inter.noise
HAMBURG 2016

Drive-by multi-frequency large-region noise-source mapping via tomographic imaging

Cagdas TUNA[1]; Shengkui ZHAO[2]; Thi Ngoc Tho NGUYEN[3]; Douglas L. JONES[4];

[1, 2, 3, 4] Advanced Digital Sciences Center (ADSC), Illinois at Singapore 138632, Singapore

[4] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

Urbana, Illinois 61820 USA

ABSTRACT

Recent studies indicate that noise pollution has hazardous impacts on human health and behavior, particularly becoming a serious problem in densely-populated cities. In order to effectively assess the environmental noise problem on a city-wide scale, there is a great need for a new low-cost environmental noise-monitoring system, which will replace conventional sound-pressure-level (SPL) meter-based techniques that require overly expensive microphone deployment at a dense grid of locations. In this paper, we propose an advanced cost-effective large-region environmental noise mapping scheme that uses a microphone array mounted on a vehicle driving around the city to make detailed acoustic noise maps without the need for point-by-point measurements. We demonstrate that the multi-frequency SPLs and locations of sparsely-distributed localized noise-sources in the neighborhood traversed by the vehicle are jointly estimated via tomography, in which the far-field delay-and-sum beamforming output power values computed at multiple locations and different times as the vehicle drives past are used for tomographic imaging. The preliminary drive-by noise-source mapping experiments with two concentric circular arrays rooftop-mounted on an electric vehicle show the promise of this new proposed technique for cost-effective city-wide large-scale environmental noise monitoring.

Keywords: Noise-mapping, Microphone-array processing, Acoustic imaging
I-INCE Classification of Subjects Number(s): 74.7, 74.8

1. INTRODUCTION

Noise pollution has become a major concern in urban areas, triggering serious health problems such as sleep disturbance, heart diseases, hearing loss and hypertension (1, 2). There is a need for an effective noise-monitoring scheme in terms of cost and time to identify urban noise hotspots for targeted noise abatement, since current sound-pressure-level (SPL) meter-based systems can only provide noise levels averaged over frequencies at the expense of extensive densely-located microphone measurements, which makes it prohibitively expensive on a city-wide scale.

Acoustic imaging with microphone directional arrays commonly used in aeroacoustic testing has the strong potential to be a cost-effective large-scale environmental noise monitoring solution (3, 4) as it becomes possible to generate acoustic maps of a region of interest at different frequencies, simultaneously identifying the dominant noise-source locations and their SPLs using acoustic beamforming. Advanced acoustic imaging algorithms including DAMAS, SC-DAMAS and CMF provide effective solutions for near-field mapping in relatively small regions (5, 6, 7).

In our recent work (8, 9), we have introduced a novel imaging model for large-region single-frequency acoustic noise-source mapping with a portable microphone array, collecting measurements at multiple locations and extending SC-DAMAS in to the multi-position data collection scenario so as to create a virtual large-aperture array. In this paper, we propose an alternative

[1] cagdas.t@adsc.com.sg
[2] Shengkui.zhao@adsc.com.sg
[3] tho.nguyen@adsc.com.sg
[4] jones@adsc.com.sg

large-scale noise-monitoring approach based on mounting a microphone array atop a vehicle to generate 2-D acoustic tomographic noise-maps of a neighborhood traversed by the vehicle. We demonstrate that the locations and SPLs of the acoustic localized noise-sources sparsely distributed over a large region can be estimated jointly at multiple frequencies using the far-field delay-and-sum beamforming output power values of the acoustic recordings at multiple locations as tomographic measurements. The preliminary acoustic imaging results with an electric vehicle with two rooftop-mounted circular microphone subarrays indicate that this new and practical sparse tomographic imaging scheme for large-scale noise-mapping offers a fundamentally new low-cost methodology for city-wide noise monitoring of urban noise hotspots.

