Study on Separation of Underwater Vehicle Noise
Based on Blind Source Separation

Xiao-bo NIU\textsuperscript{1} and Guo-qiang FU\textsuperscript{2}

\textsuperscript{1}Naval Petty Officer Academy, Bengbu, China
\textsuperscript{2}Beijing Aerospace Unmanned Vehicles System Engineering Research Institute, Beijing, China

Keywords: Blind source separation, Kernel independent component analysis, Underwater vehicle.

Abstract. Aiming at the study of underwater vehicle noise extraction in the measurement, a method based on blind source separation (BBS) is presented, and the kernel independent component analysis (KICA) is used in the separation of underwater vehicle noise. The principle and algorithm is introduced and the simulation is carried on to compare with the traditional ICA algorithm in BBS. The simulation results show that KICA not only separates the underwater vehicle noise effectively but also is more accurate than the traditional ICA algorithm.

Introduction

Underwater vehicle shock absorption and noise reduction is a hot topic in many countries, while how to improve the noise test level of underwater vehicle is an important means of noise reduction. The navigational radiation noise of an underwater vehicle is mainly composed of machinery noise, propeller noise and hydrodynamic noise. The vibration of internal mechanical equipment of the underwater vehicle that radiates into the seawater through vibration or excitation of vehicle hull is called machinery noise. Propeller noise is the noise generated by the propeller rotating in seawater, including propeller cavitation noise, propeller rotation and blade vibration. Hydrodynamic noise is the noise generated by the relative motion of the surface of the vehicle hull and the aqueous medium during navigation by the underwater vehicle.

A variety of methods are adopted to comprehensively analyze and identify noise sources, to effectively identify the distribution of major noise sources and separate the contribution of machinery noise, hydrodynamic noise and propeller noise to radiation noise. Therefore, it is of great significance to obtain the overall acoustic performance of the ship under test and evaluate the noise design indexes for noise control. So, how to separate the data of different noise sources from the acquired data is an important research content.

Blind Source Separation and Kernel Independent Component Analysis

Independent component analysis (ICA) is a typical blind source separation method. Its essence is to blind source separation of multi-channel observation signals under the assumption of statistical independence and dig out independent source components hidden in the observation signal. In 2002, F. R. Bach, M. I. Jordan et al. introduced the idea of kernel function on the basis of traditional ICA theory, and proposed a new type of ICA algorithm: Kernel Independent Component Analysis (KICA) to further improve the separation effect.

Basic Principles of Kernel Independent Component Analysis

Kernel independent component analysis is not a combination of existing independent component analysis and kernel function, but a new method. The idea is to use nonlinear mapping $\Phi : R^k \rightarrow R$ to map the nonlinear variables of original input space to a certain Reproducing Kernel Hilbert Space. Then, the mapped data are analyzed in this feature space. Traditional ICA algorithms use a single fixed nonlinear expectation function as a comparison function of component independence, such as kurtosis, negentropy and mutual information. KICA, on the hand, uses kernel function instead of
inner product operation between two input vectors to realize nonlinear transformation in Reproducing Kernel Hilbert Space. Therefore, it is not necessary to construct specific nonlinear expectation function.

**Mercer Kernel**

Any function that meet Mercer’s condition can be used as Mercer kernel and decomposed into dot product form in feature space. The Mercer condition can be described as Eq. 1 for any square-integrable function:

$$\int_{L_1 \times L_2} K(x, y)g(x)g(y)dxdy \geq 0$$

Some common Mercer kernel functions include:

