orthogonal Procustes analysis is then to determine their coordinates in the reference frame provided by the MDU. In that way, the Kabsch algorithm is performed for each support, considering the sources as anchors. The sources must be at least four, as explained in the previous paragraph. The positions of the sources and the fulcrum are known in the coordinate system of the support (Fig. 2-a). The Kabsch algorithm is then applied to find the optimal rotation and translation matrices that best map the sources coordinates in the support system onto those in the MDU reference frame. The position of each fulcrum in the MDU coordinate system is then determined.

Once more, the Kabsch algorithm is carried out to align the coordinates of the whole point set in the device reference frame. The fulcrums are defined as anchors, as their coordinates are known both in the MDU and in the device coordinate systems. Finally, the Kabsch algorithm can be performed a last time, taking the position of the microphones in the array and device coordinate systems as inputs. This provides directly the rigid transformation between the coordinate systems of the array and the device.

3. Numerical validation

3.1. Configuration

The numerical configuration is illustrated in Fig. 1. The antenna is a 50 cm diameter plane array composed of 36 microphones.

Those are arranged in three rings of 7, 11 and 18 microphones (Fig. 5), whose center is located at the origin of the coordinate system ($x = y = 0$). The array is facing towards a simplified full-scale model of a car engine with dimensions of 470 mm width, 380 mm depth and 500 mm height. The change of basis from the array to the device coordinate systems is represented by rigid transformation matrices, including a rotation/reflection ($\mathbf{R}$) matrix and a translation vector ($\mathbf{T}$). In this simulation, these matrices are:

$$\mathbf{T} = [-0.25; 0.25; 0.5](\mathrm{m}), \quad \mathbf{R} = \mathbf{R}_z\mathbf{R}_y\mathbf{R}_x, \tag{13}$$

$$\mathbf{R}_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \quad \mathbf{R}_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix},$$

$$\mathbf{R}_z(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

with the Euler angles $\alpha = 5\pi/3, \beta = 0, \gamma = 0$. The geometry of the structural support of the sources is illustrated in Fig. 2-a. The four acoustic sources form a trirectangular tetrahedron (a cube corner) whose edges are 5 cm long. The distance between the source at the center of the support and its fulcrum is equal to 5 cm too. Four supports are placed at the corners of the device in such a way that they form a three-dimensional space, as illustrated in Fig. 1.

First, the Euclidean distance matrices of the array and the acoustic sources placed on one structural support are computed, as shown in Fig. 6. The top-left diagonal block submatrix represents distances between microphones. The figure clearly reveals three fringes which are directly related to the three rings of microphones. The EDM of the sources is composed of four diagonal block submatrices while off-diagonal submatrices remain unknown. The diagonal blocks are the repeted EDM of the four sources placed on one strucural support. The distances between microphones and sources are not determined yet, then the off-diagonal submatrices of the global EDM are left blank.

3.2. EDM completion

As seen in Section 2.1.2, the off-diagonal block submatrices are obtained by cross-correlating the microphones signals with sound emitted by the synchronized reference sources. In the numerical simulation, the same linear chirp is sequentially emitted by each source. The frequency increases from 0 to 10 kHz. Three sampling

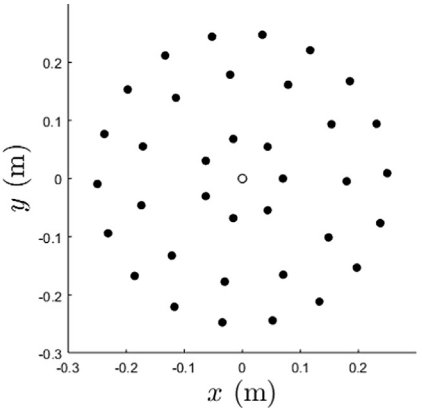


Fig. 5. Arrangement of the plane array centered at the origin of its coordinate system (white circle). The 36 microphones (black dots) are gathered in three rings of 7, 11 and 18 elements.

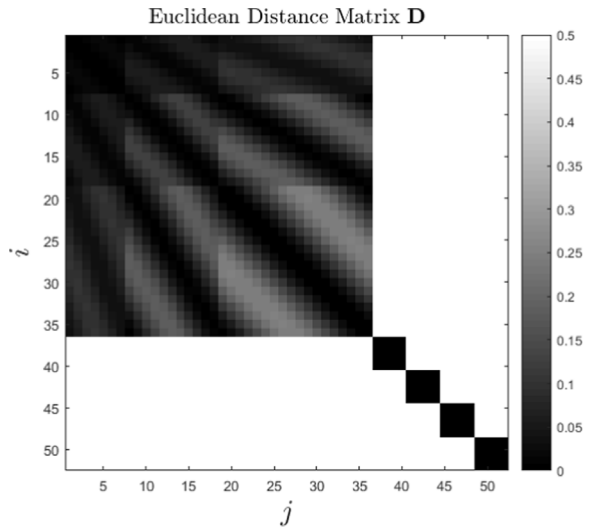


Fig. 6. Overall Euclidean distance matrix of the initial configuration. The particular EDM of the array and the acoustic sources are displayed as diagonal block submatrices.

frequencies ($f_s \in \{25, 50, 100\}$ kHz) are considered to quantify its influence on overall results. The speed of sound is set at 343 m/s. The reconstruction error Δd on the distance between sources and microphones directly depends on the sampling frequency. Indeed, the temporal position of the maximum cross-correlation peak depends on the length of a sample. The Table (1) summarizes the reconstruction errors on the distance between sources and microphones due to this sampling issue:

Finally, Fig. 7-a illustrates the obtained EDM which remains uncomplete and becomes noisy.

