Journal of Sound and Vibration 330 (2011) 545–558

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Implementation issues of the nearfield equivalent source imaging microphone array

Mingsian R. Bai*, Jia-Hong Lin, Chih-Wen Tseng

Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hsueh Road, Hsin-Chu 300, Taiwan, ROC

ARTICLE INFO

Article history:
Received 6 May 2010
Received in revised form
7 August 2010
Accepted 7 August 2010
Handling Editor: L. Huang
Available online 9 September 2010

ABSTRACT

This paper revisits a nearfield microphone array technique termed nearfield equivalent source imaging (NESI) proposed previously. In particular, various issues concerning the implementation of the NESI algorithm are examined. The NESI can be implemented in both the time domain and the frequency domain. Acoustical variables including sound pressure, particle velocity, active intensity and sound power are calculated by using multichannel inverse filters. Issues concerning sensor deployment are also investigated for the nearfield array. The uniform array outperformed a random array previously optimized for far-field imaging, which contradicts the conventional wisdom in far-field arrays. For applications in which only a patch array with scarce sensors is available, a virtual microphone approach is employed to ameliorate edge effects using extrapolation and to improve imaging resolution using interpolation. To enhance the processing efficiency of the time-domain NESI, an eigensystem realization algorithm (ERA) is developed. Several filtering methods are compared in terms of computational complexity. Significant saving on computations can be achieved using ERA and the frequency-domain NESI, as compared to the traditional method. The NESI technique was also experimentally validated using practical sources including a 125 cc scooter and a wooden box model with a loudspeaker fitted inside. The NESI technique proved effective in identifying broadband and non-stationary sources produced by the sources.

1. Introduction

Noise source identification (NSI) by sound field imaging using microphone arrays has received much research attention in recent years [1–3]. These techniques fall into two categories: far-field beamformers (FB) [4,5] and nearfield acoustical holography (NAH) [6]. This paper has different purpose than the traditional array techniques for far-field beamforming and direction of arrival (DOA) estimation in that the present work is aimed at noise source identification and sound field visualization in the nearfield, where measurement is conducted within the order of a wavelength. In the early 1980s, Fourier NAH was introduced in a series of papers by Williams and Maynard to image sources with regular geometries such as planar surfaces [6–9]. This technique revolutionized the NSI technology and later was commercialized with the name spatial transformation of sound field (STSF) by Brüel and Kjær [10–12]. Although Fourier NAH provides a clear picture of noise distribution, it suffers from a number of limitations which stem from the fact that NAH relies on the fast Fourier transform (FFT) of uniformly sampled pressure data on the hologram. To avoid

* Corresponding author. Tel.: +886 3 57 42915; fax: +886 3 5722840.
E-mail address: msbai@mail.nctu.edu.tw (M.R. Bai).

0022-460X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2010.08.017

spatial aliasing, the spacing of microphones is required to be less than one-half the wavelength. To avoid spatial wrap-around, the surrounding of the aperture would generally need to be padded with zeros. However, these two considerations often result in a large channel count. To address these issues, methods that are capable of handling sources with non-stationary noise [12] and arbitrary geometries were suggested since then [13–18]. In the 1992s, the inverse boundary element method (IBEM) [14–16] was suggested to reconstruct the sound field radiated by arbitrary shaped sources. Helmholtz integral equations alongside regularization methods [18] were exploited to reconstruct acoustic radiation from arbitrarily shaped sources. In the late 1990s, the Helmholtz equation least squares (HELS) method was suggested by Wu [19]. In the HELS, the sound field is represented by using spherical wave expansion with the coefficients determined by minimizing the errors in the least-square sense between the measured and the regenerated pressure data. This method proved effective in imaging blunt and convex sources. Another NAH method, statistically optimal NAH (SONAH) [20], was also suggested, based on plane wave expansion. In contrast to the HELS and SOHNA which are based on field representations using global basis functions, another category of NAH methods were derived by straightforward discretization of single-layer potential in the Helmholtz integral [21–23]. These methods are generally referred to as the equivalent source methods (ESM) in which the sound field is represented by a distribution of discrete point sources in the frequency-domain processing [24–27]. The advantages of ESM lie in its simplicity, robustness and accuracy in reconstruction [25]. Comprehensive coverage of NAH can be found in the monograph by Williams [28] and a recent tutorial paper by Wu [29].

In previous papers of the authors, the NAH method termed the nearfield equivalent source imaging (NESI) was proposed for noise source identification and sound filed reconstruction [30–32]. The NESI *per se* can be considered as a time-domain ESM. Multichannel inverse filters are designed offline using truncated singular value decomposition (TSVD) or Tikhonov regularization. Since all the required processing is carried out in the time domain, NESI eliminates many problems of Fourier NAH. However, NESI also suffers from various implementation issues which shall be examined in this paper. First, often times, only patch arrays with scarce sensors are allowed in practical applications. A virtual microphone technique is suggested to minimize the edge effect using extrapolation and to improve the imaging resolution using extrapolation. Second implementation issue of NESI is the enormous computation complexity incurred by the multichannel inverse filtering. To combat this problem, a state-space technique based on eigensystem realization algorithm (ERA) [33] is developed. In addition, an efficient frequency-domain overlap-add (FDOA) algorithm is also suggested. The proposed ERA method and the FDOA method are compared with the traditional filtering approach—the benchmark direct convolution (DC) method [34], in terms of operations per sample (OPS) and imaging performance. The third implementation issue of NESI is concerning sensor deployment. It is well known that randomizing sensor positions helps reducing the grating lobes in applying sparse arrays to far-field beamforming. It is interesting to see whether or not this conventional wisdom applies to the NAH as well. An informal numerical study was conducted to explore the optimal deployment of microphone arrays in near-field acoustic source imaging. The result disproved this conjecture and concluded that uniform arrays are preferred in nearfield imaging. This paper revisits this issue through experiments for a 125 cc scooter and a wooden box model with a loudspeaker fitted inside.

2. The NESI algorithm

In this section, the NESI algorithm [30] is briefly reviewed. NESI can be regarded as a time-domain ESM (Fig. 1). We begin the formulation in the frequency domain. In the step A, the vector $\mathbf{q}$ representing the complex amplitudes of the point sources on the focal surface and the vector $\mathbf{p}$ representing the pressures picked up at the microphones can be related by the propagation matrix $\mathbf{G}$ as follows:

$$p = G\mathbf{q}, \tag{1}$$

where

$$\mathbf{G} = \left\{ G(\mathbf{x}_m, \mathbf{y}_j, \omega) \right\},$$
$$G(\mathbf{x}_m, \mathbf{y}_j, \omega) = \frac{i\rho_0\omega}{4\pi} \frac{e^{-ikr_{jm}}}{r_{jm}}, \tag{2}$$

is the free-space Green's function [28] with $\boldsymbol{r}_{jm} = |\mathbf{x}_m - \mathbf{y}_j|$ being the distance between the source point $\mathbf{y}_j$ and the field point $\mathbf{x}_m$, ρ_0 is the air density, $i = \sqrt{-1}$, $k = \omega/c$ is the wavenumber, ω is the angular frequency and c is the speed of sound.