2. PROBLEM FORMULATION

2.1 Forward Model

Consider an acoustic wave-field generated by multiple noise-sources being recorded by an *M*-channel microphone array. At a measurement location, the recorded and sampled signal at each microphone is divided into *L* snapshots followed by an *N*-length fast Fourier transform (FFT) applied to every *N*-sample snapshot, resulting in *N* narrow frequency bins. At the frequency $f \in \left[f_{\min}^{n}, f_{\max}^{n}\right]$ within the nth frequency bin, the linear additive-noise frequency-domain signal model for the *l*th snapshot can be given by (10)

$$\boldsymbol{z}_{f,l} = \mathbf{A}_f \boldsymbol{s}_{f,l} + \boldsymbol{e}_{f,l}, \tag{1}$$

where the vector $\boldsymbol{s}_{f,l} = \left[s_{f,l,\theta_i}, \dots, s_{f,l,\theta_{I-1}}\right]$ corresponds to the signals arriving from *I* different angular directions, the additive-noise vector $\boldsymbol{e}_{f,l}$ of length M is zero-mean and mutually uncorrelated with the signal, and $\mathbf{A}_f = \left[\boldsymbol{a}_{f,\theta_i}, \dots, \boldsymbol{a}_{f,\theta_{I-1}}\right]$ is the fixed $M \times I$ steering matrix for any measurement location with each column corresponding to the steering vector

$$\boldsymbol{a}_{f,\theta_i} = \left[1, e^{-j2\pi f d_{1,\theta_i}/c}, \dots, e^{-j2\pi f d_{M-1,\theta_i}/c}\right], \tag{2}$$

at a particular angle θ_i, where c denotes the speed of sound, and d_{m,θ_i} is the relative Euclidean distance of the microphone *m* to the plane wave at the reference microphone $m = 0$.

The sample covariance matrix $\widehat{\mathbf{R}}_f$ is computed by averaging over *L* snapshots as follows:

$$\widehat{\mathbf{R}}_f = \sum_{l=0}^{L-1} \boldsymbol{z}_{f,l} \boldsymbol{z}_{f,l}^H = \mathbf{A}_f \mathbf{P}_f \mathbf{A}_f^H + \mathbf{U}_f, \tag{3}$$

where $\mathbf{P}_f$ and $\mathbf{U}_f$ denote the $I \times I$ angular source and $M \times M$ noise sample covariance matrices, respectively.

The far-field delay-and-sum (DAS) beamformer weighting vector at a particular angle θ_i is given by $\boldsymbol{w}_{f,\theta_i} = \boldsymbol{a}_{f,\theta_i}/M$, yielding the DAS-beamformer output power at a given angular direction θ_i

$$y_{f,\theta_i} = \boldsymbol{w}_{f,\theta_i}^H \widehat{\mathbf{R}}_f \boldsymbol{w}_{f,\theta_i} = (\boldsymbol{a}_{f,\theta_i}^H \mathbf{A}_f \mathbf{P}_f \mathbf{A}_f^H \boldsymbol{a}_{f,\theta_i} + \boldsymbol{a}_{f,\theta_i}^H \mathbf{U}_f \boldsymbol{a}_{f,\theta_i})/M^2, \tag{4}$$

It is a fair assumption to neglect the cross terms in Eq. (4) for a sufficiently large number of snapshots (L>>1) (5, 8) such that the diagonal angular source and noise sample covariance matrices $\mathbf{P}_f = \mathbf{diag}\{p_{f,\theta_0}, \dots, p_{f,\theta_{I-1}}\}$ and $\mathbf{U}_f = \mathbf{diag}\{u_{f,0}, \dots, u_{f,M-1}\}$, respectively. The beamformer output power can then be rewritten as

$$y_{f,\theta_i} = \boldsymbol{b}_{f,\theta_i}^T \boldsymbol{p}_f + \sigma_f^2, \tag{5}$$

where $\boldsymbol{b}_{f,\theta_i} = \left[\left|\boldsymbol{a}_{f,\theta_i}^H \boldsymbol{a}_{f,\theta_0}\right|^2, \dots, \left|\boldsymbol{a}_{f,\theta_i}^H \boldsymbol{a}_{f,\theta_{I-1}}\right|^2\right]^T$ refers to the DAS power beam-pattern, and the angular power-level vector $\boldsymbol{p}_f = \left[p_{f,\theta_0}, \dots, p_{f,\theta_{I-1}}\right]^T$ has length *I*, and σ_f^2 is the ambient noise at the frequency *f*. Let $\boldsymbol{y}_f = \left[y_{f,\theta_0}, \dots, y_{f,\theta_{I-1}}\right]^T$ be the DAS-beamformer output power vector of length *I* and $\mathbf{B}_f = \left[\boldsymbol{b}_{f,\theta_0}, \dots, \boldsymbol{b}_{f,\theta_{I-1}}\right]^T$ be the power beam-pattern matrix of size $I \times I$ at the frequency *f*. Then, given an array position, we arrive at the set of linear equations

