Journal of Sound and Vibration 385 (2016) 93–103

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Moving sound source localization based on triangulation method

Feng Miao, Diange Yang*, Junjie Wen, Xiaomin Lian

State Key Laboratory of Automotive Safety and Energy, Department of Automotive Engineering, Collaborative Innovation Center of Intelligent New Energy Vehicle, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form
2 September 2016
Accepted 2 September 2016
Handling Editor: R.E. Musafir
Available online 20 September 2016

ABSTRACT

This study develops a sound source localization method that extends traditional triangulation to moving sources. First, the possible sound source locating plane is scanned. Secondly, for each hypothetical source location in this possible plane, the Doppler effect is removed through the integration of sound pressure. Taking advantage of the de-Dopplerized signals, the moving time difference of arrival (MTDOA) is calculated, and the sound source is located based on triangulation. Thirdly, the estimated sound source location is compared to the original hypothetical location and the deviations are recorded. Because the real sound source location leads to zero deviation, the sound source can be finally located by minimizing the deviation matrix. Simulations have shown the superiority of MTDOA method over traditional triangulation in case of moving sound sources. The MTDOA method can be used to locate moving sound sources with as high resolution as DAMAS beamforming, as shown in the experiments, offering thus a new method for locating moving sound sources.

1. Introduction

Sound source localization is a practical research topic with many applications. Many methods have been developed to solve this problem, including beamforming, acoustic holography, and the triangulation method. The conventional beamforming method applies delay-and-sum to signals from an array of microphones, forming a beam peak in the direction of the source [1]. Since 1988, the beamforming method combined with Doppler effect removal has been used for localization of moving noise sources associated with aircraft, trains and cars [2]. For example, Stoker et al. determined Boeing 777 airframe noise from fly-over tests with a 187 microphone array and compared the results with earlier wind tunnel measurements on scale models [3]. Guerin et al. used a 168 microphone array to test the fly-over noise of an Airbus A319 aircraft [4,5]. Zechel et al. used a 32 microphone array to locate noise sources from a EuroCity train [6]. Sarradj et al. used 64 microphones to locate aeroacoustic sources from a 1:20 scale model of a high-speed train [7]. Yang et al. used 30 microphones to locate loudspeakers on a moving car [8]. Combined with deconvolution [9] or inverse methods, beamforming has also been widely used for an absolute source level estimation in many applications [5,10–12].

Unlike beamforming, the acoustic holography method is used to solve the inverse propagation problem and reconstruct acoustic fields, including sound pressure, particle velocity and intensity field. The near-field acoustic holography (NAH) was first proposed by Maynard et al. [13,14] and later widely applied and improved. The researchers have developed various

* Corresponding author.
E-mail address: ydg@mail.tsinghua.edu.cn (D. Yang).

http://dx.doi.org/10.1016/j.jsv.2016.09.001
0022-460X/

ways to locate moving sound sources using the acoustic holography method. In 1998, Nakagawa et al. used acoustic holography considering Doppler effect to detect the tire noise of a moving vehicle with an 8×8 microphone array [15]. Kwon and Kim developed the moving frame acoustical holography method by assuming that the sound source is stable, while the microphone frame sweeps over it [16]. Later, Park and Kim applied the method for locating moving sound sources on a vehicle with a 16 microphone linear array [17,18]. Yang et al. proposed quantitative analyses of pass-by noise from a high-speed vehicle with a 16 microphone cross array [19].

Normally, studies based on beamforming and acoustic holography approaches mentioned above are used to locate and quantify distributed noise sources, which means they need many microphones to produce a satisfactory map. However, in such situations, when simple sound source tracking is necessary instead of complicated sound field reconstruction, the triangulation method might be a better approach. The triangulation method depends on the time difference of arrival (TDOA) calculation and locates the sound source using geometric relationships between microphones and the sound source. Because one TDOA can be obtained using only two microphones, the sound source can be located in space with at least four microphones. The triangulation method has been widely used for real-time sound source localization with a small number of sensors. For example, Ferguson et al. used only 3 sensors to determine the bearing angle of a firing artillery gun [20]. Giraudet and Glotin used 5 omni-directional hydrophones to track marine mammals [21]. Wu and Zhu used the triangulation method to locate an arbitrary sound source in space in real time with four microphones [22]. Later, these researchers proposed a redundant version of the method using 6 microphones to improve the localization results [23,24]. However, because the Doppler effect leads to changes of time delays at microphones, the traditional triangulation method cannot be applied directly to locate moving sound sources. A method that combines TDOA and frequency difference of arrival (FDOA) has been used to locate single-frequency moving emitting sources [25]. The TDOA–FDOA method works particularly well for locating aircraft, for which the emitting frequency was preset. However, because single-frequency sound sources are rare in practice, it is difficult to locate an unknown moving sound source with the TDOA–FDOA method.

This paper proposes a moving sound source localization method based on triangulation. This method involves the following sequence of steps. First, a possible locating plane of the sound source is scanned. Secondly, the non-simplified Doppler effect removal is applied to each hypothetical source location in this possible plane based on the acoustic theory proposed by Morse [26]. The moving time difference of arrival (MTDOA) is calculated to estimate the source location based on triangulation. Thirdly, the deviation from the source location estimation to the original hypothetical source location is recorded. Because the real sound source location leads to zero deviation, the sound source of interest can be finally located by minimizing the deviation of its position. Our simulations show that the proposed method can be used to locate moving sound sources. Additionally, this theory validated our experiments with locating a loudspeaker attached to a moving vehicle for both single-frequency and white noise signals.

2. The moving sound source localization method

2.1. The moving sound source localization model

Fig. 1 shows the proposed moving sound source localization model. The microphones are located at the fixed positions, and the vector from the origin O to the ith microphone m_i is defined as $\mathbf{m}_i = (x_{m_i},\ y_{m_i},\ z_{m_i})$. The sound source s moves linearly with a known speed v, which is parallel to the x-axis. The vector from the origin O to s is $\mathbf{s}(t) = (x_s(t),\ y_s,\ z_s)$. Because the method attempts to determine the location of the sound source s, a possible locating plane is defined as $\mathbf{S}$. The plane $\mathbf{S}$ is scanned, and each grid location in $\mathbf{S}$ is considered to be a hypothetical sound source location. For the kth hypothetical location g_k, the vector from the origin O to g_k is $\mathbf{g}_k(t) = (x_{g_k}(t),\ y_{g_k},\ z_{g_k})$.