The inverse filtering problem depicted in the step B of Fig. 1 can be posed as a multiple-input–multiple-output (MIMO) model matching problem:

$$\min_{C} ||\mathbf{W} - \mathbf{CG}||_F^2, \tag{3}$$

where $||\cdot||_F^2$ symbolizes the Frobenius norm [35], $\mathbf{C}$ denotes the inverse filter and $\mathbf{W}$ denotes the matching model into which frequency weighting and window functions can be incorporated in addition to the simple identity matrix $\mathbf{I}$. TSVD [35] or Tikhonov regularization [36] is employed in the inverse filter design. Next, inverse FFT with circular shift is used to

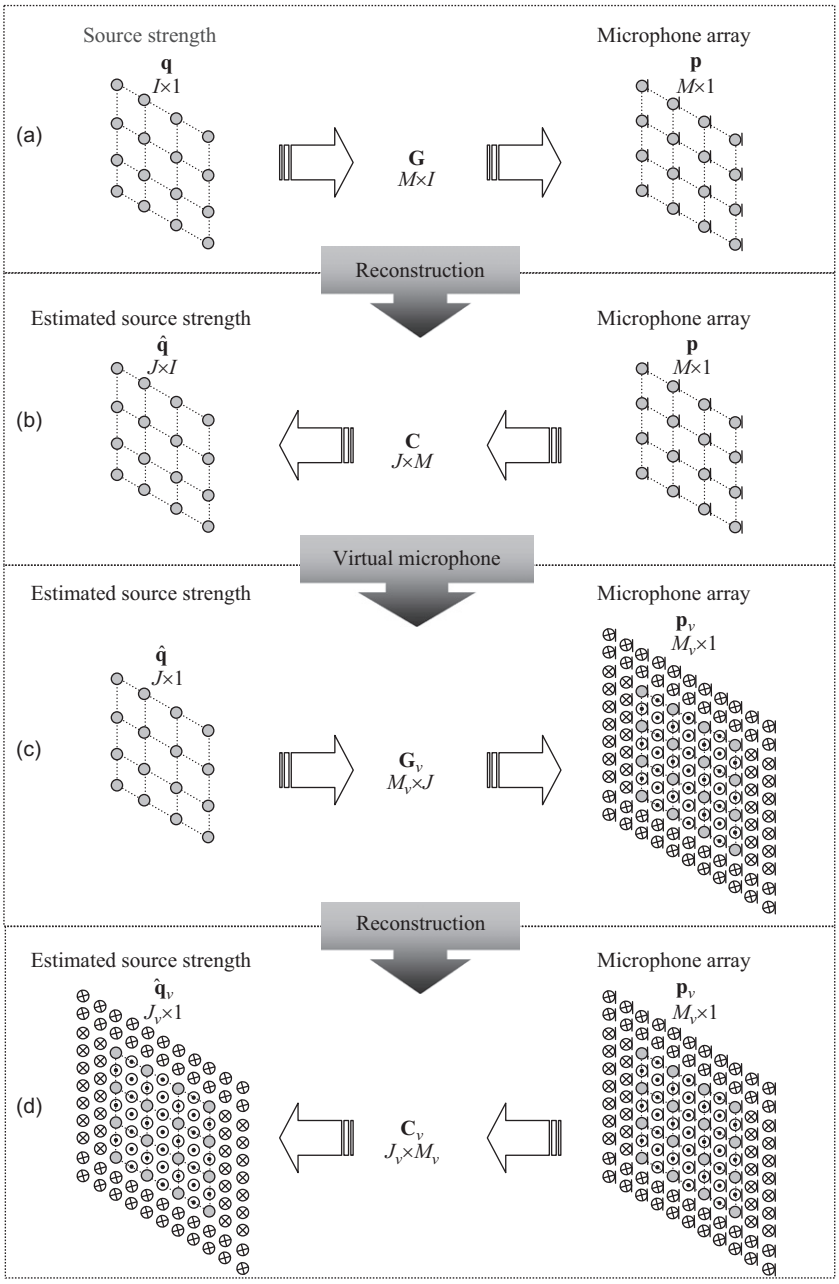

Fig. 1. The idea of the NESI with virtual microphone technique. The symbol "⊙" indicates an interpolated microphone position. The symbol "⊕" indicates an extrapolated microphone position: (a) The pressure data picked up by the microphones. (b) Reconstructed source strength at the focal points. (c) The pressure data interpolated at the virtual microphones. (d) Reconstructed source strength at the virtual focal points.

convert these frequency response functions into finite impulse response (FIR) filters in the time domain. The virtual source amplitudes at the focal points are calculated by multichannel convolutions:

$$\hat{q}_j(n) = \sum_{m=1}^{M} p_m(n) * C_{jm}(n), \quad 1 \le j \le J, \tag{4}$$

or in the matrix form

$$\underset{J\times 1}{\hat{\mathbf{q}}}(n) = \underset{J\times M}{\mathbf{C}}(n) * \underset{M\times 1}{\mathbf{p}}(n), \tag{5}$$

where n denotes the discrete-time index, $C_{jm}(n)$ denotes the impulse response of the jmth inverse filter associated with the jth focal point and the mth microphone and "*" denotes the convolution operator. In the frequency domain, Eq. (5)

can be rewritten as

$$\hat{\mathbf{q}}J\times 1(\omega)=\underset{J\times M}{\mathbf{C}}(\omega)\underset{M\times 1}{\mathbf{p}}(\omega), \tag{6}$$

where the angular frequency $\omega=2\pi f$ and f is the frequency in Hz. Thus obtained source amplitudes at the focal points serve as the basis for the subsequent calculation of acoustical quantities including sound pressure, particle velocity, active intensity and sound power.

To avoid singularity, the focal points are "retreated" slightly with a distance, e.g., $L_r=d/2$, from the reconstruction surface, as shown in Fig. 2. The time-domain sound pressure can be reconstructed as

$$p_r(\mathbf{x},n)=\sum_{j=1}^{J}\frac{\hat{q}_j(n-\Delta_{xj})}{r'_{xj}}, \tag{7}$$

where $\mathbf{x}$ is the position vector of the field point on the reconstruction surface, $\mathbf{y}_j$ is the position vector of the jth virtual point source, $r'_{xj}=|\mathbf{x}-\mathbf{y}_j|$ and $\Delta_{xj}=r'_{xj}/c$ is the propagation time delay implemented using Lagrange interpolation [28]. For the frequency domain processing, this can simply be written as

$$\mathbf{p}_r(\omega)=\mathbf{G}_r(\omega)\hat{\mathbf{q}}(\omega), \tag{8}$$

where $\mathbf{G}_r$ is propagation matrix between the focal surface and the reconstruction surface. For a specified velocity schedule, the Doppler effect and a pass-by noise of a moving source can easily be simulated [28]. With sound pressure given in Eq. (8), the normal particle velocity can be written as

$$u_n(\mathbf{x},\omega)=\frac{-1}{j\rho\omega}\frac{\partial}{\partial n}p(\mathbf{x},\omega)=\frac{-1}{j\rho\omega}\mathbf{n}\cdot\nabla_{\mathbf{x}}\left(\frac{\hat{q}(\omega)e^{-jkr}}{r}\right)=\frac{1}{j\rho\omega}(\mathbf{n}\cdot\mathbf{e}_r)\left(jk+\frac{1}{r}\right)p_r(\mathbf{x},\omega), \tag{9}$$

where ρ is the density of air, $\mathbf{e}_{r'_{xj}}=(\mathbf{x}-\mathbf{x}_0)/r'_{xj}$, $r=|\mathbf{x}-\mathbf{x}_0|$ and $k=\omega/c$ is the wavenumber. Laplace transform and Prewarped bilinear transform [34] is employed to transform Eq. (9) to the discrete-time domain. Thus, the normal particle velocity can be obtained by filtering the sound pressure using a digital filter:

$$u_n(\mathbf{x},z)=\frac{1}{\rho c}(\mathbf{n}\cdot\mathbf{e}_{r'_{xj}})\frac{\tau_0}{\tau}\frac{(1+g\tau)z+(1-g\tau)}{(1+g\tau_0)z+(1-g\tau_0)}p_r(\mathbf{x},z), \tag{10}$$

where z is the z-transform variable, $\tau=r'_{xj}/c$ is the time delay, $\tau_0>>0$ and $g=2\pi f_p/\tan(\pi f_p/f_s)$, where f_p is the bandwidth of interest and f_s is the sampling rate.

The instantaneous normal active intensity is calculated in the time domain by using

$$I_n(\mathbf{x},n)=p_r(\mathbf{x},n)u_n(\mathbf{x},n). \tag{11}$$

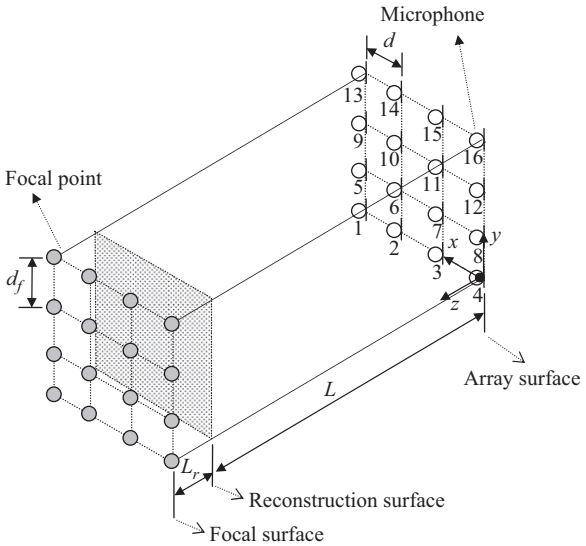

Fig. 2. The array settings for NESI using a 4×4 URA.