$$\boldsymbol{y}_f = \mathbf{B}_f \boldsymbol{p}_f + \sigma_f^2 \mathbf{1}, \tag{7}$$

where $\mathbf{1}$ refers to the vector of length *I* with all elements equal to 1.

2.2 Image Reconstruction

In this paper, we propose a two-step procedure for the multi-frequency recovery of the acoustic noise-sources, which includes an initial deconvolution step to estimate the actual angular power-levels $\boldsymbol{p}_f$, followed by tomographic reconstruction based on the multiple measurement vector (MMV) framework.

The Eq. (7) describes the convolution of the angular power-level vector $\boldsymbol{p}_f$ with the DAS beam-pattern

$\mathbf{B}_f$, resulting in blurring and some leakage from the sidelobes along with the additive ambient noise. There are deconvolution methods to estimate the direction-of-arrival (DOA) of sources using sparsity constraints (11, 12). In a similar fashion, under the assumption of sparsely distributed sources, we apply a regularized version of FOCUSS (focal underdetermined system solver) (13,14), an iteratively reweighted minimum-norm algorithm for the recovery of sparse DOAs along with the power-levels in the corresponding angular directions. For each iteration, the Tikhonov-regularized FOCUSS solves the optimization problem (13)

$$\widehat{\boldsymbol{p}}_f^{(c)} = \mathbf{argmin}_{\boldsymbol{p}_f} \left\| \boldsymbol{y}_f - \mathbf{B}_f \boldsymbol{p}_f \right\|_2^2 + \lambda_c \left\| \mathbf{W}_c^{-1} \boldsymbol{p}_f \right\|_2^2, \tag{7}$$

where λ_c is the regularization parameter at the iteration c, and the weighting matrix to enforce the weighted minimum norm constraint is set to be $\mathbf{W}_c = \mathbf{S}_c \mathbf{diag}(|\widehat{\boldsymbol{p}}_f^{(c-1)}|)$ with $\mathbf{S}_c$ being a smoothing constraint such as a 1-D Gaussian filter to increase robustness to noise. The regularized minimum-norm least-squares solution may be used as the initial low-resolution estimate for the first FOCUSS iteration.

Assuming that the acoustic field is generated by point noise-sources that are zero-mean and mutually uncorrelated, the power-level at each angle $\theta_{i'}$ may be approximated by

$$p_{f,\theta_{i'}} \approx \int_{\mathcal{R}} \frac{x(f,r,\theta_{i'})}{r^2}, \tag{6}$$

where $x(f,r,\theta_{i'})$ is the 2D unknown power-level imaging field, and $\mathcal{R} \subset \mathbb{R}^2$ is the region of support representing the angular path $\theta_{i'}$.

Given the estimated angular power values $\widehat{\boldsymbol{p}}_f$, the discretization of the imaging field x yields

$$\widehat{\boldsymbol{p}}_f = \mathbf{G}_{\theta,r} \boldsymbol{x}_f + \boldsymbol{v}_f, \tag{7}$$

where $\boldsymbol{x}_f = [x_{f,0}, \dots, x_{f,J-1}]^T$ of length J is the discretized unknown field at the frequency f, the linear matrix operator $\mathbf{G}_{\theta,r}$ of size $I \times J$ relates the Cartesian coordinates to the angular power measurements based on the weighting scheme inversely proportional to the Euclidean distance-squared from the measurement location, and $\boldsymbol{v}_f$ is the noise-vector of length I corresponding to the deconvolution error. This formulation resembles a typical tomographic measurement because the matrix operator $\mathbf{G}_{\theta,r}$ gives the total power-levels measured at an array position for particular angular directions computed along the related paths. Tomography is the recovery of a multi-dimensional signal from lower-dimensional projections collected by multiple sensors from different directions (15). Hence, if we record the acoustic data recorded at multiple locations, then it should become possible to estimate the 2D unknown acoustic field generated by sparsely distributed, localized noise-sources via tomographic reconstruction.

