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

INDOOR SOUND SOURCE LOCALIZATION USING A ROOM FEATURE-BASED METHOD

Wang Haitao, Zeng Xiangyang, Wang Lu and Liu Yanshan

Northwestern Polytechnical University, School of Marine Science and Technology, Xi'an, China
email: wht@nwpu.edu.cn

Indoor sound source localization is one of key foundations for public security, source separation, surveillance, etc. Instead of neglecting the shape and boundary information of the room in the traditional methods, a room feature-based localization method is proposed in this paper. This method uses the geometry information as the prior knowledge to construct a numerical model which simulates the characteristics of the indoor sound field. By dividing the room into a number of subdomains using nodes, the Helmholtz equation can be inversely solved using the numerical model based on the measured signals. This inverse result finally gives the location of the sound source which is represented by the location of the subdomain. The method is validated by numerical simulations. It is shown that the method can achieve high localization accuracy even in a non-convex room with the assistance of the room geometry. The verification also demonstrates that the method is valid for the multiple source localization.

Keywords: sound source localization, reverberant enclosure, inverse solution, Helmholtz equation

1. Introduction

Indoor sound source localization is a key technique in many research areas such as aircraft cabin noise source identification, intelligent robot localization and security surveillance. The methods based on the theories of beam-forming, high-resolution spectral estimation and the time delay of arrival are most frequently used in current studies[1-3]. These methods are mainly functioned based on the microphone array and can obtain good accuracy and spatial resolution in large open environment[4]. However, the features of indoor environments are quite different from the open ones. The factors of complex structure, high reverberation, acoustic interference and diffraction effect result that the sound propagations are also totally different from the free field. In such cases, many problems such as low robustness in the reverberant field, confusion on the time of arrival and low accuracy for source in the near field often emerge in the aforementioned method[5], which results that the performance and applicability of these methods degrade substantially.

Essentially, the algorithms in traditional localization methods completely rely on the absolute signal processing, where all used information is the measured signals while the information of room itself is totally neglected. However, the room features such as shape and boundary conditions actually play very important roles in forming the distribution of the sound field. They also determine the propagating path from the source to the receiver. It can be expected that the localization will be much more accurate for a specific room if the room feature information can be involved in the algorithm.

There has been method using the room geometry to achieve localization in the enclosed space. [6]The space is divided into a series of triangular grids. The finite element method is used to construct a numerical model which can denote the modal characteristics of the sound field based on those elements. This method can achieve good accuracy for non-convex space. However, this method still faces some potential problems. The accuracy of the method strongly depends on the numerical model which is determined by the distribution of the triangular grids. For the space with complex shape or boundary, the triangular grids usually cannot discrete the space in a uniform dimensions. Thus, the accuracy of the numerical model degrades obviously in such situation. Consequently, the localization method often cannot achieve robust accuracy.

In this paper, a room feature-based method for indoor sound source localization is proposed. In this method, the room is divided by independent nodes and their neighbouring domains instead of grids. The numerical model for describing the sound field is constructed using a global interpolating function derived by moving least square method. Finally, the sound source is localized by inversely solving the Helmholtz equation using the numerical model. The paper is structured as follows. The derivations of the numerical model and the global interpolating function are introduced in section 2. The method is validated by numerical simulations in section 3. At last, the conclusion is given in section 4.

2. Room feature-based localization method

2.1 Numerical model determining the indoor sound field

Assuming there is a sound source in the indoor environment as illustrated in Fig. 1. The mass of the medium provided by the source is $\rho_0 q$ in unit time. The acoustic wave equation with a source is given as follows,

$$\nabla^2 p - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} = -\rho_0 \frac{\partial q}{\partial t} \tag{1}$$

where t stands for time; p is the sound pressure; c_0 is the speed of sound in air; ρ_0 is the equilibrium density of air; and q is the volume velocity of the sound source.