Sound power can be calculated by integrating the normal active intensity over an area of interest S

$$W(n) = \int_S I_n(\mathbf{x}, n)\, dS(\mathbf{x}) \tag{12}$$

Root-mean-squares (rms) quantities can be calculated by time-averaging the instantaneous squares quantities.

Many versatile functions of the NESI including imaging of arbitrarily shaped source, pass-by noise prediction of moving sources, far-field reconstruction, far-field polar radiation pattern, etc., have been implemented, but are omitted in this paper due to limitation of space.

3. The implementation issues of NESI

3.1. Virtual microphone interpolation and extrapolation

In practical implementation of the NESI technique, an edge effect may occur when the physical extent of source is larger than the patch array aperture. In addition, the number of sensors may be too scarce to yield acceptable imaging resolution. To address these problems, a virtual microphone technique is employed with field interpolation (for improving resolution) and extrapolation (for reducing edge effect). This following example demonstrates this technique using 4×4 uniform rectangular array (URA) with microphone spacing d. This rather coarse array configuration is to be interpolated and extrapolated into 11×11 grid. The distance of reconstruction (DOR) is chosen to be $d/2$ so that the condition number of the propagation matrix $\mathbf{G}_v$ was well below 1000 [30], where $\mathbf{G}_v$ is between the virtual microphone surface and reconstruction surface. In the step C of Fig. 1, the source amplitudes on the focal surface $\hat{\mathbf{q}}J\times 1$ estimated by NESI are used to calculate sound pressure p_v for a finer grid on the microphone surface:

$$p_v(\mathbf{x}_v, n) = \sum_{j=1}^{J} \frac{\hat{q}_j(n-\Delta_{vj})}{r'_{vj}}, \tag{13}$$

where $\mathbf{x}_v$ is the position vector of the field point on the microphone surface, $r'_{vj} = |\mathbf{x}_v - \mathbf{y}_j|$, $\mathbf{y}_j$ is the position vector of the jth point source on the focal surface and $\Delta_{vj} = r'_{vj}/c$ is the time delay. The sound pressures regenerated using Eq. (13) for the interpolated and extrapolated actual/virtual sensor locations with a finer spacing can be assembled into the matrix form

$$\underset{M_v\times 1}{\mathbf{p}_v}(n) = \underset{M_v\times J}{\mathbf{G}_v}(n)\ \underset{J\times 1}{\hat{\mathbf{q}}}(n), \tag{14}$$

where $\mathbf{G}_v$ is the propagation matrix between the focal surface and the microphone surface, $M_v = 11 \times 11 = 121$ is the number of microphone and $J = 4 \times 4 = 16$ is the number of point sources on the focal surface. In the frequency domain, the sound pressure is calculated by

$$\underset{M_v\times 1}{\mathbf{p}_v}(\omega) = \underset{M_v\times J}{\mathbf{G}_v}(\omega)\ \underset{J\times 1}{\hat{\mathbf{q}}}(\omega) = \underset{M_v\times J}{\mathbf{G}_v}(\omega)\ \underset{J\times M}{\mathbf{C}}(\omega)\ \underset{M\times 1}{\mathbf{p}}(\omega), \tag{15}$$

where $\hat{\mathbf{q}}J\times 1(\omega) = \mathbf{C}_{J\times M}(\omega)\mathbf{p}_{M\times 1}(\omega)$ in Eq. (6). In Fig. 1, the interpolated and extrapolated microphones are indicated with the symbols "⊙" and "⊛", respectively. Next, choose a new point source distribution with finer spacing. The source amplitudes $\hat{\mathbf{q}}_v$ are estimated with the augmented inverse filters $\mathbf{C}_v$ in the time domain and the frequency domain as shown in the step d of Fig. 1:

$$\hat{\mathbf{q}}_{vJ_v\times 1}(n) = \underset{J_v\times M_v}{\mathbf{C}_v}(n) * \underset{M_v\times 1}{\mathbf{p}_v}(n) \tag{16}$$

and

$$\hat{\mathbf{q}}_{vJ_v\times 1}(\omega) = \underset{J_v\times M_v}{\mathbf{C}_v}(\omega)\ \underset{M_v\times 1}{\mathbf{p}_v}(\omega) = \underset{J_v\times M_v}{\mathbf{C}_v}(\omega)\ \underset{M_v\times J}{\mathbf{G}_v}(\omega)\ \underset{J\times M}{\mathbf{C}}(\omega)\ \underset{M\times 1}{\mathbf{p}}(\omega), \tag{17}$$

where $M_v = 11 \times 11 = 121$ is the number of virtual microphones and $J_v = 11 \times 11 = 121$ is the number of virtual point sources.

3.2. Multichannel fast filtering

Apparently, execution of the vast number of inverse filters in NESI is very computationally demanding for all processing is carried out in the time domain. This is becoming even more problematic in practical applications when the aforementioned virtual microphone technique is used. To overcome this problem, the ERA is employed in this paper. SVD is exploited in ERA to remove the redundancy among the inverse filters and to guarantee a balanced minimal realization of the $J \times M$ inverse filter matrix $\mathrm{C}(n)$:

$$\mathbf{x}(n+1) = \mathbf{A}_e\mathbf{x}(n) + \mathbf{B}_e\mathbf{u}(n), \tag{18}$$

$$\mathbf{y}(n) = \mathbf{C}_e\mathbf{x}(n) + \mathbf{D}_e\mathbf{u}(n), \tag{19}$$

where n is the discrete-time index, $\mathbf{x}(n)$ is the state vector, $\mathbf{u}(n)$ is the $M\times 1$ input vector, $\mathbf{y}(n)$ is the $J\times 1$ output vector and $\mathbf{A}_e$, $\mathbf{B}_e$, $\mathbf{C}_e$, $\mathbf{D}_e$ are the constant matrices defined as follows:

$$\mathbf{A}_e=\Sigma_v^{-1/2}\mathbf{U}_v{}^H\mathbf{H}(1)\mathbf{V}_v\Sigma_v{}^{-1/2},\quad \mathbf{B}_e=\Sigma_v{}^{1/2}\mathbf{V}_v{}^H\mathbf{E}_M,\quad \mathbf{C}_e=\mathbf{E}_J^H\mathbf{U}_v\Sigma_v{}^{1/2}\quad\text{and}\quad \mathbf{D}_e=\mathbf{C}(0), \tag{20}$$

where $\mathbf{E}_M^H=\begin{bmatrix}\mathbf{I}_M & \mathbf{0}_M & \cdots & \mathbf{0}_M\end{bmatrix}$, $\mathbf{E}_J^H=\begin{bmatrix}\mathbf{I}_J & \mathbf{0}_J & \cdots & \mathbf{0}_J\end{bmatrix}$, $\mathbf{I}$ is an identity matrix and 0 is a null matrix, and $\mathbf{H}$ is a Hankel matrix [33] derived from impulse responses of the matrix C(n). $\mathbf{U}_v$ and $\mathbf{V}_v$ are the unitary matrices and $\mathbf{\Sigma}$ is a diagonal matrix composed of singular values of the Hankel matrix $\mathbf{H}(0)$, i.e., $\mathbf{H}(0)=\mathbf{U}\Sigma\mathbf{V}^H$:

$$\Sigma=\begin{bmatrix}\Sigma_v & 0\\ 0 & 0\end{bmatrix}\quad\text{and}\quad \Sigma_v=\operatorname{diag}\begin{bmatrix}\sigma_1 & \sigma_2 & \cdots & \sigma_v\end{bmatrix}, \tag{21}$$

where v is the matrix rank. The ERA was detailed in Ref. [30]. The OPS for ERA is

$$\mathrm{OPS(ERA)}=v^2+vM+vJ+MJ. \tag{22}$$

In this paper, the ERA is compared to widely used methods, the DC. Direct linear convolution denoted as DC is also used as a benchmark method. The OPS of the DC method is given by

$$\mathrm{OPS(DC)}=M\times J\times N_f, \tag{23}$$

where N_f is the length of FIR filter.