Since the spatially-fixed wideband acoustic noise-sources contain information over a range of frequencies, the joint estimation of the unknown stationary acoustic field at multiple frequencies may be formulated under the MMV framework, as the unknown field vectors $\boldsymbol{x}_f$ at all frequencies of interest should share a common sparsity support (16, 17). Therefore, given K different measurement locations, the overall problem can be expressed as

$$\begin{bmatrix} \widehat{\boldsymbol{p}}_{0,0} & \cdots & \widehat{\boldsymbol{p}}_{F-1,0} \\ \vdots & \ddots & \vdots \\ \widehat{\boldsymbol{p}}_{0,K-1} & \cdots & \widehat{\boldsymbol{p}}_{F-1,K-1} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{\theta,r}^{(0)} \\ \vdots \\ \mathbf{G}_{\theta,r}^{(K-1)} \end{bmatrix} [\boldsymbol{x}_0, \dots, \boldsymbol{x}_{F-1}] + \begin{bmatrix} \boldsymbol{v}_{0,0} & \cdots & \boldsymbol{v}_{F-1,0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{v}_{0,K-1} & \cdots & \boldsymbol{v}_{F-1,K-1} \end{bmatrix} \tag{9}$$

$$\widehat{\mathbf{P}} = \mathbf{G}\mathbf{X} + \mathbf{V} \tag{10}$$

where $\widehat{\mathbf{P}}$ is the concatenated angular power-level matrix with each column corresponding to the angular power vector for the measurement locations, 0 through *K-1* at a particular frequency, $\mathbf{G}$ is the stacked-up matrix operator generated based on K measurement locations, $\mathbf{X}$ is the unknown field matrix, and $\mathbf{V}$ is the additive noise matrix. We consider here the regularized M-FOCUSS, which is the generalized version of the FOCUSS algorithm based on the Frobenius norm:

$$\widehat{\mathbf{X}}^{(c)} = \mathbf{argmin}_{\mathbf{P}} \left\| \widehat{\mathbf{P}} - \mathbf{G}\mathbf{X} \right\|_F^2 + \lambda_c \left\| \mathbf{W}_c^{-1} \mathbf{X} \right\|_F^2, \tag{7}$$

where the weighting matrix $\mathbf{W}_c = \mathbf{S}_c \mathbf{diag}\left(\left\| \boldsymbol{x}_f^{(c)}[j] \right\| \right)$ with $\left\| \boldsymbol{x}_f^{(c)}[j] \right\| = \left(\sum_{f=0}^{F-1} (x_f^{(c)}[j])^2 \right)^{1/2}$ and $\mathbf{S}_c$ represents a prior constraint matrix such as spatial smoothness to suppress background noise (which may be achieved by using a 2D Gaussian filter), making a trade-off between sparseness and image quality, resulting in slightly less sparse images with noise-sources localized within several pixels. The regularized minimum Frobenius-norm least-squares solution may be used at the first iteration As the initial low-resolution estimate (15).

3. DRIVE-BY EXPERIMENTS

Figure 1 shows the experimental setup to test the proposed joint multi-frequency acoustic tomographic noise-mapping algorithm. The microphone array mounted atop the electric vehicle (EV) was composed of two concentric 12-channel circular subarrays with diameters of $d_0 = 0.5\,m$ and $d_1 = 1\,m$, respectively. Each microphone was pre-calibrated with sensitivity $\approx 50\,mV/Pa$ and fixed at 30^o separation from each other. The EV was driven along a straight road adjacent to a large open-field, and a 60-m portion was used for the beamformer output power computation. The data synchronization was carried out by two people shouting at 0-m and 60-m marks as the EV drove past. A computer-generated square wave with harmonics at 250 Hz, 750 Hz and 1250 Hz was played as the acoustic noise-source from a loudspeaker located 15 m vertically away from the 30-m mark. A near-omnidirectional sound propagation was realized by rotating the loudspeaker toward the EV travelling along the road.