The completion algorithm called Alternating coordinate descent method (described in Section 2.2.1) is applied to this particular matrix, denoted $\tilde{\mathbf{D}}$. This method allows the determination of both a complete EDM $\widehat{\mathbf{D}}$, close to the observed EDM $\tilde{\mathbf{D}}$, and its corresponding estimated point set $\widehat{\mathbf{X}}$. Fig. 7-b represents the completed Euclidean distance matrix $\widehat{\mathbf{D}}$.

Two reconstruction errors can be considered, based on the discrepancies between the estimated EDM $\widehat{\mathbf{D}}$ and the observed EDM $\tilde{\mathbf{D}}$ on the one hand, and between $\widehat{\mathbf{D}}$ and the ground truth EDM $\mathbf{D}$ on the other hand. These two indicators are denoted $\tilde{\varepsilon}$ and ε and read:

$$\tilde{\varepsilon} = \frac{\|\mathrm{edm}\left(\widehat{\mathbf{X}}\right) - \tilde{\mathbf{D}}\|_F}{\|\tilde{\mathbf{D}}\|_F}, \quad \varepsilon = \frac{\|\mathrm{edm}\left(\widehat{\mathbf{X}}\right) - \mathbf{D}\|_F}{\|\mathbf{D}\|_F}. \tag{14}$$

The reconstruction errors $\tilde{\varepsilon}$ and ε are calculated for different sampling frequencies ($f_s \in \{25, 50, 100\}$ kHz) and gathered in the Table (2).

These indicators clearly show the weight of the sampling frequency in the accuracy of the method. The distance measurements between sources and microphones is therefore a crucial step in the global methodology. Fig. 8 shows a multidimensional unfolding of the numerical set-up. As it can be seen, the position and orientation of the overall geometric configuration are completely arbitrary, and more, the MDU result is actually a reflection of the

Table 1
Reconstruction error Δd for different sampling frequencies f_s.

f_s	25 kHz	50 kHz	100 kHz
Δd ($\times 10^{-3}$ m)	± 7	± 3.5	± 1.7

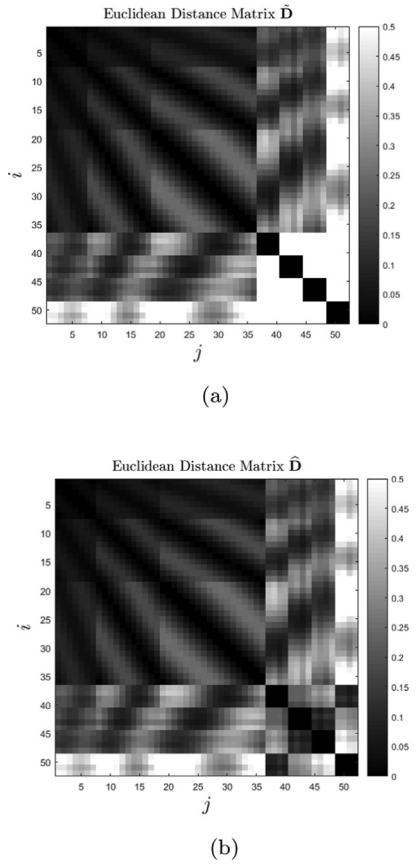


Fig. 7. Global Euclidean distance matrices. (a): Uncompleted EDM obtained after the source-microphone distances estimations. (b): Completed EDM obtained as a result of the alternating coordinate descent method.

Table 2
Reconstruction errors $\tilde{\varepsilon}$ and ε for different sampling frequencies f_s.

f_s	25 kHz	50 kHz	100 kHz
$\tilde{\varepsilon}$ (%)	1.15	0.56	0.35
ε (%)	0.59	0.35	0.16

true set-up. The antenna (dots) and the acoustic sources (squares) are clustered in subsets.

Each group of four acoustic sources corresponding to a particular location of the device are gathered in a specific subset. As the EDM refers to distances between points in a network, it involves geometric informations about the acoustic sources and microphones. This configuration does not include information about the points of contact between the structural support and the device. In order to align the point set obtained with the MDU algorithm with the ground truth configuration, the orthogonal Procustes analysis is carried in a two-step approach, as explained in Section 2.3. First, the coordinates of the fulcrum of each structural support has to be determined in the MDU coordinate system with the Kabsch algorithm.