**Polynomial kernel:**

$$K(x, y) = (x \cdot y + 1)^d, \quad d \text{ is a constant}$$

**Gaussian kernel:**

$$K(x, y) = \exp \left( -\frac{(x-y)^2}{2\sigma^2} \right), \quad \sigma > 0$$

**Kernel Independent Component Analysis Algorithm (KICA Algorithm)**

Let \(x_1, x_2 \in R^N, \ F\) be a function vector space from \(R\) to \(R\), and the correlation coefficient \(\rho_F\) of \(x_1\) and \(x_2\) is defined as the maximum correlation coefficient between the random variable \(f(x_1)\) and \(f(x_2)\)

$$\rho_F = \max_{f_1, f_2 \in F} \text{corr}(f_1(x_1), f_2(x_2))$$

$$= \max_{f_1, f_2 \in F} \frac{\text{cov}(f_1(x_1), f_2(x_1))}{(\text{var} f_1(x_1))^{1/2}(\text{var} f_2(x_2))^{1/2}}$$

\(\rho_F\) is a comparison function between random variables. Obviously, if \(x_1\) and \(x_2\) is independent, then \(\rho_F = 0\). If the space \(F\) is large enough, the opposite is true. The idea of Reproducing Kernel Hilbert Space is used to calculate the \(\rho_F\). Let \(F\) to be the Reproducing Kernel Hilbert Space on the real number set. \(K(x, y)\) is a kernel associated with this space. Let \(\Phi(x) = K(\cdot, x)\) as a mapping from the input space to \(F\), \(K(\cdot, x)\) is a functional function with \(x\) as the variable in \(F\) space. With regeneration characteristics of kernel Hilbert space:

$$f(x) = \left\{ \Phi(x) \cdot f \right\}, \quad \forall f \in F, \forall x \in R$$

Then,

$$\text{corr}(f_1(x_1), f_2(x_2)) = \text{corr}\left(\left\{ \Phi(x_1) \cdot f_1 \right\}, \left\{ \Phi(x_2) \cdot f_2 \right\}\right)$$

Let \([x_1^1, x_2^2, \ldots, x_1^N]\) and \([x_1^1, x_2^2, \ldots, x_2^N]\) to be the vectors composed of \(N\) observation values of random variable \(x_1, x_2\) respectively. \([\Phi(x_1^1), \Phi(x_2^1), \ldots, \Phi(x_1^N)]\) and \([\Phi(x_1^1), \Phi(x_2^1), \ldots, \Phi(x_2^N)]\) are the function values of \(F\). Assuming that they have been centralized, i.e., \(\sum_{i=1}^{N} \Phi(x_i^1) = \sum_{i=1}^{N} \Phi(x_i^2) = 0\). Then, Bach in literature [3] proves that the former formula can be written as:
\[ \rho_p(K_1, K_2) = \max_{a_1, a_2 \in \mathbb{R}^N} \frac{a_1^T K_1 K_2 a_2}{(a_1^T K_1^2 a_1)^{1/2} (a_2^T K_2^2 a_2)^{1/2}} \]  

(6)

Solving the above formula is equivalent to solving the following formula:

\[
\begin{bmatrix}
0 & K_1 K_2 \\
K_2 K_1 & 0
\end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \rho \begin{bmatrix} K_1^2 & 0 \\
0 & K_2^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}
\]

(7)

Where, \( K_1 \) and \( K_2 \) are the Gram matrices of the observed data, \( a_1, a_2 \) are the eigenvectors of \( K_1 \) and \( K_2 \) respectively. Eq. 7 can be extended to multiple variables:

\[
\begin{bmatrix}
K_1 K_1 & K_1 K_2 & \cdots & K_1 K_n \\
K_2 K_1 & K_2 K_2 & \cdots & K_2 K_n \\
\vdots & \vdots & \ddots & \vdots \\
K_n K_1 & K_n K_2 & \cdots & K_n K_n
\end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \lambda \begin{bmatrix} K_1^2 & 0 & \cdots & 0 \\
0 & K_2^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & K_n^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}
\]

(8)

**Algorithm Steps of KICA**

\( x_1(t), x_2(t), \ldots, x_M(t) \) are \( M \) observation vectors for mixed signals, \( K(x, y) \) is a predetermined kernel function, \( W \) is a randomly generated initialization decomposition matrix.