The NESI algorithm can also be efficiently implemented in the frequency domain. First, partition the time-domain microphone pressure data p(n) into non-overlapping frames and zero-pad the frames into $p_\Lambda(n)$, where $\Lambda=1, 2,\ldots\curlywedge$ is the frame index and $\curlywedge$ is the number of frames, as shown in Fig. 3(a) and (b). Transform each frame to the frequency domain by using the FFT. Next, multiply the transformed pressure data with the frequency-domain inverse matrix $\mathbf{C}(\omega)$ that can be

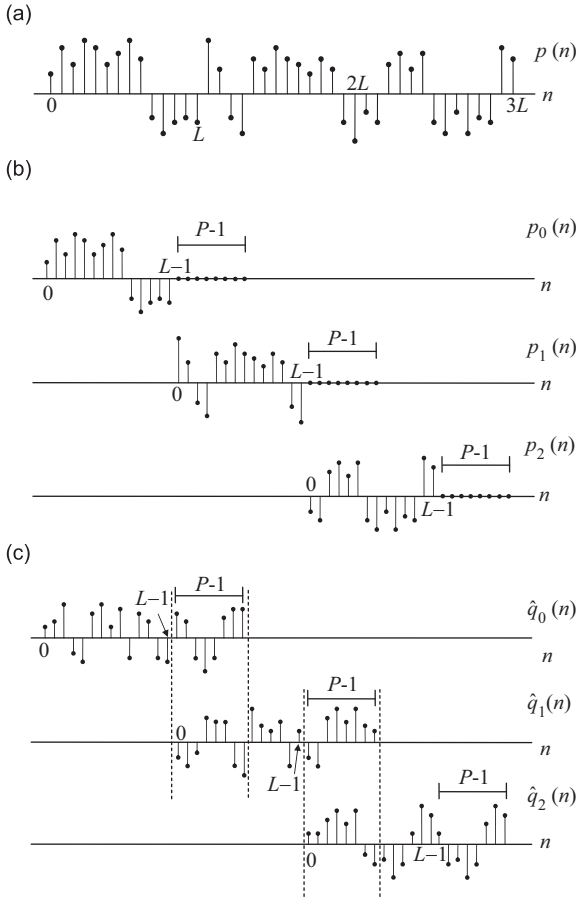

Fig. 3. Illustration of the overlap and add method: (a) the pressure data $p(n)$, (b) decomposition of $I(n)$ into non-overlapping sections of length L and (c) result of convolving each section with the inverse filter.

computed offline. Finally, calculate the time-domain source amplitudes $\hat{\mathbf{q}}_{\Lambda}(n)$ for each frame by using the inverse FFT and overlap and add the consecutive frames, as shown in Fig. 3(c).

Overlap-and-add technique can be used if continuous processing is desired. To illustrate how to choose parameters in the overlap-and-add block processing, consider the impulse response of the inverse filter matrix $\mathbf{C}(\omega)$ of length P. Assume that there are L samples in each frame $\mathbf{p}_{\Lambda}(n)$. Thus, the output of linear convolution $\hat{\mathbf{q}}_{\Lambda}(n)=C(n)*\mathbf{p}_{\Lambda}(n)$ has the length $(L+P-1)$. The linear convolution can be efficiently implemented, with the aid of FFT, by calculating the product $\mathbf{C}(\omega)\mathbf{p}_{\Lambda}(\omega)$ in the frequency domain, where $N\geq L+P-1$ point FFT must be used to avoid wraparound errors. To meet this length requirement, each frame must be padded with $(P-1)$ zeros. After inverse filtering, each frame of the source amplitude $\hat{\mathbf{q}}_{\Lambda}(n)$ is added with $(P-1)$ overlapped points. This is referred to as the frequency-domain-overlap-add (FDOA) algorithm in the following presentation. Tremendous computation efficiency can be gained because the frequency-domain inverse matrix needs be computed offline for only once. The OPS of FDOA is estimated to be

$$\mathrm{OPS(OA)}=(M+J)\log N_i+JM. \tag{24}$$

3.3. Optimal sensor deployment

For far-field imaging, arrays with sparse and random sensor deployment are known to be capable of delivering far-field images without grating lobes. Whether this idea of random deployment applies to nearfield imaging as well is an interesting question. An informal numerical study was conducted, with the aid of global optimization algorithms, to explore the optimal deployment of microphone arrays in near-field acoustic source imaging. In the nearfield optimization, a special definition of beam pattern and cost function is used for the MIMO imaging problem. Near-field pressure field beam pattern is calculated by scanning on the focal surface using a test point source. The cost function of the near-field array is defined as

$$Q=\sum_{j=1}^{J} m_j, \tag{25}$$

where m_j denotes the maximum of a main-lobe region whose level is greater than 0.707 (-3 dB) and J is the number of focal points. The objective is to find the deployment that maximizes the cost function above. Simulation annealing (SA) [37] algorithm was employed for optimizing the sensor positions. A mechanism that prevents the search from being trapped in local minima is built in the SA algorithm. This mechanism is based on a probability to accept "worse" solutions in the initial high-temperature stage

$$\begin{cases} e^{\Delta Q/T}>\gamma & \text{accepted,} \\ e^{\Delta Q/T}<\gamma & \text{rejected,} \end{cases} \tag{26}$$

where $\Delta Q=Q-Q_{\mathrm{opt}}$ is the difference between the present and the optimal cost function, T is the present temperature and γ is the random number $0\leq\gamma\leq 1$ generated subject to the uniform distribution.

A counterintuitive result obtained in this study revealed that random deployments seem to offer little benefit in nearfield imaging of planar sources and the optimal array is simply the URA. This problem will be revisited in this paper through experimental investigations on a desktop computer.

4. Numerical investigation

Numerical simulation is conducted to validate the virtual microphone technique and multichannel filtering algorithms. A URA with $M=J=4\times 4$ is employed in this simulation, as shown in Fig. 2. The spacing of the microphones (d) and the focus points (d_f) were both selected to be 0.1 m$=\lambda/2$ for 1.7 kHz.

Another important parameter to choose is distance of reconstruction (DOR) that entirely depends on the degree of ill-posedness of the inverse problem. Assume that the acoustic radiation problem can be formulated via ESM into the following matrix equation:

$$\mathbf{Ax}=\mathbf{b}, \tag{27}$$

where $\mathbf{b}$ and $\mathbf{x}$ are the hologram data and source data, respectively, which are related by the propagation matrix $\mathbf{A}$. It can be shown the perturbation term δb of the data vector such as measurement noise, numerical error, etc., and the perturbation term δx of the reconstructed data satisfy the following inequality [35]:

$$\frac{||\delta\mathbf{x}||}{||\mathbf{x}||}\leq \mathrm{cond}(\mathbf{A})\frac{||\delta\mathbf{b}||}{||\mathbf{b}||}, \tag{28}$$

where $\mathrm{cond}(\mathbf{A})=\sigma_{\max}/\sigma_{\min}$ is the condition number of the matrix $\mathbf{A}$ and $||\cdot||$ symbolizes vector 2-norm. Therefore, as an indicator of the ill-posedness inherent in the inverse filtering process, the condition number can also be regarded as a magnification factor of perturbations as well as loss of SNR after inverse filtering. For example, the SNR of data will be reduced by 60 dB of dynamic range after inverse filtering if cond(A)=1000. It is well known that condition number of the propagation matrix increases with the DOR since the evanescent wave decays rapidly with the distance. Thus, the

Table 1
Comparison of computational complexity in terms of OPS of three multichannels filtering methods for three array configurations. The block size of FFT N_i=512. The numbers of microphones and focal points are assumed to be equal, i.e., m=j. The DC method is used for benchmarking (100% in parenthesis).