Fig. 9 illustrates the alignment of each structural support (squares) with the acoustic sources positions found with the MDU technique (filled squares). The fulcrum of each support is represented by isolated squares. Then, these fulcrums become anchors, as their coordinates are known both in the MDU and in the device coordinate systems. The Kabsch algorithm is once more applied to find the coordinates of the point set in the device coordinate system. Fig. 10 shows the final result of the overall methodology, from a qualitatively standpoint. The filled markers represent

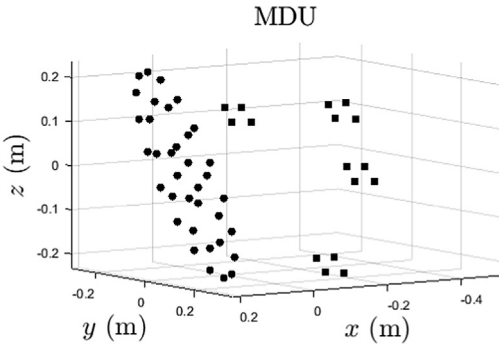


Fig. 8. Multidimensional unfolding of the numerical set-up in a arbitrary position and orientation. The dots depict the microphone array while the squares represent the acoustic sources.

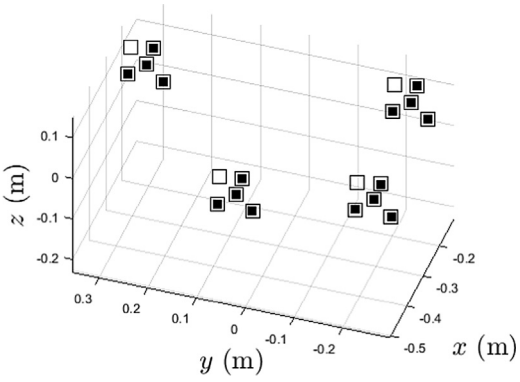


Fig. 9. Estimation of the fulcrum's position of each structural support (isolated squares).

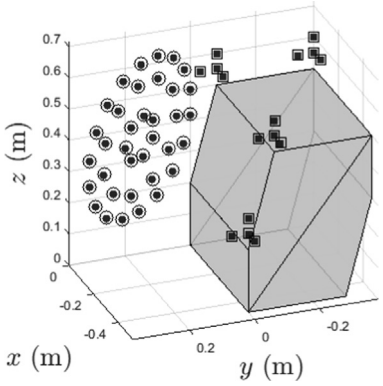


Fig. 10. Final result of the overall methodology. The ground truth positions of the microphone array and the acoustic sources are represented by the dots and filled squares, respectively, while the circles and hollow squares embody their estimated positions.

the ground truth position of the point set while the circles denotes the microphones and the squares the acoustic sources.

3.3. Parametric studies

As the overall methodology is qualitatively evaluated, two parametric studies are detailed in this section to assess the robustness of the positioning technique to sampling issues and to the number of measurement points required to perform an accurate array localization.

3.3.1. Impact of the sampling frequency on the reconstructed geometry

The agreement between the ground truth and the estimated positions of both sources and microphones is studied for the three sampling frequencies $f_s \in \{25, 50, 100\}$ kHz. A criterion based on the distance between each couple of points is defined as follows:

$$\mathbf{\Delta}_{a,s} = \|\widehat{\mathbf{X}}_{a,s} - \mathbf{X}_{a,s}\|_2^2. \tag{15}$$

Fig. 11 show the values of the criterion for each source ($\mathbf{\Delta}_s$) and for each microphone ($\mathbf{\Delta}_a$). At a fixed sampling frequency, it can be seen that the errors vary according to the reconstruction error Δd, from one point of the set to another. Therefore, the range of errors on the criterion $\mathbf{\Delta}_{a,s}$ and the magnitude of the criterion itself decrease as the sampling frequency increases. For any sampling frequency, the maximum positioning error is reached by a source, which expresses the lack of information due to the support disorientation. Concerning the position of the microphones, the maximum positioning error are 0.43 cm at $f_s = 25$ kHz, 0.36 cm at $f_s = 50$ kHz, and 0.20 cm at $f_s = 100$ kHz, which clearly shows the sampling influence.

3.3.2. Number of structural support

As detailed in Section 2.3, the orthogonal Procustes analysis needs at least four anchors to completely solve the positioning problem in a three-dimensional configuration. However, adding more measurement points also increases the number of unknowns in the completion problem and the number of noisy source-microphone distances to estimate. Fig. 12 displays a configuration with only three structural supports placed on the device. Both pictures (a) and (b) are valid solutions of the positioning problem and