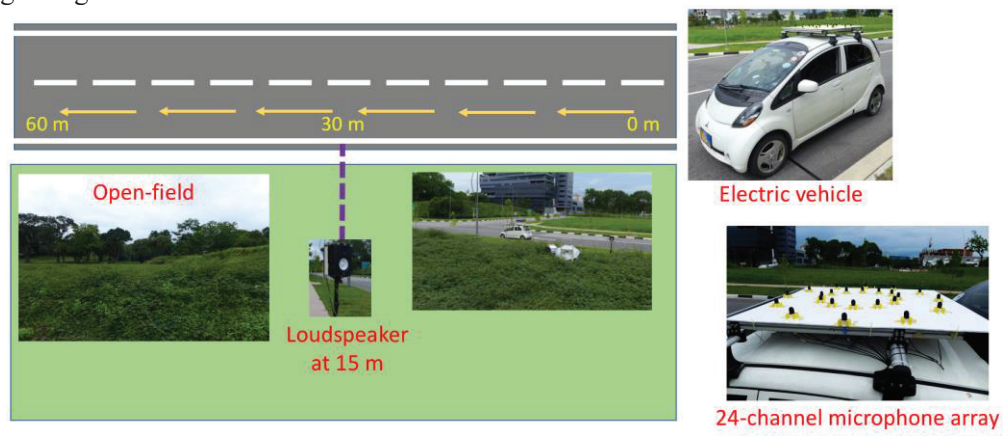

Figure 1 – The setup for drive-by acoustic noise-source tomographic mapping experiments

Figure 2 presents the joint multi-frequency acoustic imaging results for drive-by noise-source tomographic mapping at the average velocity of $24.83\,km/h$. The 24-channel acoustic data sampled at $f_s = 25.6\,kHz$ were split into smaller time segments, each referring to a specific array position along the road using the average EV velocity. The initial 0.5-sec portion of each recording was regarded for the sample covariance matrix computation. The FFT-length was $N = 512$ and a Hann window with an overlap of 50% was applied to each snapshot before the FFT.

Table 1 – The performance evaluation of tomographic imaging results for acoustic noise-source mapping

Frequency	Relative Residual Error	Measured-SPL dB(A)	Image-SPL dB(A)
250 Hz	0.6919	88.6	81.1
750 Hz	0.6895	87.4	85.7
1250 Hz	0.5831	84.8	82.4

The performance evaluation of the proposed method at the three harmonic frequencies of the square-wave noise-source is presented in Table 1 along with the relative residual error defined as the mismatch between the physical DAS-beamformer output powers and the output power estimated via the forward imaging model. The A-weighted measured SPL was calculated by collecting a separate set of SPL measurements stopping the EV every 10 m between 10-m and 50-m marks and taking a weighted-average over these measurements based on the distance-squared power attenuation. The A-weighted image-SPL was the total dB-SPL over the area visible in the reconstructed images obtained after applying a hard-thresholding at 60 dB(A) for illustration purposes. The noise-source location estimates determined as the peak locations in the images were within a few meters of error, and the dB-level difference between the average-SPL and image-SPL was less than 3 dB at 750 Hz and 1250 Hz, whereas the worst performance was achieved at 250 Hz, due the degraded performance of the

beamformer at low frequencies.