Domain	Method	4 × 4 URA	5 × 6 URA	8 × 8 URA
Time	DC	65,536 (100%)	230,400 (100%)	1,048,576 (100%)
	ERA	1936 (2.95%)	16,900 (7.34%)	69,696 (6.65%)
Frequency	FDOA	544 (0.83%)	1440 (0.63%)	5248 (0.5%)

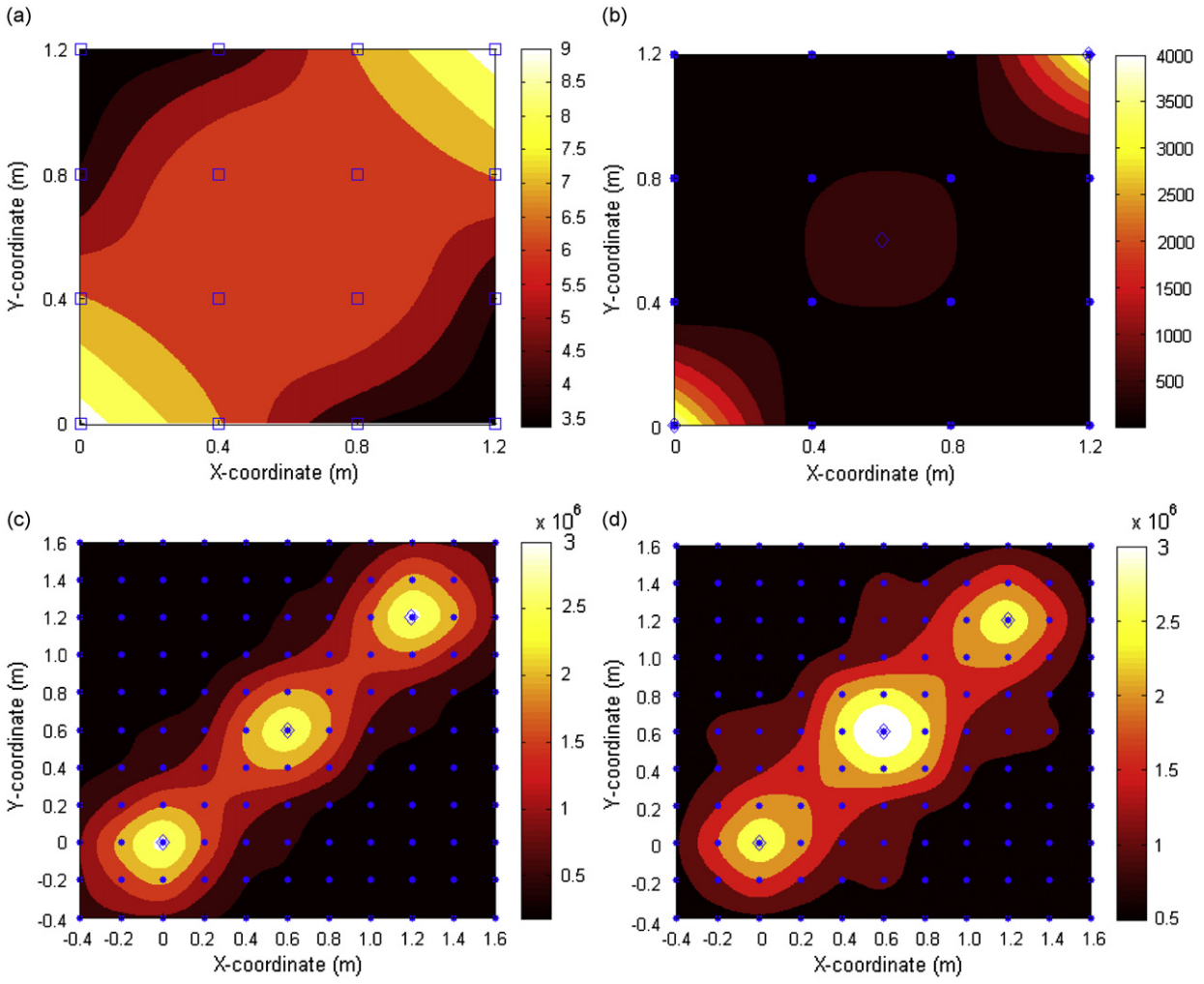

Fig. 4. The numerical simulation of NESI using the 4 × 4 URA and the virtual microphone technique: (a) the unprocessed sound pressure image received at the microphones, (b) the reconstructed active intensity image by 4 × 4 URA, (c) the reconstructed active intensity image using the virtual microphone technique in time domain processing and (d) the reconstructed active intensity image using the virtual microphone technique in frequency domain processing. The symbol "□" indicates the microphones. The symbol "●" indicates the focal points. The symbol "◇" indicates the noise sources.

condition number can be used as a useful criterion for choosing the DOR. Thus, given a 60 dB tolerance of loss of SNR, a DOR that gives a condition number less than 1000 is generally deemed appropriate. In the following examples, DOR (L) was chosen to be d/2 to yield a condition number less than 1000.

In the inverse filter design, Tikhonov regularization parameter was selected according to the L-curve method [38]. The OPS required by three different filtering methods (DC, ERA and FDOA) is compared for three different array configurations (16, 30 and 64 channels) in Table 1. The number of FFT frequency points N_i=512. The numbers of microphones and focal points are assumed to be equal, i.e., M=J. The most computationally expensive DC method is used for benchmarking as

100% (in parenthesis) OPS requirement. It is obvious from the comparison that the ERA filtering displayed remarkable computation efficiency, e.g., 6.65% for a 64-channel array, as compared with the DC processing. The computation efficiency is considerably improved using the FDOA approach, especially for large number of microphone channels (only 5% of the benchmark DC method for a 64-channel array).

In the example, three random noise sources (band-limited to 1.7 kHz) are situated at (0, 0 m), (0.6, 0.6 m) and (1.2, 1.2 m), respectively. Specifically, two sources are situated at the kitty corners and the remaining source is situated at the center on the focal surface. Fig. 4(a) shows the unprocessed sound pressure (rms) in linear scale received at the microphones. From the quite blurred image, three noise sources were barely resolvable, particularly for the noise source at the center. Fig. 4(b) shows the sound intensity (rms) calculated by the time-domain NESI. Although the image quality is slightly improved, the noise source at the center (0.6, 0.6 m) is still not identifiable because it is not located on the focal point. To overcome this problem, virtual microphone technique was applied to interpolate and extrapolate the pressure field on the microphone surface and increase the number of microphones and focal points from 4×4 to 11×11. In total, 33 and 72 microphones are uniformly distributed inside (interpolation) and outside (extrapolation), respectively, the original array aperture. With the new setting, the sound intensity reconstructed using the time-domain NESI is shown in Fig. 4(c). It can be clearly observed from the result that the quality of the reconstructed image was significantly improved. Problems due to edge effect and insufficient resolution were basically eliminated. Three sources including the one at the center are