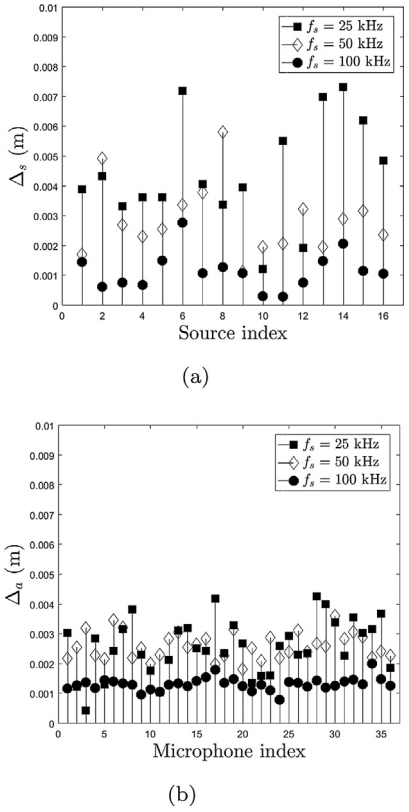


Fig. 11. Distances Δ (m) between the estimated and the ground truth positions of each point of the set, according to the sampling frequency ($f \in \{25, 50, 100\}$kHz). (a): Value of the criterion $\mathbf{\Delta}_s$ for each source. (b): Value of the criterion $\mathbf{\Delta}_a$ for each microphone.

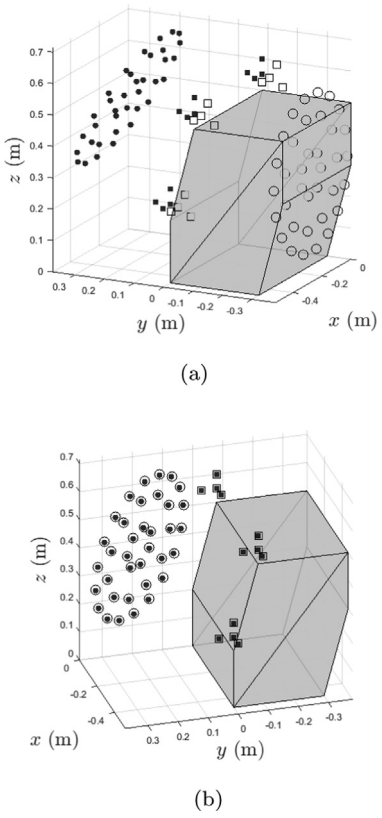
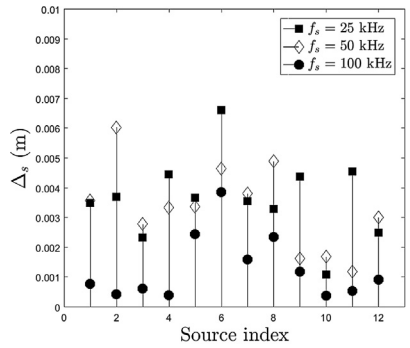


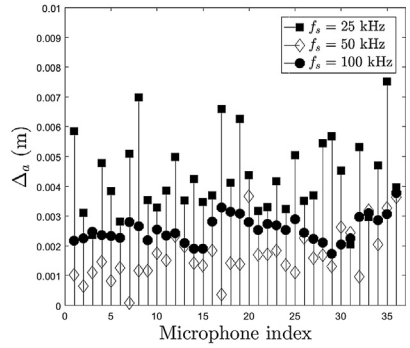
Fig. 12. Valid solutions of the positioning problem with three measurement points. (a): Result with an arbitrary value $d_{c_a} \in \{-1, 1\}$, where the obtained array position is a mirror solution. (b): Result with the opposite value $d_{c_b} = -d_{c_a}$, close to the ground truth position of the array.

are equivalent in terms of their Euclidean distance matrices. The fulcrums of the three structural supports form a plane of symmetry which create an equivocal situation.

As expressed by Eq. (12), the value of $d_c \in \{-1, 1\}$ allows the distinction of one solution from the other one. It is therefore possible to compute the solution in a two-step process. First, the whole methodology is carried out. If the obtained result is a mirror solution, the positioning algorithm is computed a second time with the opposite value of d_c. Fig. 13 displays the values of the residual criterion for each source and microphone at different sampling frequencies. For the microphones position, the maximum values of the criterion are 0.75 cm at $f_s = 25$ kHz, 0.36 cm at $f_s = 50$ kHz, and 0.38 cm at $f_s = 100$ kHz. From a qualitative standpoint, the discrepancies become significant at the sampling frequency $f_s = 25$ kHz as the efficiency of the modified method declines. In this situation, the maximum positiong error increases from the four measurement points configuration by 0.3 cm while the mean value increases from 0.36 to 0.43 cm. At $f_s = 50$ kHz, the maximum positiong error remains the same and the mean value decreases from 0.26 cm to 0.17 cm. For a sampling frequency $f_s = 100$ kHz, the maximum positioning error has nearly doubled while its mean value increases from 0.13 cm to 0.26 cm. This shows that a low sampling frequency brings significant discrepancies. Besides, the most accurate solution is found when four measurement points are considered and when the highest sampling frequency is used. Nevertheless, a three anchor configuration is found enough to solve the positioning problem and reduces the number of measurements points, while it slightly modifies the accuracy of the method by about a millimeter.

