24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

A NOVEL TWO-STAGE SOUND FIELD IMAGING TECHNIQUE USING A FARFIELD RANDOM ARRAY

Mingsian R. Bai

University of National Tsing Hua, School of Power Mechanical Engineering, Hsinchu, Taiwan
email: msbai63@gmail.com

You Siang Chen, Yi-Yang Lo, and Chung Chun

University of National Tsing Hua, School of Power Mechanical Engineering, Hsinchu, Taiwan

In this paper, a random microphone array with optimized layout is developed for locating and separate noise sources. The sensor locations are determined, with the aid of simulated annealing method, to minimize the maximum sidelobe level of the farfield beampattern. A two-stage algorithm is formulated on the basis of the spherical-wave model. In the localization stage, the active source regions are located by using the frequency-domain delay-and-sum (DAS) algorithm. In the separation stage, source signals are extracted by solving an inverse problem, based on the source positions identified in the stage one. Equivalent sources are assumed to be fewer (overdetermined) than or more (underdetermined) than microphones. Tikhonov regularization is employed to solve the inverse problem and calculate the acoustic variables including sound pressure, particle velocity, sound intensity, and sound power based on the equivalent source model (ESM). Numerical and experimental results are presented.
Keywords: sound field imaging, farfield, random microphone array

1. Introduction

Noise pollution is increasingly concerned since it always influences our living quality such as industrial noise, environmental noise, and product noise, etc. Noise source identification (NSI) is the key techniques to find the locations of noise sources and separate them in order to aim at noise sources for reducing sound pressures. Array-based NSI algorithms generally fall into two categories: farfield arrays and nearfield arrays. Using nearfield acoustic holography must be deployed in the close neighborhood of the interest source. However, the nearfield of source is inaccessible for some application scenarios such as machine tools. As opposed to the nearfield arrays, farfield array that enables larger frequency bandwidth and need not be positioned in close proximity is well suited for locating large and distant objects, but it only gives relative distribution of source field.

In this paper, a two-stage algorithm is proposed by combing nearfield and farfield concepts to solve the aforementioned limitation. In the localization stage, the active source regions are located by using the frequency-domain delay-and-sum (DAS) [1-2] algorithm based on the farfield concept. In the separation stage, source signals are extracted using the equivalent source method (ESM) by Tikhonov regularization (TIKR) to solve an inverse problem and calculate the acoustic variables including sound pressure, particle velocity, sound intensity, and sound power.

2. Stage 1: source localization

A farfield array model is given in this section. Assume that sound sources are located at the farfield of the array, so that sources can be regarded as point sources. The pressure vector $\mathbf{p}(\omega)$ due to N point sources received at M microphones can be written in the frequency domain as:

$$\mathbf{p}(\omega) = \mathbf{A}\mathbf{s}(\omega) + \mathbf{n}(\omega), \tag{1}$$

where the matrix $\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_N]$ being the steering matrix and $\mathbf{a}_n = \left[e^{-jk\cdot r_{1N}} / r_{1n} \quad e^{-jk\cdot r_{2N}} / r_{2n} \quad \cdots \quad e^{-jk\cdot r_{MN}} / r_{Mn} \right]^T, \quad n = 1, \cdots, N$, denotes the steering vector of the *n*th source, with the superscript "*T*" denoting matrix transpose. Time dependence $e^{j\omega t}$ is assumed, with ω and t denoting angular frequency and time, and $j = \sqrt{-1}$. The distance, $r_{mn} = \left\| \mathbf{r}_n - \mathbf{r}_m \right\|$, with $\mathbf{r}_n$ and $\mathbf{r}_m$ being the position vectors of the *n*th source point and the *m*th field point. The source amplitude vector $\mathbf{s}(\omega) = [s_1(\omega) \cdots s_N(\omega)]$. The vector $\mathbf{n}(\omega)$ represents the additive noise. With this model, the array output can be written as:

$$y(\omega) = \mathbf{w}^H(\omega)\mathbf{p}(\omega), \tag{2}$$

where $\mathbf{w}(\omega) = [w_1(\omega) \quad \cdots \quad w_M(\omega)]^T$ is the weight, with the superscript "*H*" denoting matrix hermitian transpose. The array output, $y(\omega)$, is a complex number. In the figures, however, array output power, $E\left\{ \left| y(\omega) \right|^2 \right\}$, is plotted instead.