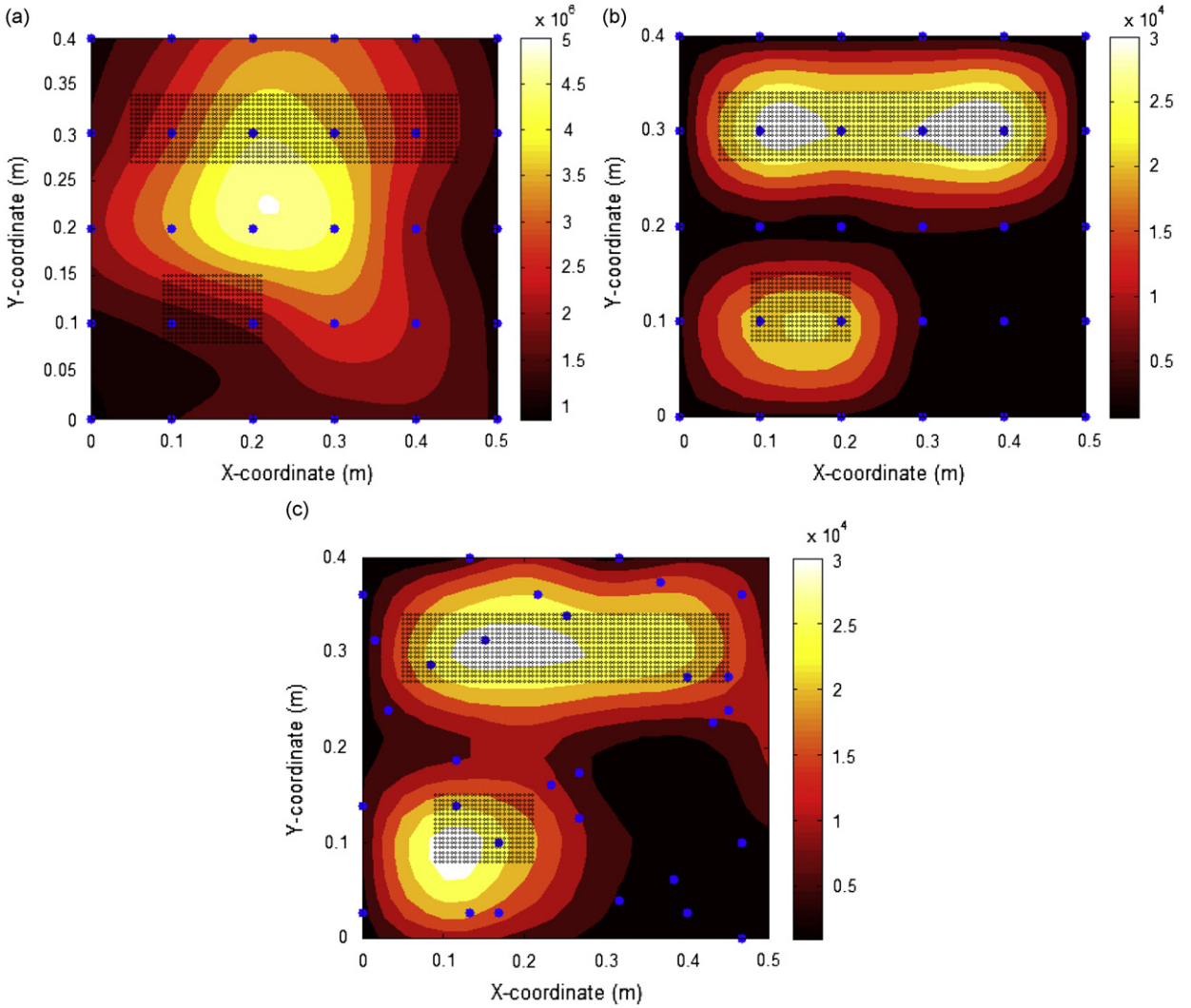

Fig. 5. The simulation results of two piston sources emanating random noise (band-limited to 1.7 kHz): (a) The unprocessed sound pressure image received at the microphones by 5×6 URA, (b) the active intensity image reconstructed using NESI by the 5×6 URA and (c) the active intensity image reconstructed using NESI by the random array.

clearly visible in the intensity map. Total sound power level is 99 dB re. 1×10^{-12} W. This array setting will be used in the following scooter experiment.

In addition to the simulation above, the frequency-domain NESI is used to identify random noise sources with the same setting as in the time-domain NESI. Here, 512 point FFT with Hamming window is employed in the simulation. Fig. 4(d) shows the active intensity (rms) calculated by the frequency-domain NESI. Although three sources are well located in Fig. 4(d), they exhibit slightly larger spreading than the result in Fig. 4(c) obtained using the time-domain NESI.

To compare which one of URA and random array perform better in nearfield imaging.

The array configuration is a 5×6 URA and a 30-channel random array that was optimized for far-field imaging in Ref. [39]. In Ref. [39], numerical simulations are undertaken to optimize the microphone deployment for both nearfield and far-field arrays by global optimization techniques. As indicated by the simulation results, the optimal nearfield array is the URA and the random deployment presents no particular benefit in nearfield imaging. For more details of the optimal array, one may consult Ref. [39]. The numerical simulation is conducted to test and verify it. Two piston sources emanating random noise (band-limited to 1.7 kHz) can be located in Fig. 5(a) at (0.25, 0.31 m) and (0.15, 0.12 m). Fig. 5(a) shows the unprocessed sound pressure (rms) in linear scale received at the microphones. Fig. 5(b) and (c) shows the particle velocity reconstructed by using URA and a random array optimized for far-field imaging, respectively. From the velocity map, the velocity distribution at URA is better than random array. This attests the conclusion drawn from the numerical simulation that random deployment offers little advantages for nearfield imaging and the URA is the optimal array configuration.

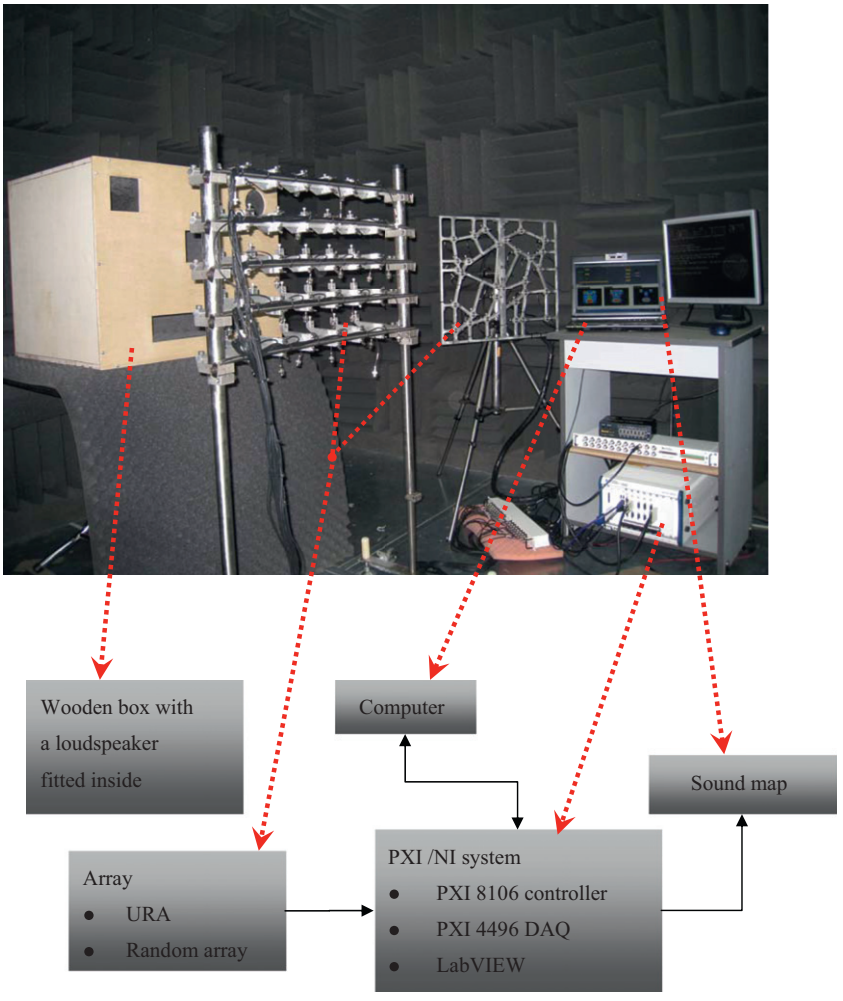

Fig. 6. The experimental arrangement for a wooden box with a loudspeaker fitted inside the URA, and a 30-channel random array optimized for far-field imaging are also shown here.

5. Experimental verifications

To validate the NESI technique, practical sources including a scooter and a wooden box model with a loudspeaker fitted inside were chosen as the test targets for experiments. Fig. 6 shows the block diagram of the experimental arrangement. In the scooter experiment, the array configuration is a 4×4 URA, while in the wooden box experiment, the array configuration is a 5×6 URA. Two PXI 4496 systems [40] in conjunction with LabVIEW [40] were used for data acquisition and processing at the sampling rate 5 kHz. A bandpass filter (20 Hz to 1.7 kHz) is used to prevent aliasing and errors occurring in the out-of-band frequencies. The source amplitude, sound pressure, particle velocity and sound intensity reconstructed using NESI can be displayed on the monitor.