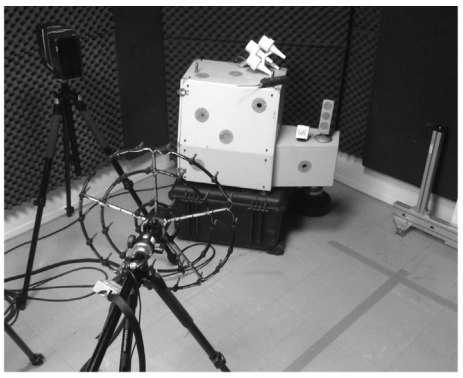


(a)

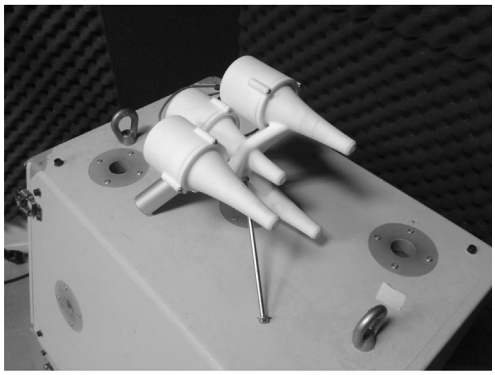


(b)

Fig. 13. Distances Δ (m) between the estimated and the ground truth positions of each point of the set, according to the sampling frequency ($f \in \{25, 50, 100\}$kHz), for a three measurement point configuration. (a): Value of the criterion $\mathbf{\Delta}_s$ for each source. (b): Value of the criterion $\mathbf{\Delta}_a$ for each microphone.



(a)



(b)

Fig. 14. Experimental test-case configuration. (a): A 36-microphone plane array is facing a full-scale car engine mock-up. A 3D optical scanner is set on the left to survey both the positions of the antenna and the device. (b): Structural support, gathering the four acoustic sources, a handle and a rod whose extremity form a fulcrum in contact with prominent locations of the device.

4. Experimental validation

As the numerical computations established the accuracy of the proposed algorithm, an experimental validation is therefore performed to assess its robustness in an industrial application.

4.1. Test-case configuration

Fig. 14-a displays the general test-case configuration. A full-scale mock-up of a car engine is employed to embody a radiating object, of the same dimensions as in the numerical parametric studies. Similarly, the same 36-microphone array (MicrodB HDcam 36) is facing towards its front face. A 3D optical scanner (Faro Focus^M 70) is set in order to survey both the positions of the mock-up and the antenna and whose measurements are considered as the reference of the experiment. The distance accuracy of the scanner is up to ± 3 mm. Fig. 14-b shows a prototype of structural support, gathering in one object four acoustic sources, a fulcrum and a handle.

Acoustic loudspeakers are embedded in a three-dimensional printed piece with specific horns to adapt the impedance between their diaphragm and the propagation medium as well as to reduce the radius of the physical acoustic sources. The four sources are located in such a way that they form a regular tetrahedron. As the position of the loudspeaker and the one considered as the source location at the extremity of the horn do not coincide, an additional delay occurs in the source-microphone time of flight and must be assessed. A geometrical calibration of the horn is then performed by comparing the distance d between the edge of the horn and one particular microphone, measured with a mechanichal calibration bench and considered as a ground truth, and the distance d_{corr} related to the time delay estimated by cross-correlating the microphone signal with the one emitted by the loudspeaker. The calibration distance $d_c = d_{\text{corr}} - d$ includes the acoustic propagation through the horn as well as electronic delays brougth by the whole set-up. Divided by the speed of sound, it therefore become a time delay to substract to the generic cross-correlation signal. The calibration microphone is located on the central source axis and is successively shifted from the edge of the horn, from 0.1 to 0.9 m. For each distance, the measurement is repeated five times. Fig. 15 reports the mean distances associated to the estimated time delays according to the selected gaps between the edge of the horn and the calibration microphone. Although the calibration distance slightly increases with the distance, the mean values spread over only one millimeter with a standard deviation lesser than the millimeter. A classical increase of the discrepancies with the distance can be related to a miscalculation of the speed of sound. As explained before, the mean calibration distance ($d_c = 0.1092$ m) does not perfectly equal the length of the horn (0.102 m), because of time delays brought by the electronic instrumentation. Finally, a calibration delay $t_c = 3.18 \times 10^{-4}$ s must be subtracted to all cross-correlation signals to recover the distance between the edge of a horn and a microphone.