In stage 1, DAS algorithm is employed for source localization based on the farfield concept. The weight vector of DAS can be expressed as:

$$\mathbf{w}(\omega) = \frac{\mathbf{a}_n(\omega)}{\mathbf{a}_n^H(\omega)\mathbf{a}_n(\omega)} = \frac{\mathbf{a}_n(\omega)}{M}, \tag{3}$$

where *M* is the microphone number.

3. Stage 2: source signal separation

In stage 2, source signals are extracted by TIKR method to solve an inverse problem and calculate the acoustic variables.

3.1 Tikhonov regularization (TIKR)

For simplicity, the ω is suppressed hereafter. Once the source locations are found in Stage 1, the source signal separation problem can be posed as the following linear system of equations (with the additive noise term omitted for simplicity):

$$\mathbf{p} = \mathbf{A}\mathbf{s}. \tag{4}$$

where **p** is the pressure vector received at the microphones, **s** is the source amplitude vector, and **A** is the steering matrix constructed on the basis of the preceding localization results. Suppose that the number of sources is less than the number of microphones, which implies the separation problem is an overdetermined inverse problem. A common way to solve this problem is the TIKR method that is based on the following least-squares optimization [3]

$$\min_{\mathbf{s}} \left(\left\| \mathbf{A}\mathbf{s} - \mathbf{p} \right\|_2^2 + \beta^2 \left\| \mathbf{s} \right\|_2^2 \right). \tag{5}$$

where $\left\| \square \right\|_2$ denotes vector 2-norm and β is the regularization parameter. The regulated solution of the problem above is [4]

$$\hat{\mathbf{s}} = \left(\mathbf{A}^H\mathbf{A} + \beta^2\mathbf{I} \right)^{-1} \mathbf{A}^H\mathbf{p}. \tag{6}$$

In this paper, the parameter β is selected to be 1% of the maximum singular value of matrix A at low frequency.

3.2 Post processing for acoustic variables

Traditionally, farfield arrays only give relative distribution of source field. However, by using the proposed method, the acoustic variables including sound pressure (p), particle velocity (u), sound intensity (I), and sound power (W) are calculated based on ESM [1]. Important surfaces involved in the imaging process using the random array are depicted in Fig. 1. A sufficient number of equivalent point sources are distributed within a 3 dB beamwidth of the DAS beampattern. It follows that the pressure vector at the microphones $\mathbf{p}$ and the source amplitude vector $\mathbf{s}_v$ can be related by a propagation matrix $\mathbf{G}_v$

$$\mathbf{p} = \mathbf{G}_v \mathbf{s}_v . \tag{7}$$

Similar to Eq. (4), the source amplitude can be calculated using the TIKR method. In general more equivalent sources than microphones are selected, so the following underdetermined least-squares minimum-norm solution applies:

$$\hat{\mathbf{s}}_v = \mathbf{G}_v^H (\mathbf{G}_v \mathbf{G}_v^H)^{-1} \mathbf{p} . \tag{8}$$

In practice, TIKR method defined in Eq. (5) are still applied to calculate the source signal

$$\hat{\mathbf{s}}_v = (\mathbf{G}_v^H \mathbf{G}_v + \beta^2 \mathbf{I})^{-1} \mathbf{G}_v^H \mathbf{p} . \tag{9}$$