5.1. Scooter experiment

In the experiment, a 125 cc scooter served as a practical source to examine the capability of NESI in dealing with non-stationary sources. The scooter is mounted on a dynamometer inside a semi-anechoic room. The array parameters are selected to be $M=J=4\times4$, $d=d_f=0.1$ m$=\lambda/2$ for 1.7 kHz and $L=d/2$. The Frequency-domain NESI was used to reconstruct the sound field on the right-hand side of the scooter in a run-up test. The engine speed increased from 1500 to 7500 rpm

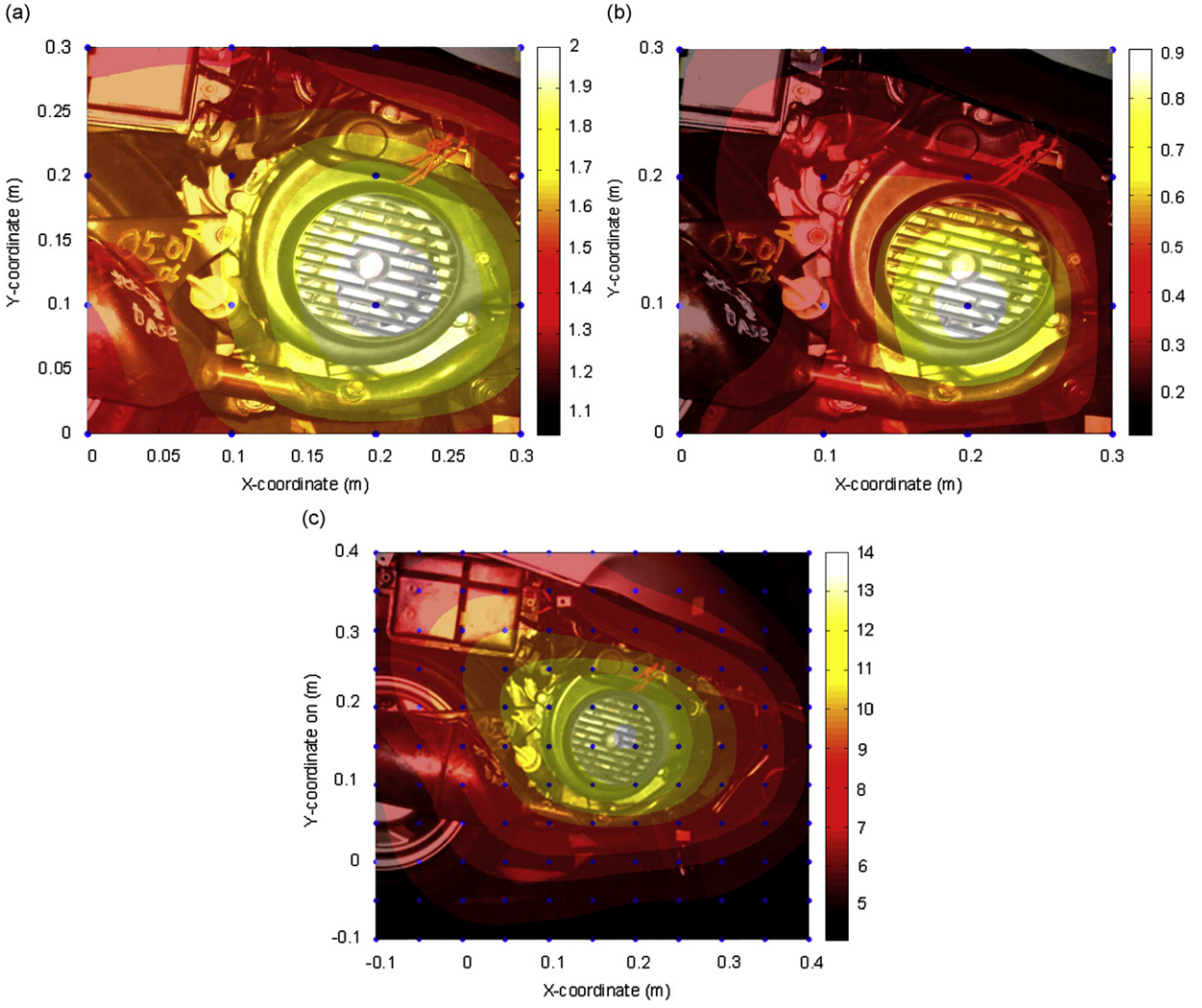

Fig. 7. The results of run-up experiment obtained using NESI with the 4×4 URA. The scooter engine was accelerated from 1500 to 7500 rpm within 10 s: (a) The unprocessed sound pressure image received at the microphones, (b) the reconstructed active intensity image and (c) the reconstructed active intensity image using the virtual microphone technique. The symbol "□" indicates the microphones. The symbol "●" indicates the focal points.

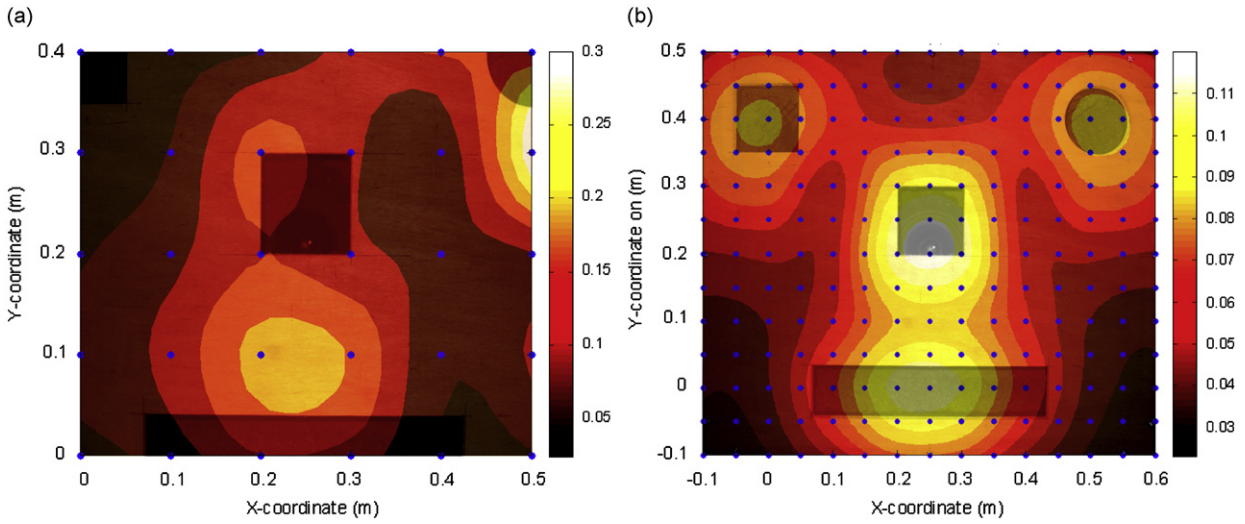

Fig. 8. The results of a wooden box with a loudspeaker fitted inside. The noise map is within the band 200 Hz to 1.6k Hz: (a) The unprocessed sound pressure image received at the microphones by 5×6 URA and (b) the particle velocity image reconstructed using NESI by the 5×6 URA.

within ten seconds. The unprocessed sound pressure received at the microphones is shown in Fig. 7(a), while the rms velocity reconstructed using the NESI is shown in Fig. 7(b). These results revealed that the cooling fan behind the vented engine cover was the major noise source. Next, the virtual microphone technique is employed to see if it is possible to further enhance the image quality by increasing the number of channels from $4 \times 4 = 16$ to $11 \times 11 = 121$. The inverse filters have been designed in the previous numerical investigation. The particle velocity was then reconstructed on the basis of the estimated source amplitude, as shown in Fig. 7(c). Total sound power level is 95 dB re. 1×10^{-12} W. Clearly visible is a larger area of image with improved resolution than that of Fig. 7(b), where again the cooling fan is the major noise source.

5.2. *Wooden box experiment*

In this experiment, a wooden box model with loudspeaker fitted inside is used to validate the NESI technique by using a 5×6 URA. As shown in Fig. 8(a), several holes with different shapes are cut in the front face of the box like a Jack_O_Lantern. A circle, two squares and a slit are located at (0.5, 0.4 m), (0, 0.4 m), (0.25, 0.25 m) and (0.25, 0 m), respectively. The loudspeaker produces random noise band-limited to 1.7 kHz. The microphone spacing d is selected to be 0.1 m ($\lambda/2$ corresponding to $f_{max} = 1.7$ kHz).

Fig. 8(a) shows the unprocessed sound pressure picked up at the microphones within the band 200–1600 Hz. From the image, the noise sources were barely resolvable, particularly for the noise source at the edge—the circle, the slot and the square at upper left corner. Also, the square at the center was difficult to distinguish. Virtual microphone technique was again applied to overcome this problem by interpolate and extrapolate the pressure field on the microphone surface and increase the number of microphones and focal points from 5×6 to 13×15. With the new setting, the particle velocity (rms) reconstructed using the frequency-domain NESI is shown in Fig. 8(b). It can be clearly observed from the result that the quality of the reconstructed image was significantly improved. Problems due to edge effect and insufficient resolution were basically eliminated.