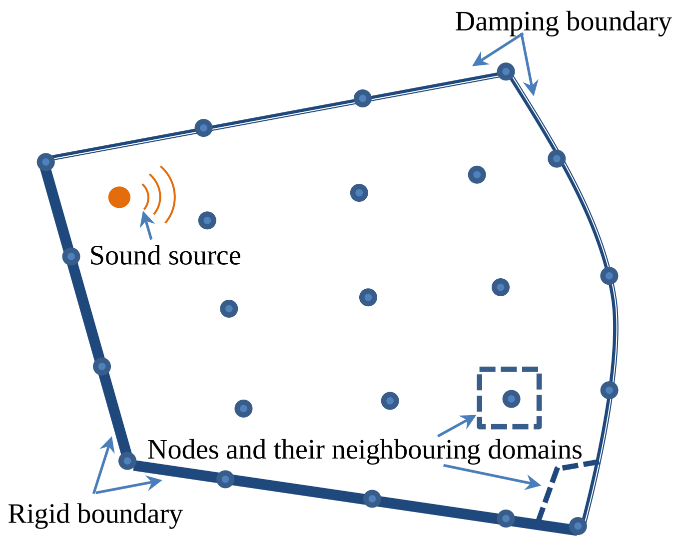


Figure 1: Schematic of the indoor environment and its discrete form based on nodes

Considering the system in frequency domain, p and q in the above equation can be expressed by

$$\begin{cases} p = p_\omega e^{\mathrm{j}\omega t} \\ q = q_\omega e^{\mathrm{j}\omega t} \end{cases} \tag{2}$$

where ω is the circular frequency, j is the imaginary unit, p_ω and q_ω are the sound pressure and the intensity of sound source in the frequency domain, respectively.

Then, the Helmholtz equation can be simultaneously deduced by Eqs. (1) and (2) as follow

$$\nabla^2 p_\omega + k^2 p_\omega = -\mathrm{j}\rho_0 \omega q_\omega \tag{3}$$

where k denotes the wave number and is defined by ω / c_0.

Helmholtz equation is the governing equation which controls the distribution of the sound field. By knowing the location and intensity of the sound source, the sound pressure of a receiving point can be obtained by solving the Helmholtz equation. Inversely, if the sound pressure has been known, the sound source also can be solved based on the Helmholtz equation. This thought builds the basic localization theory in this paper.

Indoor environment usually has complex boundary. The absorption of the boundary can be expressed as follow,

$$\begin{cases} \dfrac{\partial p_\omega}{\partial n} = -\mathrm{j}\rho_0 \omega \dfrac{p_\omega}{Z_s} & \Gamma_1 \\ \dfrac{\partial p_\omega}{\partial n} = 0 & \Gamma_2 \end{cases} \tag{4}$$

where n denotes outward normal direction, Z_s is the specific acoustic impedance of the boundary. Γ_1 stands for the damping boundary and Γ_2 stands for the rigid boundary.

In order to solve the Helmholtz equation, a trial function $\overline{p_\omega}$ should be assumed and substituted into Eqs. (3) and (4) according to the weighted residual method[7]. For each equation, there is a residual as follows,

$$\begin{cases} R_\Omega = \nabla^2 \overline{p_\omega} + k^2 \overline{p_\omega} + \mathrm{j}\rho_0 \omega q_\omega \\ R_{\Gamma_1} = \dfrac{\partial \overline{p_\omega}}{\partial n} + \mathrm{j}\rho_0 \omega \dfrac{\overline{p_\omega}}{Z_s} \\ R_{\Gamma_2} = \dfrac{\partial \overline{p_\omega}}{\partial n} \end{cases} \tag{5}$$

The weight functions, W_Ω and W_Γ are introduced here in order to make the residuals being minimum. Multiplying the residuals by the weight functions in the following form

$$W_\Omega \int_\Omega R_\Omega \mathrm{d}v + W_\Gamma \left(\int_{\Gamma_1} R_{\Gamma_1} \mathrm{d}s + \int_{\Gamma_2} R_{\Gamma_2} \mathrm{d}s \right) = 0 \tag{6}$$