In the triangulation method, signals from every two microphones can be used to produce one TDOA value. Therefore, locating a sound source in the 2-dimensional plane $\mathbf{S}$ requires at least three microphones. In this study, we consider a microphone array with four microphones to improve the robustness.

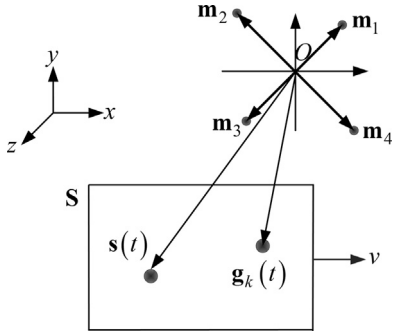

Fig. 1. The schematic of moving sound source localization.

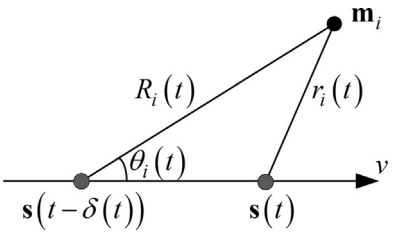

Fig. 2. The geometry of Doppler effect.

2.2. Moving sound source localization based on triangulation

First, the Doppler effect has to be removed from the microphone signals. As shown in Fig. 2 [26], at time t, the moving sound source is located at $\mathbf{s}(t)$. The distance between $\mathbf{s}(t)$ and $\mathbf{m}_i$ is $r_i(t)$. The sound wave reaching microphone m_i at time t was emitted when the sound source was at $\mathbf{s}(t-\delta(t))$, where $\delta(t)=R_i(t)/c$, c is the sound speed. $R_i(t)$ is the distance between $\mathbf{s}(t-\delta(t))$and $\mathbf{m}_i$, which can be written as

$$R_i(t)=\frac{M(x_{m_i}-x_s(t))+\sqrt{(x_{m_i}-x_s(t))^2+(1-M^2)\left[(y_{m_i}-y_s)^2+(z_{m_i}-z_s)^2\right]}}{1-M^2} \tag{1}$$

where $M=\dfrac{v}{c}$ is the Mach number.

The angle between the sound source trajectory and $\mathbf{m}_i-\mathbf{s}(t-\delta(t))$ is $\theta_i(t)$, which can be written as

$$\theta_i(t)=\arccos\frac{x_{m_i}-x_s(t)+MR_i(t)}{R_i(t)} \tag{2}$$

The velocity potential at the position of microphone m_i can be described as

$$\Psi_i(t)=\frac{1}{\rho}\frac{q(t-R_i(t)/c)}{4\pi R_i(t)(1-M\ \cos\ \theta_i(t))} \tag{3}$$

where ρ is the medium density, and q is the mass flow rate. The sound source is assumed as a monopole.

The sound pressure can be described as a time derivative of the velocity potential with the following expression

$$p_i(t)=-\rho\frac{\partial\Psi_i}{\partial t}=-\frac{\partial}{\partial t}\left[\frac{q(t-R_i(t)/c)}{4\pi R_i(t)(1-M\ \cos\ \theta_i(t))}\right] \tag{4}$$

In order to isolate the term $q'(t-R_i(t)/c)$ so that the pressure can be written in the same form as for static sound sources, a time integration and a time derivative need to be done, as will be shown below. Assuming $\Psi_i(0)=0$, the time integration of both parts of Eq. (4) is

$$\int_0^t p_i(\tau)\mathrm{d}\tau=-\frac{q(t-R_i(t)/c)}{4\pi R_i(t)(1-M\ \cos\ \theta_i(t))} \tag{5}$$

which is equal to

$$q(t-R_i(t)/c)=-4\pi R_i(t)(1-M\ \cos\ \theta_i(t))\int_0^t p_i(\tau)\mathrm{d}\tau \tag{6}$$

The time derivative of both parts of Eq. (6) can be expressed as

$$\begin{aligned}q'(t-R_i(t)/c)\frac{\partial(t-R_i(t)/c)}{\partial t}&=-4\pi R_i(t)(1-M\ \cos\ \theta_i(t))p_i(t)\\&\quad-4\pi\frac{\partial[R_i(t)(1-M\ \cos\ \theta_i(t))]}{\partial t}\int_0^t p_i(\tau)\mathrm{d}\tau\end{aligned} \tag{7}$$

In this equation, the time derivative of $R_i\ (t)$ can be written as

$$\frac{\partial R_i(t)}{\partial t}=\frac{-v\ \cos\ \theta_i(t)}{1-M\ \cos\ \theta_i(t)} \tag{8}$$

and

$$\frac{\partial}{\partial t}[R_i(t)(1-M\ \cos\ \theta_i(t))] = \frac{\partial}{\partial t}\left[\left(1-M^2\right)R_i(t) - M\left(x_{m_i}-x_s(t)\right)\right] = \frac{-v(\cos\ \theta_i(t)-M)}{1-M\ \cos\ \theta_i(t)} \tag{9}$$

Therefore, Eq. (7) is equal to

$$q'\left(t-R_i(t)/c\right) = -4\pi\left[R_i(t)(1-M\ \cos\ \theta_i(t))^2 p_i(t) + (M-\ \cos\ \theta_i(t))v\int_0^t p_i(\tau)\mathrm{d}\tau\right] \tag{10}$$

If the sound source were static and the distance between sound source and microphone were constant, i.e., $r_i(t) = r_i(t_0),\ \ \forall t$, the sound pressure at m_i can be described as

$$\overline{p}_i(t) = -\frac{q'\left(t-r_i(t_0)/c\right)}{4\pi r_i(t_0)} \tag{11}$$

Converting the sound pressure signals generated by the moving sound source to signals generated by the assumed static sound source, we obtain the Doppler effect removal as

$$\tilde{p}_i(t) = \frac{R_i(t)(1-M\ \cos\ \theta_i(t))^2 p_i(t) + (M-\ \cos\ \theta_i(t))v\int_0^t p_i(\tau)\mathrm{d}\tau}{r_i(t_0)}\ \cdot o\left(\frac{R_i(t)-r_i(t_0)}{c}\right) \tag{12}$$

where $o(\xi)$ is a time delay operator such that $o(\xi)\cdot f(t) = f(t)\cdot o(\xi) = f(t+\xi)$. The delay operator can be processed by interpolation in the time domain. The Doppler effect removed signal $\tilde{p}_i(t)$ indicates as if the sound source is static and located at $\mathbf{g}_k(t_0)$.