The NESI images apparently yielded more reliable information about noise sources than the unprocessed sound pressure.

6. Conclusions

Various implementation issues of the NESI technique have been investigated in this paper. Issues concerning sensor deployment are also investigated for nearfield imaging. The uniform array outperformed a random array previously optimized for far-field imaging, which contradicts the conventional wisdom in far-field arrays. A virtual microphone technique is suggested for minimizing edge effects using extrapolation and for improving resolution using interpolation. A state-space ERA technique is developed to ease the computation loading of the multichannel filtering in the time-domain NESI. A computationally efficient FDOA algorithm is also proposed for continuous

processing in the frequency domain NESI. Although the frequency-domain NESI (FDOA) is by far the most efficient method among all inverse filtering approaches, it can yield a noise source mapping with slightly larger spreading than the time-domain NESI. Experimental investigations have been undertaken to verify the proposed implementation technique. The NESI technique proved effective in identifying broadband and non-stationary sources produced by these sources.

Acknowledgment

The work was supported by the National Science Council in Taiwan, Republic of China, under the Project no. NSC94-2212-E-009-018. Thanks also go to National Instruments Corporation that made possible this research by donating data acquisition systems.

References

[1] V. Murino, Reconstruction and segmentation of underwater acoustic images combining confidence information in MRF models, *Journal of the Pattern Recognition Society* 34 (2001) 981–997.
[2] Y. Fan, B. Tysoe, J. Sim, K. Mirkhani, A.N. Sinclair, F. Honarvar, H. Sildva, A. Szecket, R. Hardwick, Nondestructive evaluation of explosively welded clad rods by resonance acoustic spectroscopy, *Ultrasonics* 41 (2003) 369–375.
[3] U. Benko, J. Petrovcic, D. Juricic, J. Tavcar, J. Rejec, A. Stefanovska, Fault diagnosis of a vacuum cleaner motor by means of sound analysis, *Journal of Sound and Vibration* 276 (2004) 781–806.
[4] B.D.V. Veen, K.M. Buckley, Beamforming: a versatile approach to spatial filtering, *IEEE ASSP Magazine* (1998) 4–24.
[5] H. Krim, M. Viberg, Two decades of array signal processing research, *IEEE Signal Processing Magazine* (1996) 67–94.
[6] E.G. Williams, J.D. Maynard, Holographic imaging without the wavelength limit, *Physical Review Letters* 45 (1980) 554–557.
[7] E.G. Williams, J.D. Maynard, E. Skudrzyk, Sound source reconstruction using a microphone array, *Journal of the Acoustical Society of America* 68 (1980) 340–344.
[8] E.G. Williams, The nearfield acoustic holography (NAH) experimental method applied to vibration and radiation in light and heavy fluids, *Computers and Structures* 65 (1997) 323–335.
[9] E.G. Williams, B.H. Houston, Fast Fourier transform and singular value decomposition formulations for patch nearfield acoustical holography, *Journal of the Acoustical Society of America* 114 (2003) 1322–1333.
[10] J. Hald, STSF—a unique technique for scan-based nearfield acoustic holography without restriction of coherence, *Brüel and Kjær Technical Review* 1 (1989) 1–50.
[11] J. Hald, STSF—practical instrumentation and application, *Brüel and Kjær Technical Review* 2 (1989) 1–27.
[12] J. Hald, Non-stationary STSF, *Brüel and Kjær Technical Review* 1 (2000) 1–36.
[13] W.A. Veronesi, J.D. Maynard, Digital holographic reconstruction of sources with arbitrarily shaped surface, *Journal of the Acoustical Society of America* 85 (1988) 588–598.
[14] M.R. Bai, Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries, *Journal of the Acoustical Society of America* 92 (1992) 533–549.
[15] B.K. Kim, J.G. Ih, On the reconstruction of the vibro-acoustic field over the surface enclosing an interior space using the boundary element method, *Journal of the Acoustical Society of America* 100 (1996) 3003–3016.
[16] S.C. Kang, J.G. Ih, Use of non-singular boundary integral formulation for reducing errors due to near-field measurements in the boundary element method based near-field acoustic holography, *Journal of the Acoustical Society of America* 109 (2001) 1320–1328.
[17] N.P. Valdivia, E.G. Williams, J. Klos, Reconstruction of the acoustic field using a conformal array, Proceedings of the Inter-noise, 2006.
[18] J. Hald, J. Gomes, A comparison of two patch nah methods, Proceedings of the Inter-noise, 2006.
[19] S.F. Wu, On reconstruction of acoustic pressure fields using the Helmholtz equation least squares method, *Journal of the Acoustical Society of America* 107 (2000) 2511–2522.
[20] J. Hald, Combined NAH and beamforming using the same array, *Brüel and Kjær Technical Review* 1 (2005) 1–50.
[21] G.H. Koopmann, L. Song, J.B. Fahnline, A method for computing acoustic fields based on the principle of wave superposition, *Journal of the Acoustical Society of America* 86 (1989) 2433–2438.
[22] L. Song, G.H. Koopmann, J.B. Fahnline, Numerical errors associated with the method of superposition for computing acoustic fields, *Journal of the Acoustical Society of America* 89 (1991) 2626–2633.
[23] J.B. Fahnline, G.H. Koopmann, A numerical solution for the general radiation problem based on the combined methods of superposition and singular-value decomposition, *Journal of the Acoustical Society of America* 90 (1991) 2808–2819.
[24] A. Sarkissian, Method of superposition applied to patch near-field acoustical holography, *Journal of the Acoustical Society of America* 118 (2005) 671–678.
[25] J. Gomes, J. Hald, P. Juhl, F. Jacobsen, On the applicability of the spherical wave expansion with a single origin for near-field acoustical holography, *Journal of the Acoustical Society of America* 125 (2009) 1529–1537.
[26] P.A. Nelson, S.H. Yoon, Estimation of acoustic source strength by inverse methods—part I, conditioning of the inverse problem, *Journal of Sound and Vibration* 233 (2000) 643–668.
[27] P.A. Nelson, S.H. Yoon, Estimation of acoustic source strength by inverse methods—part II, experimental investigation of methods for choosing regularization parameters, *Journal of Sound and Vibration* 233 (2000) 669–705.
[28] E.G. Williams, *Fourier Acoustics*, Academic Press, New York, 1999.
[29] S.F. Wu, Methods for reconstructing acoustic quantities based on acoustic pressure measurements, *Journal of the Acoustical Society of America* 124 (2008) 2680–2697.
[30] M.R. Bai, J.H. Lin, Source identification system based on the time-domain nearfield equivalence source imaging: fundamental theory and implementation, *Journal of Sound and Vibration* 307 (2007) 202–225.
[31] M.R. Bai, J.H. Lin, A new noise source identification technique: nearfield acoustical beamformer (NABF), Proceedings of the Inter-noise, 2006.
[32] M.R. Bai, J.H. Lin, Sound field reconstruction using the time-domain nearfield equivalence source imaging (NESI) technique, Proceedings of the Inter-noise, 2007.
[33] J.N. Juang, *Applied System Identification*, Prentice-Hall PTR, Englewood Cliffs, NJ, 1994.
[34] A.V. Oppenheim, R.W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[35] B. Noble, *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
[36] O. Kirkeby, P.A. Nelson, H. Hamada, F. Orduna-Bustamante, Fast deconvolution of multichannel systems using regularization, *IEEE Transactions on Speech and Audio Processing* 6 (1998) 189–194.

[37] S. Kirkpatrick, C.D. Gelatt, Jr., M.P. Vecchi, Optimization by simulated annealing, *Science* 220 (4598) (1983) 671–680.
[38] A. Schuhmacher, Sound Source Reconstruction Using Inverse Sound Field Calculations, Ph.D. Dissertation, Department of Acoustic Technology, Technical University of Denmark, Report No. 77, 2000, ISSN 1397–0547.
[39] M.R. Bai, J.H. Lin, K.L. Liu, Optimized microphone deployment for near-field acoustic holography: to be, or not to be random, that is the question, *Journal of Sound and Vibration* 329 (2010) 2809–2824.
[40] National Instruments, PCI extensions for Instrumentation (PXI). Available from: 〈http://www.ni.com/〉 (accessed 15.11.08).