It follows that the sound pressure on a plane with a small retracted distance can be reconstructed as

$$\hat{p}_m = \sum_{v=1}^{N_e} \hat{s}_v \cdot \frac{e^{-jkr_{mv}}}{r_{mv}} . \tag{10}$$

where r_{mv} is the distance of the vth equivalent source and the mth field point on the reconstruction surface, i.e., $r_{mv} = |\mathbf{z}_m - \mathbf{z}_v|$, with $\mathbf{z}_m$ and $\mathbf{z}_v$ being the position vectors of the vth equivalent source and the mth field point on the reconstruction surface. N_e is the number of equivalent sources. The normal component of particle velocity can also be obtained

$$\hat{u}_m = \frac{1}{j\rho_0 \omega} \sum_{v=1}^{N_e} (\mathbf{n} \cdot \mathbf{e})(jk + \frac{1}{r_{mv}}) \hat{p}_m . \tag{11}$$

where ρ_0 is the air density, $\mathbf{e} = (\mathbf{z}_m - \mathbf{z}_v)/r_{mv}$ and $\mathbf{n}$ is the unit vector normal to the reconstruction surface. In addition, sound intensity is readily calculated as:

$$I_m = \frac{1}{2} \mathrm{Re}\{\hat{p}_m^* \cdot \hat{u}_m\} . \tag{12}$$

where "*" denote complex conjugation and "Re" denotes the real part. The total time-averaged, acoustic power radiated from a vibrating structure can be expressed in terms of its equivalent source distribution. The radiated sound power can be estimated with two approaches. First, the power is approximated by numerical integration of preceding sound intensities.

$$W = \int_{S_W} I dS \approx \frac{S_W}{N_W} \sum_{m=1}^{N_W} I_m . \tag{13}$$

where S_W is the area covering the equivalent sources, N_W is the number of intensity components, and I_m is the mth normal intensity component. However, this method tends to underestimate the

sound power, as will be seen in the following simulation. Alternatively, an analytical expression for the sound power produced by an array of monopole sources can be used [5].

$$W = \frac{2\pi}{\rho_0 c}\mathbf{s}^H \mathbf{R}\,\mathbf{s}\,, \tag{14}$$

where c is the speed of sound and the radiation resistance matrix $\mathbf{R}$ is defined as

$$\mathbf{R} = \begin{bmatrix} 1 & R(\mathbf{r}_1/\mathbf{r}_2) & \cdots & R(\mathbf{r}_1/\mathbf{r}_{N_e}) \\ R(\mathbf{r}_1/\mathbf{r}_2) & 1 & \cdots & R(\mathbf{r}_2/\mathbf{r}_{N_e}) \\ \vdots & \ddots & \ddots & \vdots \\ R(\mathbf{r}_{N_e}/\mathbf{r}_1) & R(\mathbf{r}_{N_e}/\mathbf{r}_2) & \cdots & 1 \end{bmatrix}, \tag{15}$$

which is real and symmetric.

$R(\mathbf{r}_p/\mathbf{r}_q) = \sin(k|\mathbf{r}_p - \mathbf{r}_q|)/(k|\mathbf{r}_p - \mathbf{r}_q|), \quad p, q = 1, \cdots, N_e$. $R(\mathbf{r}_p/\mathbf{r}_q) = 1$ for $p = q$.

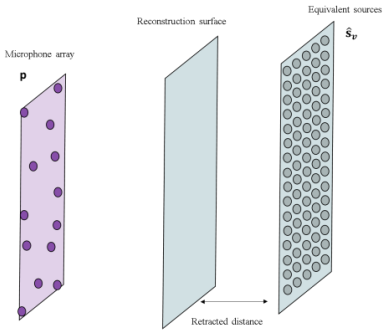


Figure 1: Important surfaces defined for the ESM.

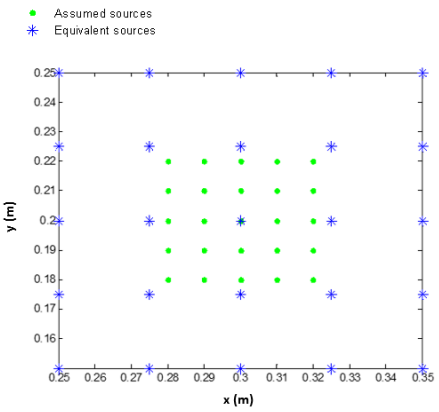


Figure 2: Sound power estimation for the overdetermined layout.