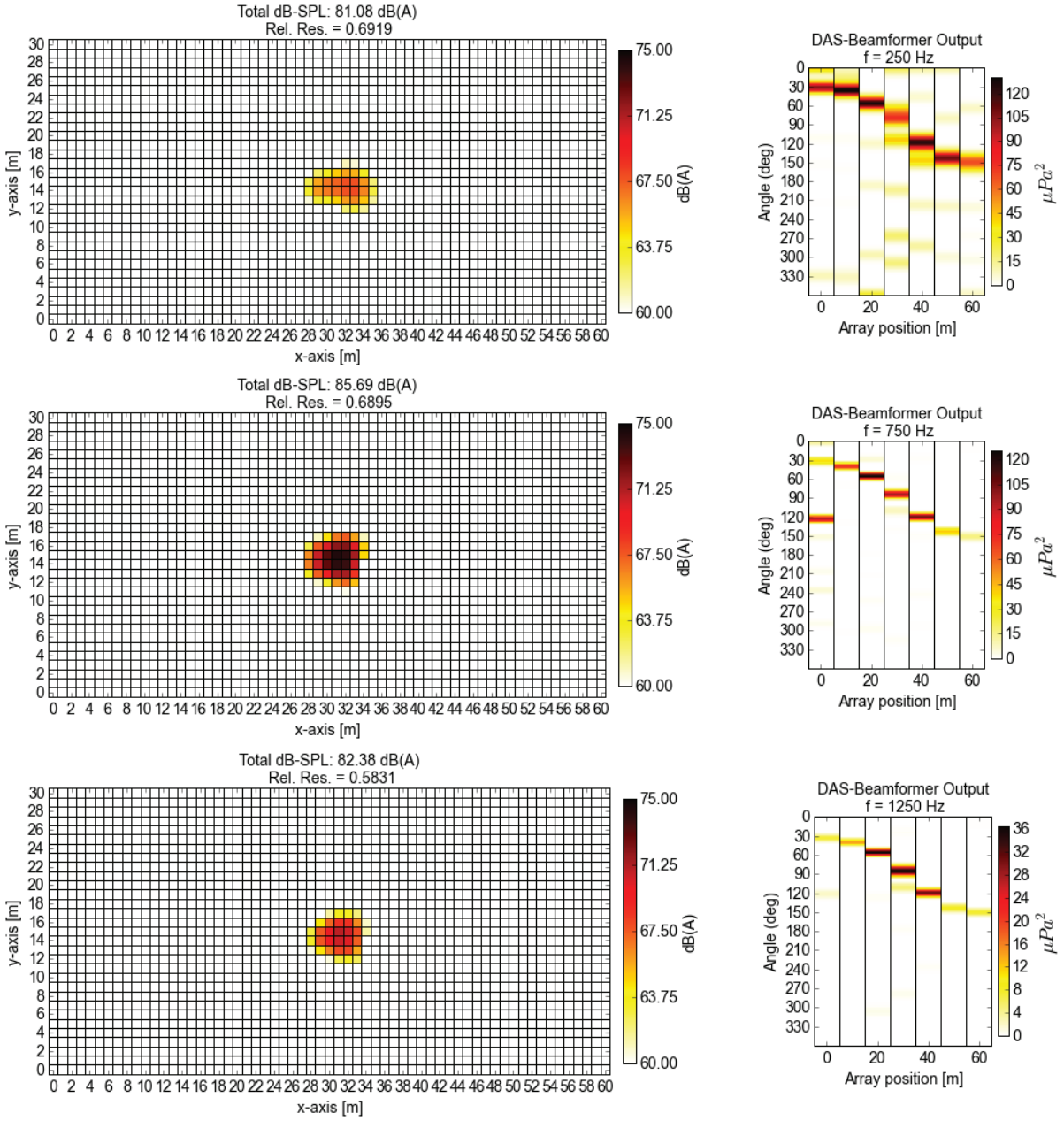

Figure 2 – The drive-by joint multi-frequency acoustic noise-source tomographic imaging results: Left column: The static dB-SPL images of the acoustic source-field. Right column: The DAS beamformer output powers computed with 1^o angular resolution at locations separated by 10 m.

The Doppler effect was insignificantly observed, as the frequency shift generally stayed within the same frequency band as a result of operating at a relatively slower speed and lower frequencies. On the contrary, the constructive/destructive ground reflections had a strong impact on the DAS-beamformer output power, particularly at 250 Hz as expected, further increasing the relative residual error. The collection of measurements at sufficiently large number of locations mitigates this issue by the imaging model formulation implicitly averaging over the measurements. Wind may also become a major problem, particularly at lower frequencies and high velocities, which may be avoided by high-quality wind screens.