Because the method attempts to determine the real sound source location, the Doppler effect removal process can only be performed for the hypothetical source locations. For $\mathbf{g}_k(t)$, the Doppler effect removal is accomplished using

$$\tilde{p}_i\left(t,\mathbf{g}_k(t)\right) = \frac{R_i\left(t,\mathbf{g}_k(t)\right)\left(1-M\ \cos\ \theta_i\left(t,\mathbf{g}_k(t)\right)\right)^2 p_i(t)}{r_i\left(t_0,\mathbf{g}_k(t_0)\right)}\cdot o\left(\frac{R_i\left(t,\mathbf{g}_k(t)\right)-r_i\left(t_0,\mathbf{g}_k(t_0)\right)}{c}\right) + \frac{\left(M-\ \cos\ \theta_i\left(t,\mathbf{g}_k(t)\right)\right)v\int_0^t p_i(\tau)\mathrm{d}\tau}{r_i\left(t_0,\mathbf{g}_k(t_0)\right)}\cdot o\left(\frac{R_i\left(t,\mathbf{g}_k(t)\right)-r_i\left(t_0,\mathbf{g}_k(t_0)\right)}{c}\right) \tag{13}$$

where $r_i\left(t_0,\mathbf{g}_k(t_0)\right)$ is the distance between $\mathbf{m}_i$ and $\mathbf{g}_k(t_0)$. $R_i\left(t,\mathbf{g}_k(t)\right)$ and $\theta_i\left(t,\mathbf{g}_k(t)\right)$ can be written as

$$R_i\left(t,\mathbf{g}_k(t)\right) = \frac{M\left(x_{m_i}-x_{g_k}(t)\right) + \sqrt{\left(x_{m_i}-x_{g_k}(t)\right)^2 + \left(1-M^2\right)\left[\left(y_{m_i}-y_{g_k}\right)^2 + \left(z_{m_i}-z_{g_k}\right)^2\right]}}{1-M^2} \tag{14}$$

$$\theta_i\left(t,\mathbf{g}_k(t)\right) = \arccos\frac{x_{m_i}-x_{g_k}(t)+MR_i\left(t,\mathbf{g}_k(t)\right)}{R_i\left(t,\mathbf{g}_k(t)\right)} \tag{15}$$

Using the signals after the Doppler effect removal, the MTDOA can be calculated with the generalized cross correlation (GCC) [27] or the cross-power spectral phase (CSP) [28] method. After defining the function $T[x(t), x(t+\Delta t)] = \Delta t$, and $\Delta t_{ij,k}$ as the MTDOA between microphone m_i and m_j for $\mathbf{g}_k(t)$, there exists

$$\Delta t_{ij,k} = T\left[\tilde{p}_i\left(t,\mathbf{g}_k(t)\right), \tilde{p}_j\left(t,\mathbf{g}_k(t)\right)\right]\quad (i=1\sim 4,\ \ j=1\sim 4,\ \ i\neq j, k=1\sim K) \tag{16}$$

The MTDOA can be used to perform the triangulation

$$\left\|\hat{\mathbf{g}}_k-\mathbf{m}_j\right\| = \left\|\hat{\mathbf{g}}_k-\mathbf{m}_i\right\| + c\Delta t_{ij,k} \tag{17}$$

where $\hat{\mathbf{g}}_k$ is the sound source location estimation for time t_0, $\hat{\mathbf{g}}_k = \left(\hat{x}_{g_k}, \hat{y}_{g_k}, \hat{z}_{g_k}\right)$. The symbol $\|\ \cdot\ \|$ represents the length of a vector.

For a spatial four microphone array, Eq. (17) can be divided into 3 independent three-variable quadratic equations. The solutions have been derived in the study [22], which demonstrates that there are always two real roots, leading to two source locations; therefore, additional data processing is necessary to discard the false location. For a planar array, however, the two roots would locate on two sides of the plane ($\hat{z}_{g_k}$ is either plus or minus the square root in Eq. (18)), so the ambiguity would become trivial. Thus solutions for planar arrays are derived in this paper. Assuming $z_{m_i} = 0\quad (i=1\sim 4)$, Eq. (17) can be simplified into linear equations [29]. Choose m_1 as the reference microphone, i.e., $\Delta t_{12,k}$, $\Delta t_{13,k}$, and $\Delta t_{14,k}$ are calculated, the three coordinates can be solved as

$$\hat{x}_{g_k} = \frac{b_{2,k}-b_{1,k}}{a_{1,k}-a_{2,k}}$$

$$\hat{y}_{g_k} = \frac{a_{1,k}b_{2,k}-a_{2,k}b_{1,k}}{a_{1,k}-a_{2,k}}$$

$$\hat{z}_{g_k} = \sqrt{\left(\frac{2(x_{m_1}-x_{m_2})\hat{x}_{g_k}+2(y_{m_1}-y_{m_2})\hat{y}_{g_k}+e_k}{2c\Delta t_{12,k}}\right)^2 - (\hat{x}_{g_k}-x_{m_1})^2 - \left(\hat{y}_{g_k}-y_{m_1}\right)^2} \tag{18}$$

where

$$a_{1,k} = \frac{(x_{m_1}-x_{m_3})\Delta t_{12,k} - (x_{m_1}-x_{m_2})\Delta t_{13,k}}{(y_{m_1}-y_{m_2})\Delta t_{13,k} - (y_{m_1}-y_{m_3})\Delta t_{12,k}}$$

$$a_{2,k} = \frac{(x_{m_1}-x_{m_4})\Delta t_{12,k} - (x_{m_1}-x_{m_2})\Delta t_{14,k}}{(y_{m_1}-y_{m2})\Delta t_{14,k} - (y_{m_1}-y_{m_4})\Delta t_{12,k}}$$

$$b_{1,k} = \frac{c_3\Delta t_{12,k} - c_2\Delta t_{13,k}}{2\left[(y_{m_1}-y_{m_2})\Delta t_{13,k} - (y_{m_1}-y_{m_3})\Delta t_{12,k}\right]}$$

$$b_{2,k} = \frac{c_4\Delta t_{12,k} - c_2\Delta t_{14,k}}{2\left[(y_{m_1}-y_{m_2})\Delta t_{14,k} - (y_{m_1}-y_{m_4})\Delta t_{12,k}\right]}$$

$$e_k = x_{m_2}^2 + y_{m_2}^2 - x_{m_1}^2 - y_{m_1}^2 - (c\Delta t_{12,k})^2$$

The normalized deviation at $\mathbf{g}_k$ is defined as

$$D(\mathbf{g}_k) = \frac{\|\hat{\mathbf{g}}_k - \mathbf{g}_k(t_0)\|}{\|\mathbf{g}_k(t_0)\|} \times 100\% \tag{19}$$

If $\mathbf{g}_k(t) = \mathbf{s}(t)$, $\tilde{p}_i(t, \mathbf{g}_k(t)) = \tilde{p}_i(t)$, the Doppler effect removal is correct, the MTDOAs are correct, and the location estimation $\hat{\mathbf{g}}_k$ is the same as $\mathbf{g}_k(t_0)$, thus $D(\mathbf{g}_k = \mathbf{s}) = 0$. Otherwise, $D(\mathbf{g}_k \neq \mathbf{s}) \neq 0$.