Step 1: The observed data was undergone whitening process;

Step 2: Let \( Y = WX \), obtain the estimated value \( y_1, y_2, \ldots, y_N \) of the source signal \( s_1, s_2, \ldots, s_N \), and calculate the estimated Gram matrix \( K_1, K_2, \ldots, K_N \) of estimated value by Cholesky decomposition;

Step 3: Defined as the maximum eigenvalue of the following formula:

\[
\begin{bmatrix}
K_1 K_1 & K_1 K_2 & \cdots & K_1 K_n \\
K_2 K_1 & K_2 K_2 & \cdots & K_2 K_n \\
\vdots & \vdots & \ddots & \vdots \\
K_n K_1 & K_n K_2 & \cdots & K_n K_n
\end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \lambda \begin{bmatrix} K_1^2 & 0 & \cdots & 0 \\
0 & K_2^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & K_n^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}
\]

(9)

Step 4: Set the objective function \( C(W) = -\frac{1}{2} \log \lambda_W \), and to minimize it;

Step 5: Make a convergence judgment, if \( W \) does not converge, return to the second step;

Step 6: Use \( Y = WX \) to estimate the source signal.

**Simulation Verification of Algorithm Feasibility**

In this paper, the submarine vibration simulation data in literature [7] is used to carry out simulation verification of the algorithm to verify the feasibility of the algorithm. As shown in Fig. 1, a sinusoidal signal was used to simulate power frequency interference, a pulse signal and a sawtooth wave signal were used to simulate two different kinds of mechanical vibration noise signals, and a Gaussian white noise signal was used to simulate the background noise of the marine environment.
A mixed matrix is generated randomly, and the matrix composed of four signal values is multiplied by the mixed matrix, that is, the source signal is physically transmitted through a transmission channel which response function is unknown. The results obtained are shown in Fig. 2. At this time, it is difficult to distinguish the characteristics of the original signal. The traditional ICA algorithm based on cumulants and KICA algorithm were used to calculate the randomly mixed signals, and the results obtained successfully separated the simulated source signals that participated in the verification calculation, as shown in Fig. 3 and Fig. 4.
Through simulation verification using simulated data, both traditional ICA algorithm and KICA algorithm can effectively separate all source signals from mixed signals. The results obtained by KICA separation are less affected by noise interference than those obtained by traditional ICA, and the results are more accurate. Although some signals are found to be flipped on the time domain feature in the processing result, and the order of each source signal is changed, this is caused by the principle of independent component analysis algorithm. At present, there is no effective method to solve this problem, but this phenomenon does not affect the estimated frequency domain feature of each source signal.

Summary
In this paper, based on the introduction of the basic principle and specific algorithm of KICA, the performance of traditional ICA and KICA are verified and compared quantitatively. The results show that KICA has better signal separation performance than the traditional ICA algorithm. After that, the algorithm is applied to the separation calculation of underwater vehicle navigation noise signal. After the simulation verification of simulation data, the signal of underwater vehicle vibration noise source is separated, and the feasibility and effectiveness of the algorithm are verified.

References
[1] S.J. Zhu, L. He. Vibration Control of Marine Machinery. National Defense Industry Press, Beijing, 2006.
[2] J.C. Ma, Y.L. Niu, H. Y. Chen. Blind Signal Processing. National Defense Industry Press, Beijing, 2006.
[3] F.R. Bach, M.I. Jordan. Kernel independent component analysis. Journal of Machine Learning Research, 3(2002) 1-48.
[4] G. Marc G. Classes of kernels for machine learning: a statistics perspective. Journal of Machine Learning Research, 2(2001) 299-312.
[5] B. Scholkopf, A.J. Smola. Learning with kernels. The MIT Press, Cambridge, 2001.
[6] S. Saitoh. Theory of reproducing kernels and its application. Longman Scientific & Technical, Harlow, 1988.
[7] W. Cheng, Z.S. Zhou, Z.J. He. Quantitative Calculation of Vibration Source Contributions of Submarines Based on Independent Component Analysis, Chinese Journal of Mechanical Engineering, 46(2010) 82-87.