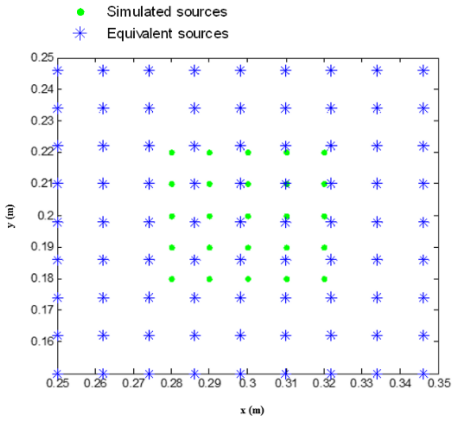


Figure 3: Sound power estimation for the underdetermined layout.

4. Numerical and experimental validations

In the section, DAS method is utilized in the localization stage. In the separation stage, the equivalent sources are distributed within the 3 dB beamwidth of the DAS beampattern. TIKR method is employed in solving the inverse problems to calculate acoustic variables.

4.1 Verification of sound power estimation

An example is given to demonstrate power estimation of a 5×5 matrix of 4 kHz point sources with designated amplitudes, as shown in Fig. 4. The sources are equally spaced with 1 centimeter and centered at (0.3 m, 0.2 m, 1 m). The microphone array is positioned 1m away from the sources, where microphones are indicated by squares. The sound power calculated with Eq. (16), W_t =7.04 Watts, serves as the theoretical value. TIKR method is employed in solving the inverse problems. The equivalent sources are distributed within the 3 dB beamwidth of the DAS beampattern. Both overdetermined and underdetermined problems are investigated with layouts depicted in Fig. 2 and Fig. 3. The equivalent source surface is retracted distance from the reconstruction surface by 5mm. In the overdetermined case, 25 equivalent sources which is less than the number of microphones (30) are assumed. The power estimated by Eq.(16) is W = 2.24 Watts which is significantly underestimated. However, in the underdetermined case where the 81 equivalent sources are assumed, the power estimated by Eq.(16) is W = 6.64 Watts which is in close agreement with the theoretical value. Furthermore, the sound power estimated using average intensity method in Eq.(15), W =6.16 Watts, is underestimated. Therefore, power estimation method of Eq.(16) is used hereafter.

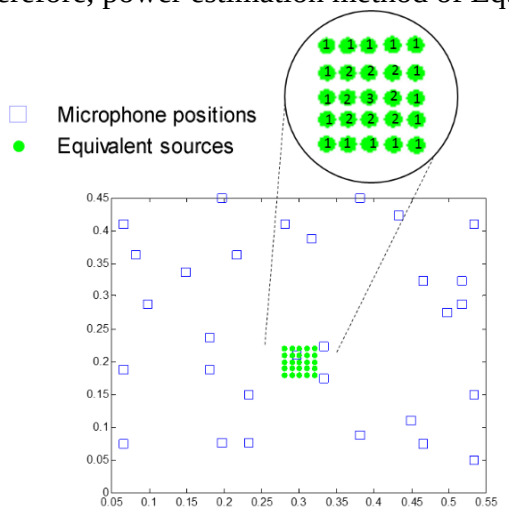


Figure 4: Sound power estimation using the ESM and the TIKR method. The simulated point sources are placed in a 5×5 matrix with amplitudes indicated. The lattice spacing is 1cm.