In Galerkin method, the weight function is the same as the trial function[8]. Then, in this problem, W_Ω and W_Γ can be replaced by $\overline{p_\omega}$ and $-\overline{p_\omega}$, respectively. The following equation can be obtained with substituting the residuals and weight functions into Eq. (6)

$$\int_\Omega \overline{p_\omega} \cdot (\nabla^2 \overline{p_\omega} + k^2 \overline{p_\omega} + \mathrm{j}\rho_0 \omega q_\omega) \mathrm{d}v - \left(\int_{\Gamma_1} \overline{p_\omega} \cdot (\frac{\partial \overline{p_\omega}}{\partial n} + \mathrm{j}\rho\omega \frac{\overline{p_\omega}}{Z_s}) \mathrm{d}s + \int_{\Gamma_2} \overline{p_\omega} \cdot \frac{\partial \overline{p_\omega}}{\partial n} \mathrm{d}s \right) = 0 \tag{7}$$

According to Green's first identity,

$$\int_\Omega (\varphi \nabla^2 \Phi + \nabla \varphi \nabla \Phi) \mathrm{d}v = \int_\Gamma \varphi \frac{\partial \Phi}{\partial n} \mathrm{d}s \tag{8}$$

The reduced equation of Eq. (7) can be written as

$$\int_\Omega (\nabla \overline{p_\omega} \cdot \nabla \overline{p_\omega} - k^2 \overline{p_\omega} \cdot \overline{p_\omega} - j\rho_0 \omega \overline{p_\omega} q_\omega) \mathrm{d}v + \int_{\Gamma_1} j\rho_o \omega \frac{\overline{p_\omega}}{Z_s} \cdot \overline{p_\omega} \mathrm{d}s = 0 \tag{9}$$

The Helmholtz equation has been transformed to an integral equation spreads on the global space domain and its boundary. Usually, this equation cannot be directly solved because of the complex space shape and boundary condition. As a natural thought, the complex integral calculations over the global domain are replaced by simple calculations over subdomains.

As illustrated in Fig. 1, the room is divided by *n* nodes. Based on the theory of the numerical calculation, the sound pressure at any position, for example *r*, in the problem domain can be expressed by

$$p_r = \boldsymbol{N}_r{}^{\mathrm{T}} \boldsymbol{p} = [N_{r1}, N_{r2}, \cdots, N_{rn}] \begin{Bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{Bmatrix} \tag{10}$$

where *n* is the number of nodes, $\boldsymbol{N}_r$ is the interpolating function for position *r*, it denotes the relationship of r and all nodes, $\boldsymbol{p}$ is the vector of nodal sound pressure. Substituting the $\overline{p_\omega}$ in Eq. (9) with the form described in Eq. (10), the following equation can be obtained

$$\int_\Omega [\boldsymbol{p}_\omega{}^{\mathrm{T}} (\nabla \boldsymbol{N})(\nabla \boldsymbol{N})^{\mathrm{T}} \boldsymbol{p}_\omega - k^2 \boldsymbol{p}_\omega{}^{\mathrm{T}} \boldsymbol{N}\boldsymbol{N}^{\mathrm{T}} \boldsymbol{p}_\omega - \mathrm{j}\rho_0 \omega \boldsymbol{p}_\omega{}^{\mathrm{T}} \boldsymbol{N} q_\omega]\mathrm{d}v + \int_{\Gamma_1} \frac{\mathrm{j}\rho_0\omega}{Z_s} (\boldsymbol{p}_\omega{}^{\mathrm{T}} \boldsymbol{N}\boldsymbol{N}^{\mathrm{T}} \boldsymbol{p}_\omega)\mathrm{d}s = 0 \tag{11}$$

This equation can be converted into the form as follows.

$$\int_\Omega [(\nabla \boldsymbol{N})(\nabla \boldsymbol{N})^{\mathrm{T}}]\mathrm{d}v \cdot \boldsymbol{p}_\omega - \int_\Omega (k^2 \boldsymbol{N}\boldsymbol{N}^{\mathrm{T}})\mathrm{d}v \cdot \boldsymbol{p}_\omega - \int_\Omega (\mathrm{j}\rho_0 \omega \boldsymbol{N} q_\omega)\mathrm{d}v + \int_{\Gamma_1} \frac{\mathrm{j}\rho_0\omega}{Z_s} (\boldsymbol{N}\boldsymbol{N}^{\mathrm{T}})\mathrm{d}s \cdot \boldsymbol{p}_\omega = 0 \tag{12}$$