In practice, the plane $\mathbf{S}$ is meshed into discrete grids, thus the sound source location at time t_0 can be determined by finding the minimum value of D

$$D(\hat{\mathbf{s}}) = \min\left[D(\mathbf{g}_k)\right] \tag{20}$$

2.3. Array discussion

As we stated in Sections 2.1 and 2.2, at least three microphones were required to locate a sound source in a plane with the triangulation method. A fourth redundant microphone improves the robustness of the procedure. For a spatial four-microphone array, there are always two source locations, therefore, additional data processing is necessary to discard the false location. In this study, to avoid the source location ambiguity problem, we use a planar four-microphone array.

Starting from Eq. (18), if four microphones are located in a line, the parameter $a_{1,k} = a_{2,k}$, and, thus, the source location is unsolvable. Therefore, the planar array discussed in this study is non-linear.

For different source types, there are different array design principles. Define d as the maximum distance between all 3 microphones and the reference microphone. For single-frequency sound source localization, the distances between all 3 microphones and the reference microphone should be smaller than half wavelength, i.e., $d < \frac{\lambda}{2}$, to avoid phase wrapping. For wide band or impulsive sound sources, phase wrapping might not be a problem, therefore, the array can be designed in any shape, as long as all 4 microphones can receive signal from the same sound source. In conditions when the sound source frequency characteristics are unknown, the sound pressure signals captured by the microphones must go through a low-pass filter before data processing. The cutoff frequency is $f_c = \frac{c}{2d}$.

3. Simulations

Simulations have been designed to compare TDOA and MTDOA methods in case of moving sound sources. It should be noted that TDOA is computed in the 3-D space, while MTDOA calculation is limited to the z_s plane ($\forall \mathbf{g}_k, \quad z_{g_k} = z_s$). We used two common types of sound sources, single-frequency and white noise sources.

3.1. Single-frequency sound source localization

In this simulation, the source generates sound of 1000 Hz. As stated in Section 2.3, the distances between all 3 microphones and the reference microphone should be smaller than half wavelength, which is 0.17 m. A cross array was designed as shown in Fig. 3. The coordinates of the microphones are $\mathbf{m}_1$ = (0.05, 0.05, 0) m, $\mathbf{m}_2$ = (−0.1, 0.1, 0) m, $\mathbf{m}_3$ = (−0.05, −0.05, 0) m, $\mathbf{m}_4$ = (0.1, −0.1, 0) m. m_1 was chosen as the reference microphone.

The sound source moves with a speed of 20 m/s along the x-direction. At time t_0, the sound source is located at (0, 0, 5) m. Independent background white noises with signal-to-noise ratio SNR=20 dB were added to signals received by the microphones. The sampling frequency is 8192 Hz. The sampling length is 0.1 s.

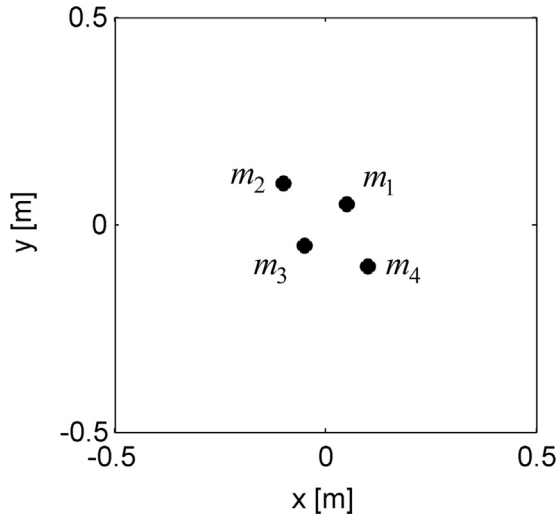

Fig. 3. The microphone array for simulations.

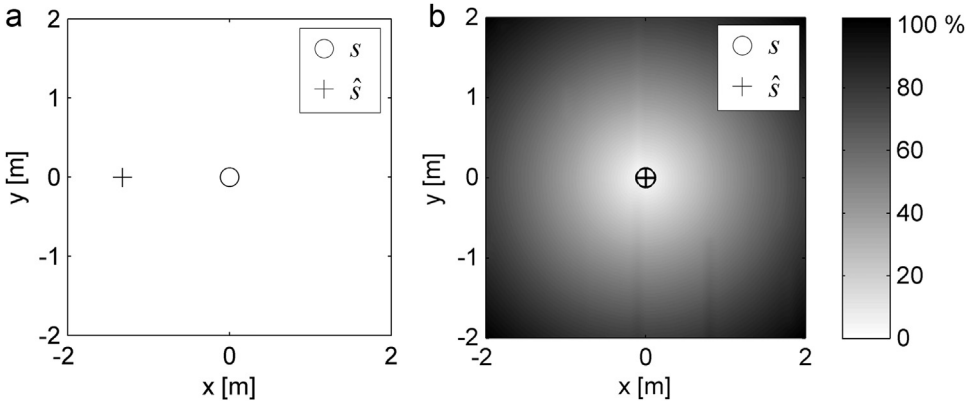

Fig. 4. 1000 Hz sound source localization results: (a) the traditional triangulation method; (b) the triangulation-based moving sound source localization method.

Fig. 4(a) shows the result of the traditional triangulation sound source localization. The real sound source s at t_0 is marked as a circle at (0, 0, 5) m. Because the sound source movement has led to a change in time delay, the traditional triangulation localization result $\hat{s}$ is incorrect, marked with a cross at (-1.32, 0.01, 5.23) m. We define localization error as $e=\frac{\|\hat{\mathbf{s}}-\mathbf{s}(t_0)\|}{\|\mathbf{s}(t_0)\|}\times 100\%$. The localization error in this case is 26.8%.