The fulcrum of the structural support is placed at three prominent locations of the mock-up (some corners), without knowing its true orientation. All the acoustic sources emit at once one particular uncorrelated broadband noise, allowing the distance estimation by cross-correlating the microphones signals with each referenced

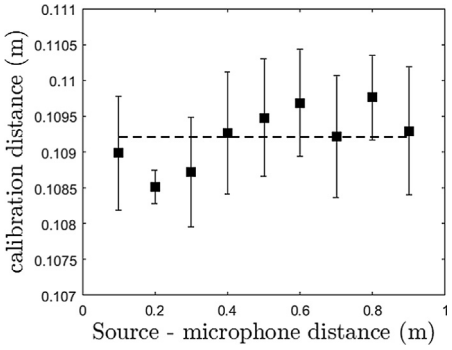


Fig. 15. Mean calibration distance according to the selected gaps between the edge of the horn of the central source and the calibration microphone (squares). Error bars represent the standard deviations based on five measurements for each gap. The dashed line embodies the overall mean calibration distance, as selected in the algorithm.

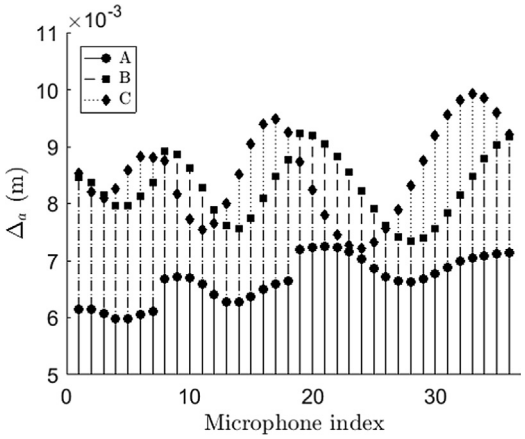


Fig. 16. Distances Δ_a (m) between the estimated and the reference positions of each microphone of the array, according to the configuration (A (dots): 20 cm, B (squares): 50 cm, C (diamonds): 100 cm).

noise. The sampling frequency is set at 102400 Hz to reduce its influence on the positioning errors. A special care should be taken to select the first peak of the cross-correlation function, representative of the direct path between the source and the microphone. In the designed algorithm, the first peak might not be the maximum of the cross-correlation function, due to strong reflection of the mock-up. Besides, if this function is too noisy, the first detected peak and the direct path can be unrelated.

To deal with this, the maximum of the cross-correlating function is first evaluated. A threshold is defined to select all the peaks which are related to the direct path or the possible reflections and remove the noisy fluctuations of the function. This threshold is half of the maximum in the present study. Finally, the sample corresponding to the earliest of this peak selection is considered to assess the distance between the source and the microphone. The whole algorithm is included in an executable Matlab code, working on any laptop computer. Its calculation time depends of the sampling frequency and the number of microphone channels and ranges from a few secondes to less than half a minute.

4.2. Positioning results

Three positions have been tested, representative of typical configurations encountered in acoustic imaging. A first position (configuration A) is set with an array 20 cm away from the mock-up to assess the robustness of the method to the near-field effects such as the diffraction of the object itself and the radiating pattern of the sources. Two other positions at 50 and 100 cm (configurations B and C) are defined to evaluate the possible deviations due to the ground reflections and the inaccuracy of the speed of sound, estimated according to Cramer [17] with a measurement of both the temperature and the humidity rate in the room. As in the numerical studies, a positioning error $\mathbf{\Delta}_a$ is defined as the distance between the microphone positions as found by the algorithm and by the optical scanner, considered as the reference. More precisely, the algorithm outputs are a rotation matrix and a translation vector. The microphone positions are then the rigid transformation of the known positions of the microphones in the array coordinate system, calculated with the output matrices. To prevent a misestimation of the microphone positions with the optical scanner, the Kabsch algorithm is performed to find the rigid transformation which corresponds at most to this obtained array position. In both cases, the arrays of microphones are identical to the original one, but have been modified by different rigid transformations. An estimation of the antenna positioning errors is displayed in Fig. 16, representing the three configurations (dot markers: 20 cm, square markers: 50 cm and diamond markers: 100 cm). The characteristic positioning errors follow two distinguishable patterns. At 50 and 100 cm, the pattern expresses an increasing periodic curve whose mean equals 8.5 mm in both cases. The maximum errors are 9 mm and 10 mm for the B and C configurations, respectively, and they do not correspond to the same microphone. For the configuration A, the positioning errors can be gathered in three groups, corresponding to the first six microphones, the following eleven and the last eighteen ones. These collections are clearly related to the three rings of microphones the array is made of. In this case, the maximum positioning error does not exceed 7.5 mm. Those two patterns can be explained with the analysis of the rotation and translation errors, defined as the differences between the rigid transformations determined with the proposed positioning method and with the optical scanner (Tables (3) and (4)).

The translation errors of the rigid transformation range from 0.7 mm to 8 mm. In configuration A, the errors are maximum and similar in the x and y directions (-4.11 mm and -4.21 mm, respectively), in the plane of the array. For the B and the C configurations, the maximum errors occur in the x direction (-7.70 mm) and the y direction (-6.65 mm), respectively. As shown in Table (4), the angular discrepancies stay below a half degree, except for the angular error on the z-axis in the configuration A (0.71 degree). The latter dominant value explains the two patterns in the antenna

Table 3
Discrepancies between the translation vector of the estimated and the reference positions of the array (m).