In the above equation, using $\boldsymbol{K}$ to represent $\int_\Omega (\nabla \boldsymbol{N})(\nabla \boldsymbol{N})^{\mathrm{T}}\mathrm{d}v$, $\boldsymbol{M}$ to $\frac{1}{\boldsymbol{c}^2}\int_\Omega \boldsymbol{N}\boldsymbol{N}^{\mathrm{T}}\mathrm{d}v$, $\boldsymbol{C}$ to $\frac{\rho_0}{Z_s}\int_{\Gamma_1} \boldsymbol{N}\boldsymbol{N}^{\mathrm{T}}\mathrm{d}s$, and $\boldsymbol{F}$ to $\int_\Omega \mathrm{j}\rho_0\omega \boldsymbol{N} q_\omega \mathrm{d}v$. Finally, a system equation can be obtained as follows.

$$(\boldsymbol{K} - \omega^2 \boldsymbol{M} + \mathrm{j}\omega \boldsymbol{C})\boldsymbol{p}_\omega = \boldsymbol{F} \tag{13}$$

Imitating the elastic system, $\boldsymbol{K}$ is called as the stiffness matrix, $\boldsymbol{M}$ as the mass matrix, $\boldsymbol{C}$ as the damping matrix and $\boldsymbol{F}$ as the load matrix. If the sound source is a point source at r_0, the intensity of the source can be described as $q_\omega = q_\omega(r_0)\delta(r - r_0)$. By defining $Q = q_\omega(r_0)\mathrm{d}v$ as the volume velocity, the load matrix can be consequently written as $\boldsymbol{F} = \int_\Omega \mathrm{j}\rho_0\omega \boldsymbol{N} q_\omega(r_0)\delta(r - r_0)\mathrm{d}v = \mathrm{j}\rho_0\omega \boldsymbol{N}_{r_0} Q$.

In this paper, the localization is realized based on Eq. (13). When solving the Helmholtz equation using Eq. (13), the intensity of the sound source is decomposed and scattered on all nodes through the interpolating function. As an inverse result, if the sound signal is measured at a receiving point with known location, the nodal pressure also can be obtained based on the interpolating function of the receiving point. Then, Eq. (13) can be inversely solved as follows

$$(\boldsymbol{K} - \omega^2 \boldsymbol{M} + \mathrm{j}\omega \boldsymbol{C})\boldsymbol{N}_r p_r = \boldsymbol{F} \tag{14}$$

where p_r is the signal in frequency domain measured at receiving point *r*, $\boldsymbol{N}_r$ is the interpolating function of location *r* to all nodes.

If there are a certain number of microphones of which the locations are known, the initial distribution of the intensity of the source can be calculated simultaneously according to Eq. (14) and the location of the source can be obtained consequently.

2.2 Constructing the interpolating function

From the localization setup and derivation of the numerical model, it can be seen that the interpolating function plays very important role in the method. On one hand, the system matrices are mainly obtained by the interpolating functions and their derivatives. The accuracies of the stiffness matrix , mass matrix and damping matrix are closely related with the interpolating functions. On the other hand, the interpolating function gives the final localization result. Thus, it is very important to construct a type of interpolating function with high performance. In this paper, the Moving Least Square (MLS) method is used to construct the interpolating function.

In the MLS method, an approximate function $u^h(\boldsymbol{x},\bar{\boldsymbol{x}})$ in the neighborhood of the field function $u(\boldsymbol{x})$ is defined as[9]

$$u^h(\boldsymbol{x},\bar{\boldsymbol{x}}) = \sum_{i=1}^{m} p_i(\bar{\boldsymbol{x}})a_i(\boldsymbol{x}) = \boldsymbol{p}^{\mathrm{T}}(\bar{\boldsymbol{x}})\boldsymbol{a}(\boldsymbol{x}) \tag{15}$$

where $\bar{\boldsymbol{x}} = (x, y, z)$ denotes the coordinates of the node which is located in the neighborhood of the calculating point $\boldsymbol{x}$; $\boldsymbol{p}^{\mathrm{T}}(\bar{\boldsymbol{x}}) = [p_1(\bar{x}), p_2(\bar{x}), \cdots, p_m(\bar{x})]$ is the vector of the bases; $\boldsymbol{a}(x) = [a_1(x), a_2(x), \cdots a_m(x)]^{\mathrm{T}}$ is the vector of the unknown coefficients and m is the number of the bases.