4. CONCLUSIONS

The preliminary drive-by imaging experiments show that our proposed tomographic mapping approach can estimate the short-term average SPL and location of an acoustic noise-source jointly at

multiple frequencies with high accuracy. Environmental noise-monitoring on a city-wide scale may become very practical with this new alternative methodology, generating frequent maps of the neighborhoods traversed by the vehicle on a daily basis. Future work includes the acoustic imaging of real acoustic noise-sources possibly with a GPS-device for more precise recordings, and the development of a more advanced imaging model by accounting for other factors including acoustic reflections, ground and atmospheric absorptions, and variations in wind and temperature conditions.

ACKNOWLEDGEMENTS

This material is based on research/work supported in part by the Singapore Ministry of National Development and National Research Foundation under L2 NIC Award No. L2NICCFP1-2013-7 and in part by the research grant for the Human-Centered Cyber-physical Systems Programme at the Advanced Digital Sciences Center from Singapore's Agency for Science, Technology and Research (A*STAR).

REFERENCES

1. Alberti PW. Noise, the most ubiquitous pollutant, Noise Health. 1998. http://www.noiseandhealth.org/text.asp?1998/1/1/3/31783 (accessed 14 Apr 2016).
2. WHO Regional Office for Europe. Burden of disease from environmental noise. 2011. http://www.euro.who.int/__data/assets/pdf_file/0008/136466/e94888.pdf (accessed 14 Apr 2016).
3. Humphreys WM, Brooks TF, Hunter WW, Meadows KR. Design and use of microphone directional arrays for aeroacoustic measurements. Proc 36th AIAA Aerospace Sciences Meeting and Exhibit; 12-15 Jan 1998; Reno, NV 1998.
4. Lanslots J, Deblauwe F, Janssens K. Selecting sound source localization techniques for industrial applications. J Sound Vib. 2010; 44(6): 6-10.
5. Brooks TF, Humphreys WM. A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays. J Sound Vib. 2006;294(4):856-79.
6. Yardibi T, Li J, Stoica P, Cattafesta III LN. Sparsity constrained deconvolution approaches for acoustic source mapping, J Acoust Soc Am. 2008;123(5):2631-42.
7. Yardibi T, Li J, Stoica P, Zawodny NS, Cattafesta III LN. A covariance fitting approach for correlated acoustic source mapping, J Acoust Soc Am. 2010;127(5):2920-31.
8. Zhao S, Nguyen TNT, Jones DL. Large region acoustic source mapping using movable arrays. Proc 2015 IEEE ICASSP; 19-24 Apr 2015; Brisbane, Australia 2015. p. 2589-93.
9. Zhao S, Tuna C, Nguyen TNT, Jones DL. Large region acoustic source mapping: a generalized sparse constrained deconvolution approach. Proc 2016 IEEE ICASSP; 19-26 Mar 2016; Shanghai, China 2016.
10. Van Veen BD, Buckley KM. Beamforming: a versatile approach to spatial filtering. IEEE ASSP Mag. 5(2):4-24, (1998).
11. Fuchs JJ. On the application of the global matched filter to DOA estimation with uniform circular arrays. IEEE Trans Signal Process. 2001;49:702-9.
12. Wang Y, Leus G, Pandharipande A. Direction estimation using compressive sensing array processing. Proc 2009 IEEE/SP 15th Workshop on Statistical Signal Processing; 31 Aug-3 Sep 2009; Cardiff, United Kingdom 2009. p. 626-9.
13. Gorodnitsky I, Rao BD. Sparse signal reconstruction from limited data using FOCUSS: a reweighted minimum norm algorithm. IEEE Trans Signal Process. 1997; 45:600-16.
14. Rao BD, Engan K, Cotter S, Palmer J, Kreutz-Delgado K. Subset selection in noise based on diversity measure minimization. IEEE Trans Signal Process. 2003;51:760-70.
15. Groetsch CW. Inverse Problems in the Mathematical Sciences. Braunschweig/Wiesbaden, Germany: Vieweg; 1993.
16. Cotter SF, Rao BD, Engan K, Kreutz-Delgado K. Sparse solutions to linear inverse problems with multiple measurement vectors. IEEE Trans Signal Process. 2005;53(7):2477-88.
17. Zdunek R, Cichocki A. Improved M-FOCUSS algorithm with overlapping blocks for locally smooth sparse signals. IEEE Trans Signal Process. 2008;56(10):4752-61.