	Translation errors (m)		
Configuration	ε_x	ε_y	ε_z
A	−4.11e-03	−4.21e-03	−1.07e−03
B	−7.06e-04	−6.65e-03	4.73e-03
C	−7.70e-03	−3.26e-03	1.31e-03

Table 4
Discrepancies between the Euler angles of the estimated and the reference positions of the array (degrees).

	Angular errors (degrees)		
Configuration	$\Delta\Theta_x$	$\Delta\Theta_y$	$\Delta\Theta_z$
A	−1.19e-01	−3.48e-01	7.13e-01
B	2.40e-02	−1.25e-01	−3.40e-01
C	2.49e-01	−3.54e-01	−2.57e-02

positioning errors in Fig. 16. This underlying analysis is illustrated in Fig. 17.

Fig. 17-a represents the configurations B and C, as a typical disorientation of the array. The diagram amplifies the real discrepancies and simplifies the positioning errors such as only a translation error in the z-direction and an angular error according to the x axis occur. In this case, the combination of the translation and the angular deviations reduce the errors on one side of the array while they are increased on the other side. In this way, the deviations from the reference indeed show a increasing oscillating function according to the index of the microphones and to the radius of the rings. The configuration A is illustrated by the Fig. 17-b. Here, the angular deviation on the z-axis, normal to the array plane, dominates the others. A rotation of the array around its axis of symmetry occurs in such a way that the positioning errors of each microphone depend mostly of their ring. On the Fig. 16, the discrepancies are clearly gathered in three groups, representing the three sets of microphones. The similar translation errors in the x and y directions add a mild oscillating form to the deviations.

Fig. 16 shows that the maximum positioning error stays below one centimeter, which assesses the reliability of the proposed algorithm. The mean discrepancies remain equivalent for both the configurations B and C, implying an accurate speed of sound estimation. Among the other sources of uncertainty, the main parameters are the distance accuracy of the optical scanner, the estimation of the calibration distance of the sources and the ability to make the measurement points of the physical object and its representation in the numerical mesh coincide. A special care must then be taken during the experimental campaign to carry out the measurement, since the numerical studies show that a very small part of the discrepancies comes from the algorithm itself.

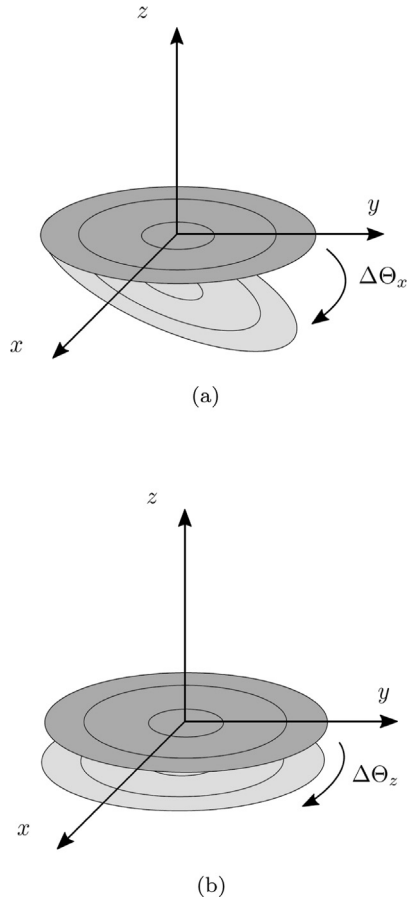


Fig. 17. Physical interpretation of the two main patterns encountered in the array positioning problem. Dark grey ellipses depict the three rings of the reference array while light gray ellipses represent the array obtained with the proposed algorithm. (a): The combination of a translation in the z-direction and a rotation around the x-axis reduces the deviations on one side of the array while they are increased on the other side. (b): A dominating rotation around the z-axis, normal to the array plane, implies that the positioning discrepancies of each microphone depend mostly of their ring.

5. Conclusion

The accuracy of acoustic retro-propagating methods are directly related to the precise estimation of the microphone array arrangement and its position and orientation in the experimental environment, in relation to the radiating device. Thus, this application-oriented paper aimed to propose an array localization methodology for acoustic imaging, gathering a collection of techniques based on the Euclidean distance geometry. The antenna positioning problem is tackled with the help of the Euclidean distance matrix, which reports on the complete geometric configuration of a set-up. This set-up is here composed of the microphone array, some prominent locations of the device plus acoustic sources employed to link them together. Each feature geometry is initially known in its own coordinate system and this paper detailed a practical tool to collect each microphone coordinates in the device mesh coordinate system to perform accurately the retro-propagation method onto it. The complete theoretical methodology is first presented to detail the underlying principles of the algorithms involved and the benefits of their application to find the optimal rotation and translation matrices to align the array with its ground truth position. A numerical validation is conducted to assess the proposed algorithm in a typical experimental configuration encountered in acoustic imaging. As the distance between acoustic sources and microphones is determined by cross-correlating the received signals with the referenced ones, the sampling frequency is found as a decisive parameter of the method, the higher being the more accurate. Also, the methodology requires at least three measurement points to achieve the antenna positioning problem. With a 100 kHz sampling frequency, the reconstruction error of the antenna brought by the algorithm itself is 0.35 %, and the distances between the obtained and the ground truth positions of the microphones stay below 3 mm. Finally, an experimental validation with an 3D optical scanner has been performed to test the robustness of the method in an industrial application. Three configurations have been selected with an array 20, 50 and 100 cm away from the device and the largest discrepancy equals one centimeter, which illustrates that experimental misestimations greatly affect the accuracy of the method. The most decisives factors are the determination of the in situ speed of sound, the calibration distance of the acoustic sources and the coincidence of the measurement points and their representation in the numerical mesh.

