Support vector machine-based sound pressure extrapolation to overcome the finite aperture used in nearfield acoustic holography

Hefeng Zhou¹,²,³, Yun Zhao¹, Dongbao Gao¹, Zhangfu Tian¹, Xinwu Zeng¹ and Weikang Jiang²

¹College of Meteorology and Oceanology, National University of Defense Technology, Changsha, Hunan 410073, China
²School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
³E-mail: zhouhefeng@nudt.edu.cn

Abstract. Sound pressure extrapolation based on support vector machine (SVM) is proposed in order to reduce the reconstruction error in Fourier-based nearfield acoustic holography (NAH) when the aperture is small compared to the source size. The measured pressure on the aperture is introduced into SVM and then a regression model is obtained to extrapolate the sound pressure outside the aperture. Numerical simulations show that the new method allows a relatively accurate and quick reconstruction in comparison with the results using zero-padding and linear predictive border-padding.

1. Introduction
Nearfield acoustic holography is a technique to obtain a visible sound image through a reconstruction based on recording the radiated field of an object [1, 2]. A key issue to NAH is to obtain the complex sound pressure on the measurement surface in the nearfield, namely the hologram plane. According to the strict NAH theoretical model, the sound pressure should be measured continuously on an infinite hologram plane. In practice, the holographic sound pressure measurements can only be carried out at a series of discrete positions on a finite aperture [3]. This is a mathematical approximation to the theory, leading to a reconstruction error. Therefore, in order to reduce the error caused by the finite aperture, the measurement aperture in conventional NAH is generally required to be at least four times the size of the sound source [4]. However, this requirement may be achievable for some small sound sources but difficult to meet for large sound sources such as cars or planes. Moreover, if only partial radiation characteristics of the large sound source attract attention, it will serve little real purpose to measure in accordance with the requirement.

In recent years, there have been some developments to overcome the above limitation. Saijyou and Yoshikawa found that the data extrapolation can effectively reduce the reconstruction error due to the finite measurement aperture when dealing with the large-scale implementation of NAH, and proposed a wavenumber extrapolation method and a real space extrapolation method for planar sources [5]. Williams studied the continuation of sound nearfields and developed an extrapolation method by combining the singular value decomposition (SVD) with the regularization technique [6]. This method can be applied to arbitrarily shaped sources. Williams introduced an extrapolation technique into the
reconstruction-patch NAH [7]. Patch NAH gets rid of the measurement aperture size requirement of conventional NAH and only the measured data on the corresponding hologram part of interest is needed. A fast Fourier transform and SVD based patch NAH approach for planar sources was proposed [8] and generalized to cylindrical geometry subsequently [9]. Lee and Bolton described a patch NAH procedure that allows the sound source reconstruction from the hologram pressure measured over multiple discontinuous regions [10]. In general, this kind of iterative patch NAH has a low computational efficiency and a convergence condition difficult to determine (sometimes a priori knowledge is necessary).

In addition to the above methods, another sound field iterative extrapolation technique called linear predictive border padding (LPBP) was developed by Scholte et al. [11]. Linear predictive filtering is used to determine samples outside a finite area through an approximation based on a certain number of previous samples. Specifically, a single row or column of the two-dimensional data matrix is first applied to calculating the impulse response coefficients. Next, a digital filter is generated with the corresponding impulse response coefficients. Finally, a zero-valued vector of the border-padded width is fed into the filter to extrapolate the data to either direction. This process is repeated for all rows and columns and the padded samples can be included in new cycles. In contrast with other iterative data extrapolation methods, there is no convergence condition involved in the calculation of LPBP such that a priori knowledge is not required. However, due to the multiple extrapolation steps needed for the whole border-padded matrix, LPBP takes much processing time, especially when padding a small matrix to a relatively large size.