Fig. 4(b) shows the triangulation result based on moving sound source localization approach. The real location of the sound source is also marked with a circle at (0, 0, 5) m. The possible locating plane at $z=5$ m was meshed into grids of 0.1 m $\times$ 0.1 m. The deviation at each grid is represented by a grey scale from black to white, corresponding to the deviation value from 100% to 0%. The localization result is determined by identifying the minimum deviation location, which is marked as a cross at (0, 0, 5) m. The localization error in this case is 0.

3.2. White noise sound source localization

In this simulation, the sound source generates white noise. The data processing is based on band filtered signals in the 1 kHz third-octave band. As stated in Section 2.3, the microphone array can have any shape. For simplicity, the same array as the one shown in Fig. 3 was used for this sound source localization example. Microphone m_1 was chosen as the reference microphone.

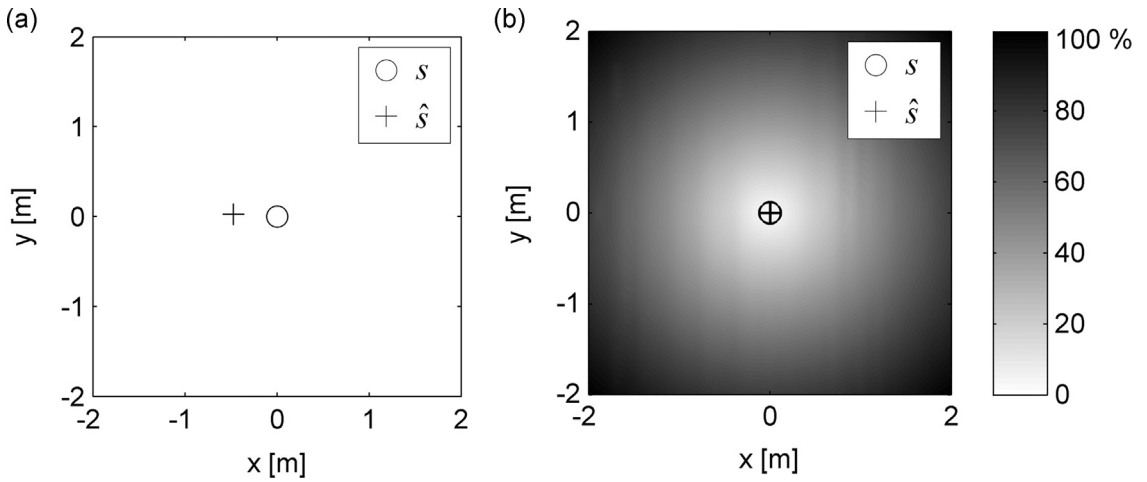

Fig. 5. White noise sound source localization results: (a) the traditional triangulation method; (b) the triangulation-based moving sound source localization method.

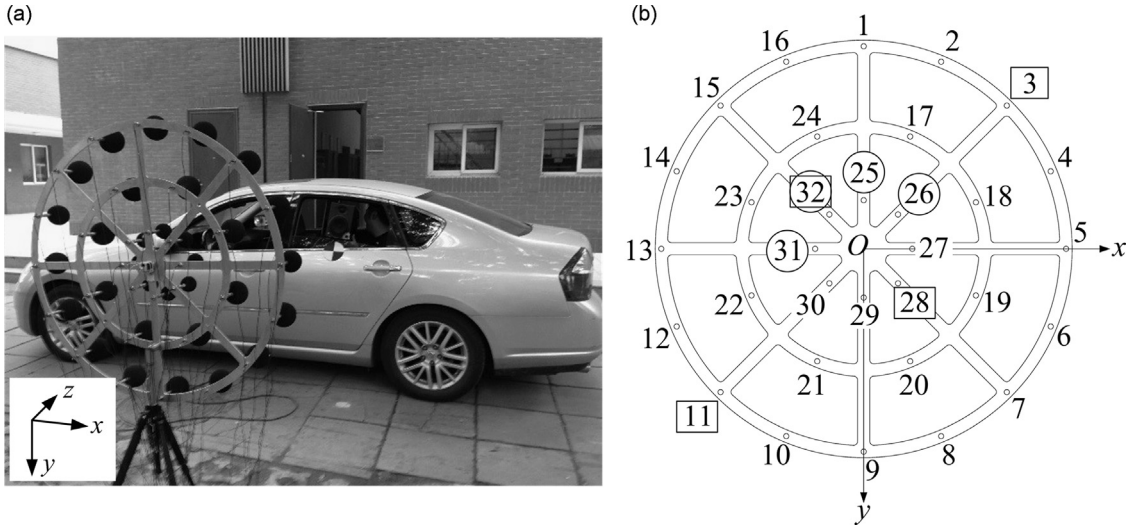

Fig. 6. Experimental setup: (a) experiment photo; (b) microphone array.

The sound source moves with a speed of 20 m/s along the x-direction. At time t_0, the sound source is located at (0, 0, 5) m. Independent background white noises with SNR = 20 dB were added to signals received by the microphones. The sampling frequency is 8192 Hz. The sampling length is 0.1 s.

Fig. 5(a) shows the result obtained with the traditional triangulation sound source localization method. The real sound source s location at t_0 is marked with a circle at (0, 0, 5) m. The localization result $\hat{s}$ is marked with a cross at (−0.47, 0.03, 9.85) m. The localization error in this case is 97.5%.

Fig. 5(b) shows the result obtained with the triangulation-based moving sound source localization approach. The possible plane at $z = 5$ m was meshed into grids of 0.1 m × 0.1 m. The deviation at each grid is represented by a grey scale from black to white, corresponding to the deviation value from 100% to 0%. The localization result is determined by finding the minimum deviation location, which is marked with a cross at (0, 0, 5) m. The localization error in this case is 0.

These simulations show that the traditional triangulation method cannot be used for moving sound source localization. Conversely, the proposed method can locate both single-frequency and wide-band sound sources. We demonstrated that the localization results were accurate for the cases with a small background noise.

4. Experiments

We have conducted experiments to validate the proposed moving sound source localization method. The experiment setup photo is shown in Fig. 6(a). A loudspeaker was placed in a moving car near the left rear side window, generating

1000 Hz tone and white noise in different cases. A cross-shaped marker was pasted onto the left rear door just below the loudspeaker. The car moved linearly with a speed of 30 km/h. A planar microphone array was set up 3.7 m from the car. A camera was placed at the center of the array for video recording. The array center is at the origin. The x-axis is horizontal, parallel to the line along which the car moved, the y-axis is vertical, pointing downward to the ground, and the z-axis is horizontal, pointing toward the car.