Instead of 3D optical scanners, the proposed methodology forms an industrial tool which can be quickly deployed in an experimental campaign with space constraints. It automatically exploits the temporal signals measured by the microphones to find directly their position toward the device, instead of extracting manually the orientation of the array in the mesh one would obtain with an optical scanner, also dependant on its refinement. At last, from an industrial point of view, the use of the microphones themselves to find their own position limits the need of both external hardware and software, and drastically reduces the costs.

Beyond its simplicity, one of the main advantages of the proposed method is that it can benefit from extensions provided by the EDM literature to tackle many issues faced in domains of source localization and acoustic array processing, such as array cal-

ibration in wind tunnel, multiple array localization or the measurement of experimental transfer functions between the sources and the microphones to assess the diffraction due to the device.

CRediT authorship contribution statement

Simon Bouley: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. **Charles Vanwynsberghe:** Conceptualization, Writing - review & editing. **Thibaut Le Magueresse:** Project administration. **Jérôme Antoni:** Supervision, Writing - review & editing. **Allan Outrequin:** Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study has been produced in the framework of LUG2 supported by Région Auvergne Rhône-Alpes and BPIFrance (FUI22). It was performed within the framework of the Labex CeLyA of Université de Lyon, operated by the French National Research Agency (ANR-10-LABX-0060/ANR-11-IDEX-0007).

References

[1] Gilquin L, Bouley S, Antoni J, Le Magueresse T, Marteau C. Sensitivity analysis of two inverse methods: Conventional beamforming and Bayesian focusing. J Sound Vib 2019;455:188–202.
[2] Plinge A, Jacob F, Haeb-Umbach R, Fink GA. Acoustic microphone geometry calibration: an overview and experimental evaluation of state-of-the-art algorithms. IEEE Signal Process Mag 2016;33(4):14–29.
[3] Wang L, Hon T-K, Reiss JD, Cavallaro A. Self-localization of ad-hoc arrays using time difference of arrivals. IEEE Trans Signal Process 2016;64(4):1018–33.
[4] Valente SD, Tagliasacchi M, Antonacci F, Bestagini P, Sarti A, Tubaro S. Geometric calibration of distributed microphone arrays from acoustic source correspondences. Multimedia Signal Processing (MMSP), 2010 IEEE International Workshop on. IEEE; 2010. p. 13–8.
[5] Gower JC. Properties of Euclidean and non-Euclidean distance matrices. Linear Algebra Appl 1985;67.
[6] Parhizkar R. Euclidean distance matrices: Properties, algorithms and applications, Ph.D. thesis. École Polytechnique Fédérale de Lausanne (EPFL) 2013.
[7] Torgerson WS. Multidimensional scaling: I. Theory and method. Psychometrika 1952;17(4):401–19.
[8] Borg I, Groenen PJ. Modern multidimensional scaling: theory and applications. 2nd Edition. New York: Springer-Verlag; 2005.
[9] Dokmanić I, Parhizkar R, Ranieri J, Vetterli M. Euclidean distance matrices: essential theory, algorithms, and applications. IEEE Signal Process Mag 2015;32(6):12–30.
[10] Gower JC. Euclidean distance geometry. Math Sci 1982;7(1):1–14.
[11] Schönemann PH. On metric multidimensional unfolding. Psychometrika 1970;35(3):349–66.
[12] Crocco M, Del Bue A, Murino V. A bilinear approach to the position self-calibration of multiple sensors. IEEE Trans Signal Process 2012;60(2):660–73.
[13] Takane Y, Young FW, De Leeuw J. Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. Psychometrika 1977;42(1):7–67.
[14] Gower JC. Procrustes methods. Wiley Interdisciplinary Rev: Comput Stat 2010;2(4):503–8.
[15] Kabsch W. A solution for the best rotation to relate two sets of vectors. Acta Crystallographica Section A 1976;32(5):922–3.
[16] Kabsch W. A discussion of the solution for the best rotation to relate two sets of vectors. Acta Crystallographica Section A 1978;34(5):827–8.
[17] Cramer O. The variation of the specific heat ratio and the speed of sound in air with temperature, pressure, humidity, and CO_2 concentration. J Acoust Soc Am 1993;93(5):2510–6.