In this paper, the relative simplicity, accuracy and rapidity of Fourier-based NAH is acknowledged. A new method, fundamentally different from the aforementioned ones, is developed to extend the finite aperture quickly before SFT without a priori knowledge. This is achieved by introducing a support vector machine (SVM) into the sound pressure extrapolation. SVM is a state-of-the-art tool in machine learning for linear and nonlinear input-output relation discovery [12]. The discovery task usually involves separating data into training and test datasets. Each sample in the training dataset contains a target value and several attributes. Based on the training dataset, SVM aims at building a model which is able to predict the target values of the test dataset given only attributes of the test dataset. The originality of the proposed method stems from the characteristics of SVM. Specifically, by setting the measured acoustic parameters, e.g., sound pressures, on the original aperture as the training dataset, where the target values correspond to the sound pressure values and the planar coordinate of each measurement point represents the attributes of this sample, a model is quickly generated using SVM without the involvement of a priori knowledge. Defining sound pressures on the exterior region as the test dataset and inputting their planar coordinates to the model, the values of the extrapolated sound pressures are rapidly obtained. Unlike LPBP and other patch NAH techniques where a large number of iterations is needed, a one-step procedure is sufficient to lead to a satisfactory extrapolation result with the proposed method.

2. SVM-based sound pressure extrapolation
The complex sound pressures on the planar measurement aperture serve as the training dataset of the proposed method. Specifically, \( x_i = (x_i, y_i) \) is the planar coordinate of the \( i \)th measurement point. Consider that the SVM algorithm is designed for real number computation, thus \( x_i \) is chosen to be the real part or the imaginary part of \( p(x_i, y_i, z_{ii}) \), corresponding to the regression function \( f_{\text{Re}}(x_i) \) or \( f_{\text{Im}}(x_i) \).

The detailed procedure of the SVM-based sound pressure extrapolation is illustrated in Figure 1. Firstly, the planar coordinates and the acoustic quantities on the aperture \( A \) constitute the training dataset, which is put into SVM to learn and generate a regression function \( f_{\text{Re}}(x_i) \) or \( f_{\text{Im}}(x_i) \). Next, \( (x, y) \) on the extrapolation area \( A^+ \) is introduced into the regression model, leading to the
corresponding part of \( p_{A'}(x_i, y_i, z_H) \), that is, the real part by \( f_{Re}(x_i) \) and the imaginary part by \( f_{Im}(x_i) \). Further, adding the originally measured data \( p_A(x_i, y_i, z_H) \) to the predicted \( p_{A'}(x_i, y_i, z_H) \) yields the complete extrapolated sound pressure on the hologram plane \( p(x_i, y_i, z_H) \). To smoothen the data at the edge of \( A' \), a two-dimensional Tukey window is applied to the newly generated pressure, whose constant part is placed exactly over the original aperture. Finally, the windowed sound pressure \( \hat{p}(x_i, y_i, z_H) \) is used for the NAH reconstruction.

Figure 1. Synopsis of SVM-based sound pressure extrapolation.

3. Numerical simulation
The learning and generalization abilities of SVM depend on the selection of the kernel function and other parameters. Research shows that in the absence of a priori knowledge, the radial basis function is a better option than other kernel functions, e.g., linear, polynomial, and sigmoid. The radial basis function can suppress the high frequency components of data and make the analysis results smoother. Furthermore, the cross-validation method is used to determine the SVM parameters. As a library for SVM, LIBSVM was developed by Lin et al. and is currently one of the most widely used SVM software packages [13]. In this paper, LIBSVM is utilized to predict the acoustic quantities in the extension area with the model obtained by using the measured data as the training set.

3.1. Modal source
In this section, numerical results obtained by using a source comprising a point-driven, simply supported aluminum plate in an infinite baffle are presented to demonstrate the applicability of the approach described above.

The dimension of the rectangular plate is 0.5 m × 0.5 m with a thickness of 3 mm. Its density is 2.7 g/cm³. The Young's modulus is 7.1×10¹⁰ Pa and the Poisson's ratio is 0.33. The force of 4 N is exerted at (0.1 m, 0.4 m). The excitation frequency is chosen to be 1155 Hz, which is approximately the natural frequency of mode \((4, 3)\).

Figure 2 indicates the simulation set-up. The source is located in a plane at \( z_s = 0 \). A microphone array is laid in the hologram plane at \( z_H = 4 \) cm. Considering the trade-off between the computation and the spatial resolution, the hologram plane consists of 41×41 points with a 1.25-cm lattice spacing in the \( x \)- and \( y \)-directions. The central 21×21 points are selected to be the first aperture \( A_1 \), while the second aperture \( A_2 \) covers 21×21 points on its top right corner. They are expanded to four times the
original size, that is, double the length and similarly for the width (21×21→41×41). The reconstruction plane lies at \( z_R = 2 \) cm. The areas on the reconstruction plane corresponding to \( \Lambda_1 \) and \( \Lambda_2 \) are \( R_1 \) and \( R_2 \), respectively. The sound pressure distribution of the modal source is shown in Figure 3(a). Its hologram pressure in Figure 3(b) is determined by Rayleigh's second integral formula. The two apertures are marked with white boxes.