There are several kinds of functions which can be used as the bases, such as the monomial function, the trigonometric function, and the wavelet function. The most widely used basis is the monomial function which is defined as follows in three dimensional problems

$$\boldsymbol{p}^{\mathrm{T}}(\bar{\boldsymbol{x}}) = [1, x, y, z] \quad \text{(m=4, linear basis)} \tag{16}$$

A discrete weighted norm needs to be constructed as follow in order to obtain the vector of the unknown coefficients $\boldsymbol{a}(\boldsymbol{x})$

$$J = \sum_{I=1}^{N} w_I(\boldsymbol{x}-\boldsymbol{x}_I)\left[\boldsymbol{p}^{\mathrm{T}}(\boldsymbol{x}_I)\boldsymbol{a}(\boldsymbol{x}) - u_I\right]^2 = \sum_{I=1}^{N} w_I(\boldsymbol{x}-\boldsymbol{x}_I)\left[\sum_{i=1}^{m} p_i(\boldsymbol{x}_I)a_i(\boldsymbol{x}) - u_I\right]^2 \tag{17}$$

where N is the number of nodes; u_I is the nodal value on $\boldsymbol{x}_I$, and $w_I(\boldsymbol{x}-\boldsymbol{x}_I)$ is the weight function on $\boldsymbol{x}_I$.

The weight function is only available in the support domain which is usually a sphere or a box. A common weight function is listed as follows

$$w(r) = \begin{cases} \dfrac{e^{-r^2\beta^2} - e^{-\beta^2}}{1 - e^{-\beta^2}} & r \le 1 \\ 0 & r > 1 \end{cases} \tag{18}$$

where r stands for the radius providing the support domain is a sphere.

With minimizing J

$$\frac{\partial J}{\partial a_j(x)} = 2\sum_{I=1}^{N} w_I(x)[\sum_{i=1}^{m} p_i(x_I)a_i(x) - u_I]p_j(x_I) = 0 \quad j = 1, 2, \cdots, m \tag{19}$$

the following equation can be obtained after collocation

$$\boldsymbol{A}(\boldsymbol{x})\boldsymbol{a}(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} \tag{20}$$

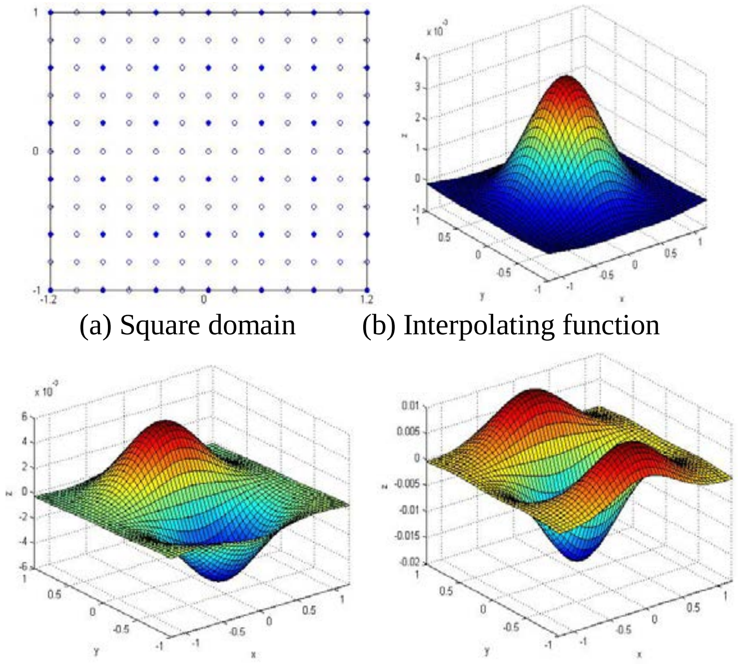
where $\boldsymbol{A}(\boldsymbol{x})$, $\boldsymbol{B}(\boldsymbol{x})$ are matrices defined as