4.2 Verification of sound power estimation

To verify the two-stage NSI algorithm, experiments are undertaken in a 5.4m×3.5m×2m anechoic room, as shown in Fig. 5. An electrically hand cutter and a hair dryer are used as two noise sources. In the experimental arrangement, 30 channels of 1/4" PCB® condenser microphones are utilized to capture the noise signals. NI PXI-1042Q® is used as the data acquisition system operated under the sample rate 16 kHz. The origin of coordinate is anchored at the bottom-left corner of the rectangular array holder. The source locations measured by using a laser distance meter are (-0.01m, 0.03m, 1m) and (0.65m, 0.27m, 1m), respectively. Similarly, the DAS output is shown in Fig. 6. In addition, 81 distributed equivalent sources within the 3 dB beamwidth of the DAS beampattern are used to compute the acoustic variables for a plane with a 5mm retracted distance, as shown in Fig. 7. All images in particular the velocity map show the source locations very clearly. The sound powers produced by the hand cutter and the hair dryer are estimated to be 1.5×10^{-3} Watts and 3×10^{-2} Watts, respectively.

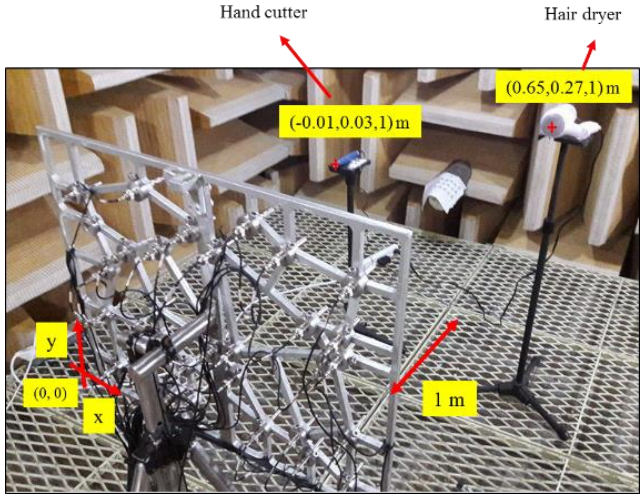


Figure 5: Experimental arrangement of two practical noise sources: a hand cutter and a hair dryer.

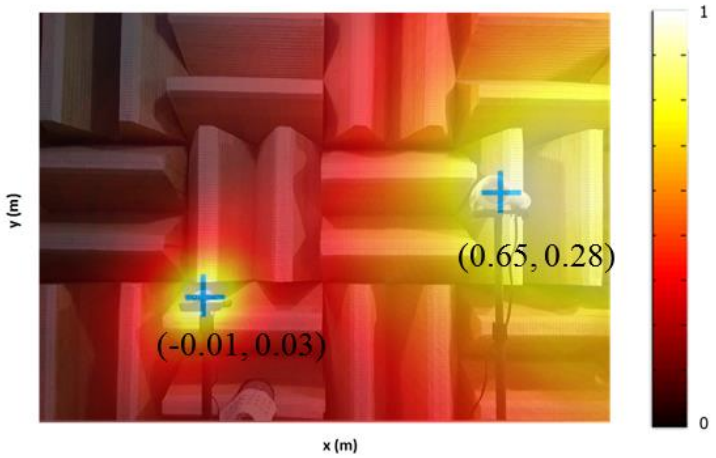
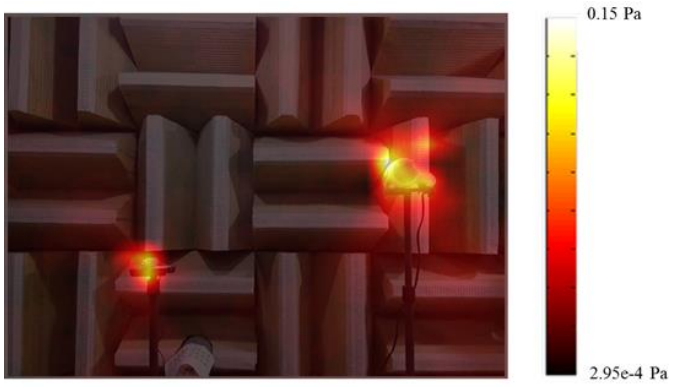
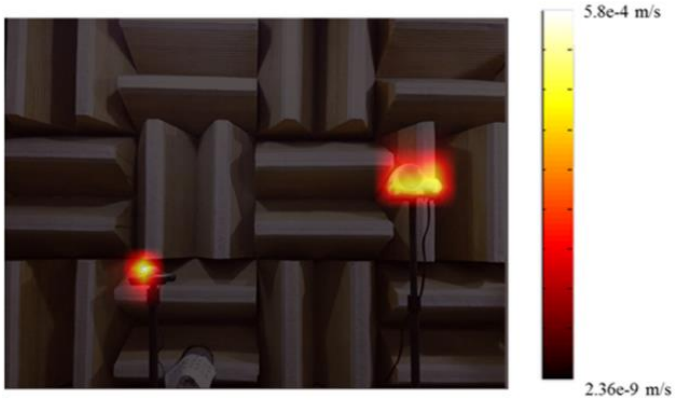