The second case keeps most configurations the same as in the previous example, except for replacing the modal source with 10 discrete points as represented in Figure 3(c) and changing the excitation frequency to 1000 Hz. The corresponding sound pressure on the hologram plane is given in Figure 3(d). Since point sources contain energy over a broader wavenumber domain, they are a more challenging type of source to reconstruct with Fourier-based NAH.

![Figure 2. Schematic representation of the modal source case.](image)

(a) Pressure @ \( z_S \) (modal)  
(b) Pressure @ \( z_H \) (modal)  
(c) Pressure @ \( z_S \) (point)  
(d) Pressure @ \( z_H \) (point)

![Figure 3. Spatial pressure fields of modal source (top) and point sources (bottom) on the source plane at \( z_S = 0 \) (left) and on the hologram plane at \( z_H = 4 \) cm (right).](image)
For comparison, the reconstructions with zero-padding (ZP) and LPBP processed sound pressures on the hologram plane are considered. To exclude the impact of regularization, the same $k_c$ and $\gamma$ that are determined by L-curve [11] are applied to each case.

Figure 4 shows the comparison of three different reconstructions to the theoretical solutions on the areas corresponding to both apertures. With respect to $R_1$, we can see that although the two main acoustic radiation zones corresponding to yellow are clearly shown in each reconstruction, the result using ZP (Figure 4(b)) or LPBP (Figure 4(c)) is not as accurate as with SVM (Figure 4(d)). Specifically, the acoustic radiation at the top and bottom edges is reduced in both size and strength for the reconstruction using ZP or LPBP relative to the exact result (Figure 4(a)). Moreover, the area of the main radiation spots in Figure 4(b) or 4(c) is smaller than that in Figure 4(a).

Figure 4. Three reconstructions in comparison with the exact result on $R_1$ (top) and on $R_2$ (bottom) for the modal source.

Regarding $R_2$ at the bottom of Figure 4, the reconstruction using ZP (Figure 4(f)) has a similar problem to that on $R_1$. Specifically, the four yellow main radiation zones are clear but not well arranged compared with the exact result (Figure 4(e)). In addition, its acoustic radiation at the bottom edge moves up relative to the corresponding area in the exact result. In contrast with the ZP result, the reconstructions using LPBP (Figure 4(g)) and SVM (Figure 4(h)) show a certain improvement. That is, a sound pressure distribution close to the real value is described.

3.2. Point sources

Figure 5 describes the performance of the three reconstructions in the second case. Relative to the result on $R_1$ for the modal source, the reconstruction using ZP (Figure 5(b)) shows a poorer performance here. The shape and value of the large central radiation zone in the reconstruction using ZP are apparently different from the reference (Figure 5(a)). In terms of the four small radiation spots in the four corners of the exact result, they are moved towards the center of $R_1$ and expanded to a long and narrow area in Figure 5(b). On the contrary, the reconstruction using LPBP (Figure 5(c)) performs much better. Although the central radiation zone is somewhat expanded to the bottom and the sound pressure in the four corners has a slight increase, its sound pressure distribution is close to the exact result. In comparison with the result using LPBP, the reconstruction using SVM (Figure 5(d)) has no significant differences except a small decrease of the sound pressure in the top right corner and at the bottom edge, which makes it closer to the real value.
Regarding $R_1$, although the large radiation zone is shown in the reconstruction using ZP (Figure 5(f)), it has a different shape and the sound pressure extends to the right edge compared to the exact result (Figure 5(e)). In addition, a long and narrow area with the relatively large sound pressure which does not exist in the exact result appears at the top and bottom edges. However, the reconstruction using either LPBP (Figure 5(g)) or SVM (Figure 5(h)) leads to a more accurate reconstruction. Yet, taking the exact result as reference, the sound pressure near the bottom left corner in the SVM reconstruction is slightly better than that in the reconstruction using LPBP.