$$\begin{cases} \boldsymbol{A}(\boldsymbol{x}) = \sum_{I=1}^{N} w_I(\boldsymbol{x})p(x_I)p^{\mathrm{T}}(x_I) \\ \boldsymbol{B}(\boldsymbol{x}) = [w_1(\boldsymbol{x})p(x_1), w_2(\boldsymbol{x})p(x_2), \cdots, w_N(\boldsymbol{x})p(x_N)] \end{cases} \tag{21}$$

After solving Eq. (20), the following equation can be obtained by substituting $\boldsymbol{a}(\boldsymbol{x})$ into Eq. (14) and $\boldsymbol{N}(\boldsymbol{x})$ is obtained consequently.

$$\boldsymbol{u}^h(\boldsymbol{x},\bar{\boldsymbol{x}}) = \boldsymbol{p}^{\mathrm{T}}(\bar{x})\boldsymbol{A}^{-1}(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x})\boldsymbol{u} = \boldsymbol{N}(\boldsymbol{x},\bar{x})\boldsymbol{u} \tag{22}$$

The interpolating function obtained by MLS method has smooth interpolating feature over the whole problem domain. A schematic of interpolating function, first order and second order derivatives over a square domain calculated using MLS are illustrated in Fig. 2



(a) Square domain (b) Interpolating function

(c) First order derivative (d) Second order derivative

Figure 2: Schematic of the interpolating function and its derivatives for a 2D square area

The dots in Fig. 2 are the discrete nodes spread over the square domain and the circles are the integrating points. It can be seen that the interpolating function of the original point over all nodes is smooth, as well as its first and second order of derivatives. This feature can make the interpolation has high accuracy, which is beneficial for the localization.

3. Numerical verifications

3.1 Localization of multi-sources

In order to evaluate the accuracy of the localization method proposed in sec. 2, a multi-source localization is numerically simulated. The room is L-shaped and its dimension is illustrated in Fig. 3. In this paper, the measured signals are all obtained using Virtual. Lab Acoustics which is a commercial simulation software based on the finite element and boundary element method. Eq. 14 demonstrates that the system equation can be inversely solved even there is just the measure signals at only one frequency when there are several microphones. In this verification, we used 9 microphones and the signals at 200Hz. The localization result is illustrated in Fig. 3.

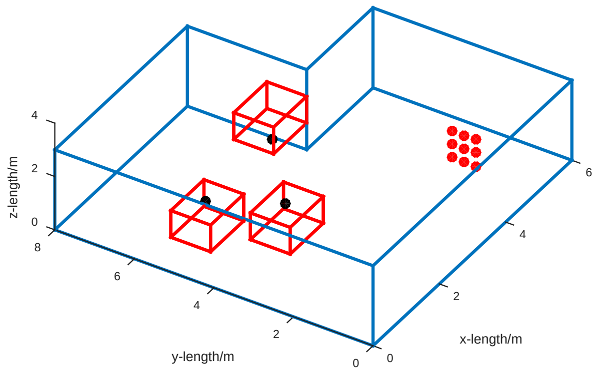
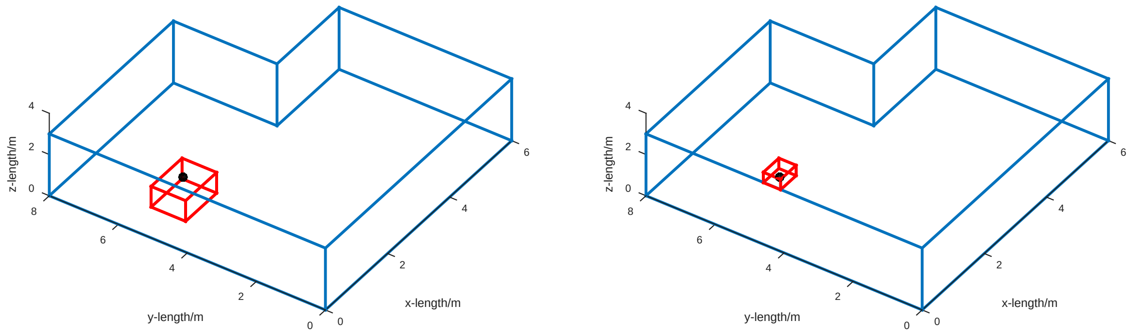