Fig. 6(b) shows the microphone array in detail. The microphones were placed on three concentric circles. Diameters of the three circles are 0.24 m, 0.6 m, and 1 m. There were 32 microphones in total. The microphone Numbers 1 to 16 were located uniformly on the largest circle, Numbers 17 to 24 were on the medium circle, and Numbers 25 to 32 were on the smallest circle. As stated in Section 2.3, for a single-frequency sound source, the distances between 3 microphones and the reference microphone in the triangulation method must be smaller than half wavelength, which is 0.17 m for 1000 Hz. Therefore, microphones 25, 26, 31 and 32 were chosen as the triangulation array for 1000 Hz sound, in which the microphone 25 was the reference microphone. The corresponding microphone numbers are marked with circles in Fig. 6(b). For white noise, the microphone array for the triangulation method can have any shape. Considering the beamforming method, microphones 32, 28, 11 and 3 were selected, among which the microphones 32 and 28 were on the smallest circle, to guarantee a large spatial sampling frequency, and the microphones 11 and 3 were on the largest circle, to obtain a large array aperture. The microphone 32 was the reference microphone. The corresponding microphone numbers are marked with rectangles in Fig. 6(b).

Fig. 7 shows localization results for loudspeaker generating single-frequency sound, and Fig. 8 shows localization results for loudspeaker generating white noise. The beamforming analyzing frequency for Fig. 7 is 1000 Hz, and that for Fig. 8 is 1000 Hz third octave band.

Figs. 7(a) and 8(a) show the de-Dopplerized beamforming results using signals from all 32 channels. They are considered as accurate localization references. The contour lines show the sound pressure levels on the reconstruction plane, which was 3.7 m from the microphone array. The sound source is located at the rear side window.

Figs. 7(b) and 8(b) show the de-Dopplerized beamforming results using signals from the four chosen channels. For Fig. 7(b), the four channels are 25, 26, 31 and 32. The microphone array aperture is $d = 0.22$ m, and the Rayleigh radius is $r = 0.61\frac{\lambda z}{d}$, which is equal to 3.49 m. The main lobe area in beamforming is so large that it is difficult to tell the exact source location. For Fig. 8(b), the four channels are 32, 28, 11 and 3. As the four microphones are located at vertexes of a diamond, the shapes of the main lobe and side lobes are somehow like diamonds. The highest side lobe is only 1 dB lower than the main lobe, which may lead to false localization problem. Traditional beamforming results using only four microphones is not satisfactory.

In Figs. 7(c) and 8(c) DAMAS deconvolution [9] has been extended to moving sources using the present de-Dopplerization method. Beamforming results are deduced from the signals of the four chosen channels. The wide main lobe and high side lobe problems are solved by this deconvolution process. The localization results are accurate points.

Figs. 7(d) and 8(d) show the traditional triangulation localization results. The localization results are marked with white crosses. Because the Doppler effect has led to a change in time delay, the localizations point to wrong positions. In Fig. 7(d), the localization result is (1.57, −1.98, 5.37) m. In Fig. 8(d), the localization result is (−3.63, −12.29, 41.05) m.

Figs. 7(e) and 8(e) show the MTDOA results. The localization results are marked with black crosses. The grey-scale map shows deviations at the grids from 30% to 0%, corresponding to a grey scale from black to white. The sound source was located by finding the minimum position in the deviation matrix. In Fig. 7(e), the localization result is (0.83, −0.18, 3.7) m. In Fig. 8(e), the localization result is (0.74, −0.08, 3.7) m. The MTDOA method is comparable with DAMAS beamforming, presenting an accurate localization point.

These experiments have validated that the proposed method can produce moving sound source localization results as good as DAMAS beamforming. The MTDOA method has offered a new way for locating moving sound sources, while DAMAS has the additional benefit of providing the strength of the sources. Considering time consumption, it took 14 s to produce result in Fig. 7(c) at 1000 Hz, and 72 s to produce result in Fig. 8(c) at 5 frequency points (900 Hz, 950 Hz, 1000 Hz, 1050 Hz and 1100 Hz) in 1000 Hz third octave band. On the other hand, for both Fig. 7(e) and Fig. 8e), the MTDOA method took about 90 s. Since the MTDOA method processes in time domain, the analyzing frequency band has no effect on time consumption. Thus MTDOA method is comparable with DAMAS considering a wide band sound source. For a single-frequency sound source, however, MTDOA is more time consuming. Future studies would focus on reducing calculation amount, for example, using a quick iterative search method instead of the traversal search method.

5. Conclusions

This study proposed a triangulation-based moving sound source localization method. First, the method requires scanning of a possible sound source locating plane. Next, for each hypothetical source location on the possible plane, the Doppler effect has to be removed, and MTDOA has to be calculated. In this method, the sound source location is estimated and compared to the original hypothetical location. Next, the deviation from the source location estimation to the original hypothetical location is calculated. Finally, the sound source can be located by searching for the minimum in the deviation matrix.