A quantitative measure to calculate the error between the theoretical value and the reconstructed result is defined as

$$E = \frac{\| p(x, y, z_R) - p_{\text{rec}}(x, y, z_R) \|_2}{\| p(x, y, z_R) \|_2} \times 100\%$$  \hspace{1cm} (1)

where $p(x, y, z_R)$ is the exact pressure on the reconstruction area and $p_{\text{rec}}(x, y, z_R)$ the reconstructed pressure over the same region.

The effects of the three pre-processing methods are compared quantitatively in Figure 6. As illustrated above, the reconstruction for point sources generally has a larger error than that for the modal source. Moreover, compared to the result with LPBP or SVM, the reconstruction using ZP leads to a much larger error for both types of sound source. However, the performance of the reconstruction using SVM is better than using LPBP in the same case. It should be related to the good small-sample learning ability of SVM. This quantitative analysis is consistent with our qualitative observation with Figures 4 and 5.

![Figure 5. Three reconstructions in comparison with the exact result on $R_1$ (top) and on $R_2$ (bottom) for the point sources.](image)

![Figure 6. Reconstruction errors on the reconstruction plane with three processing ways for the modal source (a) and the point sources (b).](image)
Besides the reconstruction accuracy, another factor of interest is the processing speed. Consider that the reconstruction improvement by using ZP is limited, LPBP and SVM are compared here. The speed test is carried out in a MATLAB R2014a environment on an average speed PC system (Intel Core i5-2500 CPU 3.3 GHz with 8 GB memory). The extrapolation process in our cases, that is, from a matrix of $21 \times 21$ to another of $41 \times 41$, requires about 9.4 s for LPBP and 3.2 s for SVM.

4. Conclusions

For conventional NAH, the reconstruction errors induced by the finiteness of the measurement aperture and the discrete solution are inevitable. Although a series of techniques have been developed to settle this issue, they still have some shortcomings. In this paper, a data extrapolation method is proposed to reduce these errors. It treats the measured pressure on the original aperture as the training dataset and makes use of SVM to derive a regression model and extrapolate the sound field outside the aperture. Numerical simulations demonstrate that the new method can give access to a quick and reliable reconstruction, qualitatively and quantitatively.

Acknowledgements

The authors thank to partly financial support from the National Natural Science Foundation of China (No. 11904406) and the 67th China Postdoctoral Science Foundation (No. 2020M671089).

References

[1] Williams E G and Maynard J D 1980 Holographic imaging without the wavelength resolution limit Physical Review Letters 45(7) 554-557
[2] Zhou H F, Lopez-Arteaga I, et al. 2016 Broadband planar nearfield acoustic holography based on one-third-octave band analysis Applied Acoustics 109 18-26
[3] Rowell M A and Oldham D J 1995 Determination of the directivity of a planar noise source by means of near field acoustical holography, 1: Theoretical background Journal of Sound and Vibration 180(1) 99-118
[4] Maynard J D, Williams E G and Lee Y 1985 Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH The Journal of the Acoustical Society of America 78(4) 1395-1413
[5] Saijyou K and Yoshikawa S 2001 Reduction methods of the reconstruction error for large-scale implementation of near-field acoustical holography The Journal of the Acoustical Society of America 110(4) 2007-2023
[6] Williams E G 2003 Continuation of acoustic near-fields The Journal of the Acoustical Society of America 113(3) 1273-1281
[7] Williams E G 2003 Approaches to patch NAH Proceedings of International Congress and Exposition on Noise Control Engineering 2003 2187-2194
[8] Williams E G, Houston B H and Herdic P C 2003 Fast Fourier transform and singular value decomposition formulations for patch nearfield acoustical holography The Journal of the Acoustical Society of America 114(3) 1322-1333
[9] Lee M and Bolton J S 2005 Patch near-field acoustical holography in cylindrical geometry The Journal of the Acoustical Society of America 118(6) 3721-3732
[10] Lee M and Bolton J S 2007 Reconstruction of source distributions from acoustic pressures measured over discontinuous regions: multipatch holography and interpolation The Journal of the Acoustical Society of America 121(4) 2086-2096
[11] Scholte R, Lopez Arteaga I, et al. 2009 Truncated aperture extrapolation for Fourier-based nearfield acoustic holography by means of border-padding The Journal of the Acoustical Society of America 125(6) 3844-3854
[12] Vapnik V N 1998 Statistical Learning Theory Wiley
[13] Chang C C and Lin C J 2011 LIBSVM: a library for support vector machines ACM Transactions on Intelligent Systems and Technology 2(3) 1-26