Figure 3: Localization results of three sound sources

The black dots in Fig. 3 represent the real locations of sources. The red boxes are the locations obtained using the localization method. Note that the localization result obtained using the proposed method is an approximate domain but not an exact coordinate in the space. This is because that the localization is realized by inversely solving Eq. (14), where the result is actually picked from the distribution of the sound intensity over the nodes. Thus, the location of the source is represented by the coordinate of a node and its neighboring domain.

Figure 3 illustrates that the localization method is capable of giving the correct locations of the multi-sources even in a non-convex room. In multi-experiments for sound sources at different locations, the accuracy rate is 93%. The experiments also demonstrates that the localization method is capable of giving the correct locations of even 5 sources. More evaluations on the performance of the localization method will be conducted in further studies.

In the proposed localization method, the resolution of the localization can be changed by adjusting the density of the discrete nodes. A localization problem of a single source in the same space with the above example is further solving under two conditions of different nodal densities. The comparison of the localization results are illustrated in Fig. 4.



(a) Localization under a large nodal density (b) Localization under a small nodal density

Figure 4: Comparison of the resolution of localizations based on different nodal densities

Figure 4 illustrates that the resolutions of the localizations under different nodal densities are obviously different. Large nodal densities can give high resolution. However, it is should be note that large densities causes the system matrices are also large scaled at same time. The computational efficiency will decrease. Thus, a balance between the localization time and resolution is important.

4. Conclusion

An indoor sound source localization method is proposed in this paper. The method uses the room features as the prior knowledge to assist the localization. The numerical model for representing the indoor sound field is derived according to the theory on discretely solving the Helmholtz equation. The interpolating function is constructed using the Moving Least Square method to connect the sound pressure on discrete nodes. By inversely solving the Helmholtz equation, the location of the sound source can be obtained, which is represented by the location of the nodes. The method is validated by numerical simulations for a multi-sources localization problem. It is shown that the method can achieve high localization accuracy even in a non-convex room with the assistance of the room geometry.

REFERENCES

1 Huang, F., Sheng, W. and Ma, X. Modified projection approach for robust adaptive array beamforming, *Signal Processing*, **92**(7), 1758-1763, (2012).

2 MN, K. E., Boyer, R., Renaux, A. and Marcos, S. Conditional and unconditional Cramer-Rao Bounds for near-field source localization, *IEEE Transactions on Signal Processing*, **58**(5), 2901-2907, (2010).

3 Yuan, Q., Chen, Q. and Sawaya, K. Accurate DOA estimation using array antenna with arbitrary geometry, *IEEE Transactions on Antennas and Propagation*, **53**(4), 1352-1357, (2005).

4 Krim, H. and Viberg, M. Two decades of array signal processing research: the parametric approach, *IEEE Signal Process Mag.*, **13**(4), 67-94, (1996).

5 Roman, N. and Wang, D. L. Binaural tracking of multiple moving sources., *IEEE Transactions on Audio, Speech, and Language Processing*, **16**(4), 728-739, (2008).

6 Dokmanic, I. and Vetterli, M., *Room helps: Acoustic localization with finite elements*, *ICASSP2012*, Kyoto, Japan, 25-30 March, (2012).

7 Finlayson, B. A. and Scriven, L. E. The method of weighted residuals - a review, *Appl. Mech. Rev.*, **19**, 735-748, (1966).

8 Nayroles, B., Touzot, G. and Villon, P. Generalizing the finite element method: diffuse approximation and diffuse elements, *Comput. Mech.*, **10**(5), 307-318, (1992).

9 Lancaster, P. and Salkauskas, K. Surfaces generated by moving least squares methods, *Math. Comput.*, **37**(155), 141-158, (1981).