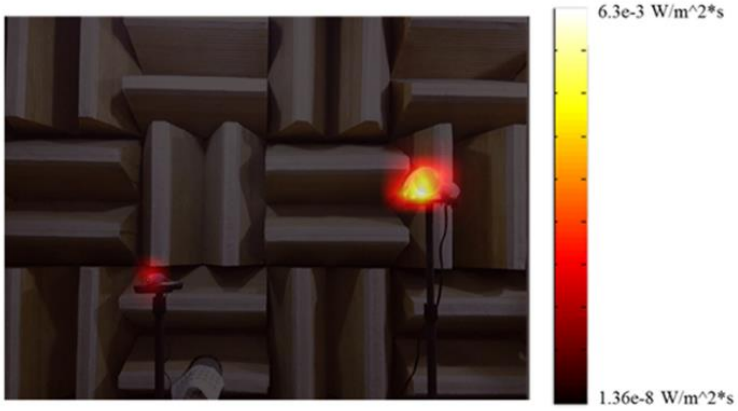
Figure 6: Experimental result of locating two practical sources by using DAS.



(a) Sound pressure



(b) particle velocity



(c) active intensity

Figure 7: Acoustic variables calculated using the ESM for two practical noise sources.

5. Conclusions

The two-stage procedure is executed in the context of source localization and separation. In the localization stage, noise map is obtained using the DAS beamformer. In separation stage, the TIKR method is utilized to solve inverse problem. Post-processing for acoustic variables gives information of pressure, particle velocity, intensity, and radiation power of the noise sources, which is rarely done in conventional farfield arrays. Results of experiments have demonstrated the efficacy in localizing noise sources and the associated characteristics.

6. Acknowledgments

The work was supported by the Ministry of Science and Technology (MOST) in Taiwan, Republic of China, under the project number 102-2221-E-007-029-MY3. Thanks also go to Mr. Eric Chin-Pu Tsai for his enthusiastic support of the Telecom Electroacoustics Audio laboratory (TEA lab)

REFERENCES

1 Bai, M. R., Ih, J. G., and Benesty, J., *Acoustic Array Systems: Theory, Implementation, and Application*, John Wiley & Sons, Solaris South Tower, Singapore (2013).

2 Benesty, J., Chen, J., Huang, Y., *Microphone Array Signal Processing*, Springer-Verlag, Berlin, Germany (2008).

3 Bertero, M., Poggio, T. and Torre V. Ill-Posed Problems in Early Vision, *Proceedings of the IEEE*, **76**(8), 869-889, (1988).

4 Wikipedia. *Tikhonovregularization*. [Online.] available: https://en.wikipedia.org/wiki/Tikhonov_regularization

5 Song, L., Koopmann, G. H. and Fahnline, J. B. Active control of the acoustic radiation of a vibrating structure using a superposition formulation, *J. Acoust. Soc. Am.*, **89**(6), 2786-2792, (1991).