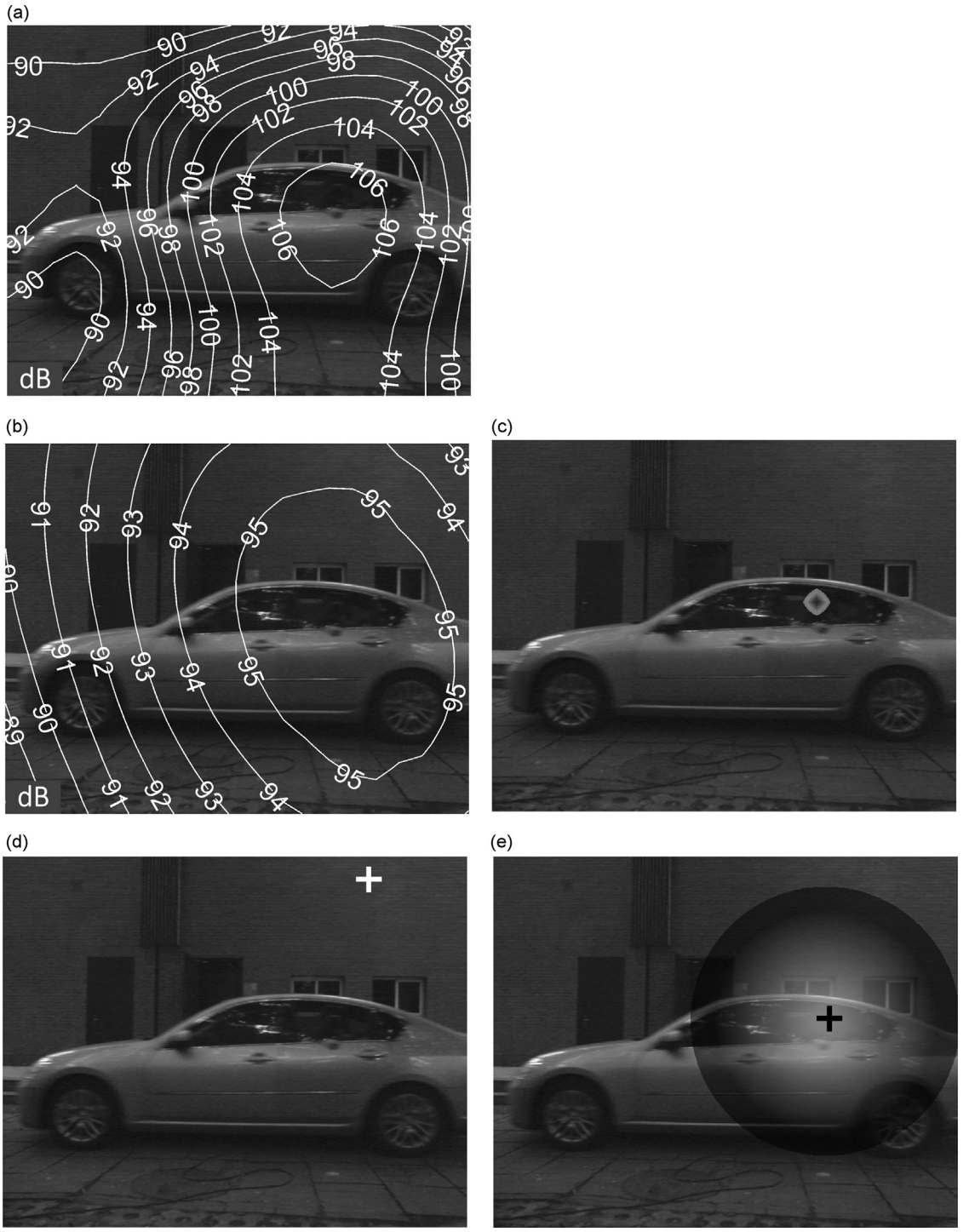

Fig. 7. Single-frequency loudspeaker localization results: (a) the beamforming method using 32 channels; (b) the beamforming method using 4 channels; (c) the DAMAS beamforming using 4 channels; (d) the traditional triangulation method; (e) the MTDOA method.

Our simulations demonstrated that the proposed method can be used to locate both single-frequency and wide band moving sound sources. Additionally, our experiments have validated the proposed method. Experimental results show that the MTDOA method can locate moving sound sources with equivalent accuracy and spatial resolution as DAMAS beamforming, while DAMAS has the additional benefit of providing the strength of the sources. The MTDOA has offered a new way for locating moving sound sources. Future work would focus on reducing the amount of calculation needed to perform quasi-real-time localization.

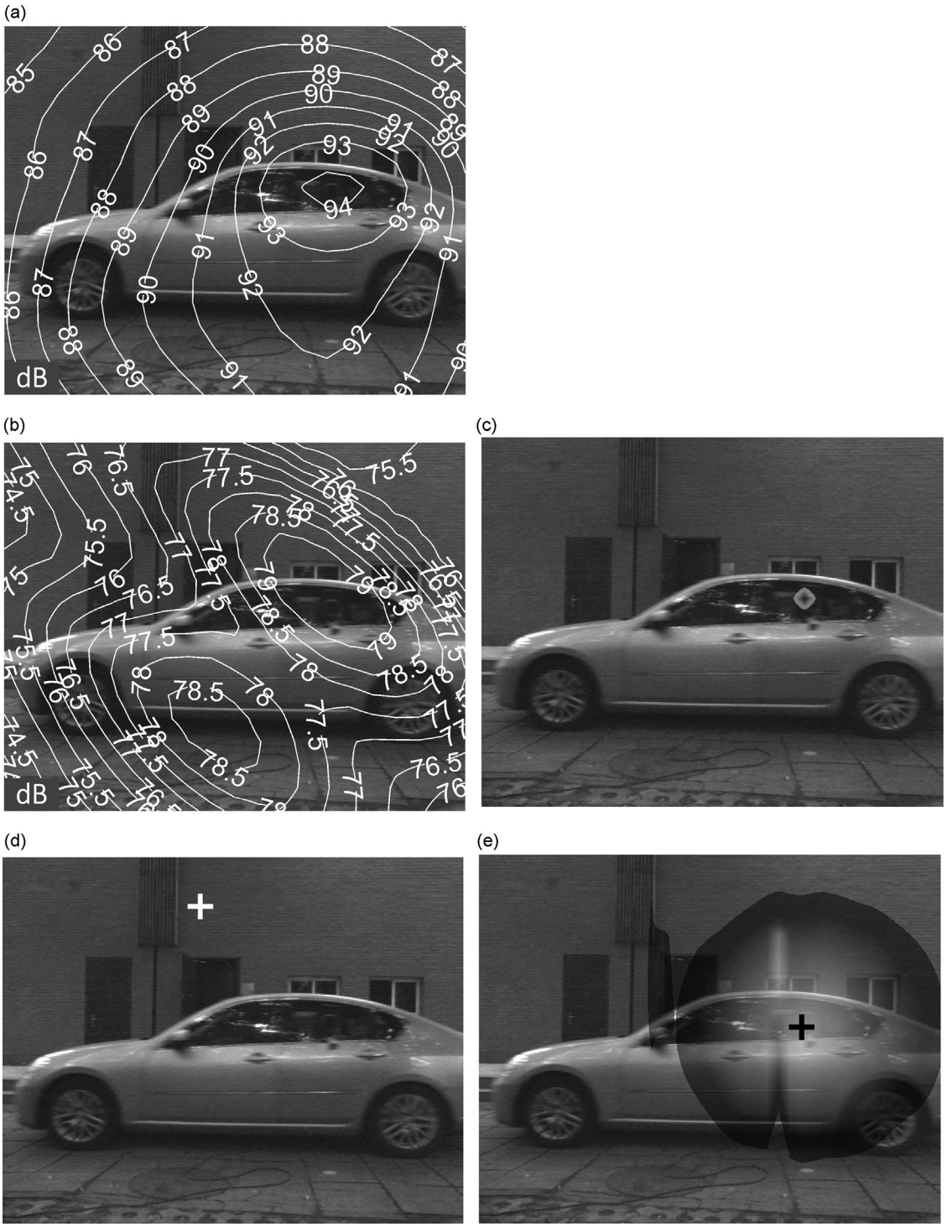

Fig. 8. White noise loudspeaker localization results: (a) the beamforming method using 32 channels; (b) the beamforming method using 4 channels; (c) the DAMAS beamforming using 4 channels; (d) the traditional triangulation method; (e) the MTDOA method.

Acknowledgments

This study has been supported by the National Natural Science Foundation of China (51375252) and the Program for New Century Excellent Talents in University (NCET-11-0284).

References

[1] D.H. Johnson, D.E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Simon & Schuster, New York, 1992.
[2] B. Barsikow, W.F. King, On removing the Doppler frequency shift from array measurements of railway noise, *Journal of Sound and Vibration* 120 (1988) 190–196.
[3] R.W. Stoker, Y. Guo, C. Streett, N. Burnside, Airframe Noise Source Locations of a 777 Aircraft in Flight and Comparisons with Past Model Scale Tests, AIAA paper 2003-3232, 9th AIAA/CEAS Aeroacoustics Conference, Hilton Head, South Carolina, USA, 2003.
[4] S. Guerin, U. Michel, H. Siller, U. Finke, G. Saueressig, Airbus A319 Database from Dedicated Flyover Measurements to Investigate Noise Abatement Procedures, AIAA paper 2005-2981, 11th AIAA/CEAS Aeroacoustics Conference, Monterey, California, USA, 2005.
[5] S. Guerin, U. Michel, Aero-engine Noise Investigated from Flight Tests, *AIAA paper 2006-2463, 12th AIAA/CEAS Aeroacoustics Conference*, Cambridge, Massachusetts, USA, 2006.
[6] G. Zechel, A. Zeibig, M. Beitelschmidt, Time-domain beamforming on moving objects with known trajectories, *Proceedings of the BeBeC-2010-12, Berlin Beamforming Conference*, Berlin, Germany, 2010.
[7] E. Sarradj, T. Geyer, H. Brick, K.R. Kirchner, T. Kohrs, Application of beamforming and deconvolution techniques to aeroacoustic sources at highspeed trains, INTER-NOISE and NOISE-CON Congress and Conference Proceedings, New York, USA, 2012.
[8] D. Yang, B. Li, Z. Wang, S. Zheng, X. Lian, Video visualization for moving sound sources based on binocular vision and short-time beamforming, *Noise Control Engineering Journal* 58 (2010) 382–388.
[9] T.F. Brooks, W.M. Humphreys, A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays, *Journal of Sound and Vibration* 294 (2006) 856–879.
[10] P. Sijtsma, R.W. Stoker, Determination of Absolute Contributions of Aircraft Noise Components using Fly-Over Array Measurements, AIAA paper 2004-2958, 10th AIAA/CEAS Aeroacoustics Conference, Manchester, UK, 2004.
[11] S. Oerlemans, P. Sijtsma, B.M. López, Location and quantification of noise sources on a wind turbine, *Journal of Sound and Vibration* 299 (2007) 869–883.
[12] I.C. Dedoussi, T.P. Hynes, H. Siller, Investigating landing gear noise using fly-over data: the case of a Boeing 747-400, AIAA paper 2013-2115, 19th AIAA/CEAS Aeroacoustics Conference, Berlin, Germany, 2013.
[13] J.D. Maynard, E.G. Williams, Y. Lee, Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH, *Journal of the Acoustical Society of America* 78 (1985) 1395–1413.
[14] W.A. Veronesi, J.D. Maynard, Nearfield acoustic holography (NAH) II. Holographic reconstruction algorithms and computer implementation, *Journal of the Acoustical Society of America* 81 (1987) 1307–1322.
[15] H. Nakagawa, H. Tsuru, T. Tanaka, I. Sakamoto, Detection and visualization of moving sound source through acoustic holography, *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Christchurch, New Zealand, 1998.
[16] H. Kwon, Y. Kim, Moving frame technique for planar acoustic holography, *Journal of the Acoustical Society of America* 103 (1998) 1734–1741.
[17] S. Park, Y. Kim, An improved moving frame acoustic holography for coherent band limited noise, *Journal of the Acoustical Society of America* 104 (1998) 3179–3189.
[18] S. Park, Y. Kim, Visualization of pass-by noise by means of moving frame acoustic holography, *Journal of the Acoustical Society of America* 110 (2001) 2326–2339.
[19] D. Yang, Z. Wang, B. Li, Y. Luo, X. Lian, Quantitative measurement of pass-by noise radiated by vehicles running at high speeds, *Journal of Sound and Vibration* 330 (2011) 1352–1364.
[20] B.G. Ferguson, L.G. Criswick, K.W. Lo, Locating far-field impulsive sound sources in air by triangulation, *Journal of the Acoustical Society of America* 111 (2002) 104–116.
[21] P. Giraudet, H. Glotin, Echo-robust and real-time 3d tracking of marine mammals using their transient calls recorded by hudrophones array, IEEE Transactions on Acoustics, Speech and Signal Processing, Toulouse, France, 2006.
[22] S.F. Wu, N. Zhu, Locating arbitrarily time-dependent sound sources in three dimensional space in real time, *Journal of the Acoustical Society of America* 128 (2010) 728–739.
[23] N. Zhu, S.F. Wu, Sound source localization in three-dimensional space in real time with redundancy checks, *Journal of Computational Acoustics* 20 (2012) 1250007.
[24] S.F. Wu, N. Zhu, Passive sonic detection and ranging for locating sound sources, *Journal of the Acoustical Society of America* 133 (2013) 4054–4064.
[25] S.C. Lee, W.R. Lee, K.H. You, TDOA/FDOA based aircraft localization using adaptive fading extended Kalman filter algorithm, Power control and optimization, *Proceedings of the 3rd Global Conference on Power Control and Optimization*, Gold Coast, Australia, 2010.
[26] P.M. Morse, K.U. Ingard, *Theoretical Acoustics*, Princeton University Press, Princeton, 1968.
[27] C.H. Knapp, G.C. Carter, The generalized correlation method for estimation of time delay, *IEEE Transactions on Acoustics, Speech, and Signal Processing* (1976) 320–327.
[28] M. Omologo, P. Svaizer, Acoustic event localization using a crosspower-spectrum phase based technique, IEEE transactions on Acoustics,Speech, and Signal Processing, Adelaide, Australia, 1994, pp. 273–276.
[29] M. Feng, Y. Diange, W. Rujia, W. Junjie, W. Ziteng, L. Xiaomin, A triangulation method based on phase difference of arrival estimation for sound source localization, *Proceedings of the 21st International Congress on Sound and Vibration*, Beijing, China, 2014.