International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.9 (2016), pp.211-220
http://dx.doi.org/10.14257/ijmue.2016.11.9.22

Research on Sound Source Localization for Device Fault Diagnosis*

Liang Zhang [1,2,a], Tiejun Li [1,b,*], Yinxue Zong[1] and Yang Tan [1]

[1]*School of Mechanical Engineering, Hebei University of Technology, Tianjin, 300130, China*
[2]*School of Mechanical Engineering, Jiujiang Vocational and Technology College, Jiujiang, 332000, China*
[a.]zhangliang860226@163.com(L.Z.), leochang315@163.com(T.L.),
[b.]zongyinxue@126.com(Y.Z.), tanyangstar@gmail.com(Y.T.)
**Corresponding Author*

Abstract

A sound source localization method is presented herein around the utilization with the sound signal for device fault diagnosis. Firstly, Butterworth wavelet and its filter banks were designed aimed at the background noise separation in the industrial environment, it is proved that the Butterworth wavelet filter can extract the weak fault features in source signals without frequency aliasing phenomenon. Then the algorithms of sound source localization was studied in depth, the generalized cross-correlation(GCC) based time delay estimation(TDE) was presented. Finally the coordinates of sound source location was obtained in the quaternion microphone array model. The experiment results show that in the industrial environment, the error value of positioning accuracy is less than 4cm within an appropriate distance, it can meet the needs of practical application, meanwhile the reliability and the real-time performance is satisfactory.

***Keywords**: Fault Diagnosis; Sound Source Localization; Butterworth Wavelet; Generalized Cross-correlation; Time Delay Estimation*

1. Introduction

With the rapid development of signal processing technology, the utilization with the sound signal in device fault diagnosis is becoming a research hotspot [1]. Fault feature extraction and sound source localization algorithm are the key problems in practical engineering application. Sound source localization technology is a new kind of non-contact detection technology [2], it can be used to confirm the origins and the intensity of the device fault sounds, speculate the fault positions, analyze the failure modes and guide the maintenance of devices.

However, in the industrial environment, there are scattering and reverberation in sound signals, so the accuracy of sound source positioning is seriously affected, and this is the main reason that why the technology is not widely used in the area of fault diagnosis.

A kind of sound source localization method for device fault diagnosis is proposed herein. And the content is as follows. Firstly, in Section 2, the filter combined with Butterworth filter and wavelet denoising is constructed to solve the aliasing and reverberation of source signals in the industrial environment. In Section 3, the method of generalized cross correlation algorithm based time delay estimation is presented. Next in Section 4, the coordinates of fault location in the microphone quaternion matrix model are obtained. In Section 5, it is proved that the proposed method is feasible through an example. Finally in Section 6, the results are summarized.

ISSN: 1975-0080 IJMUE

2. Butterworth-wavelet Filter Banks

Butterworth filter is a kind of common IIR filter with the characteristic of steep frequency cut-off, and its performance is better with the degree increases [3], it is widely used in conventional signal denoising processing. However, in the industrial environment, the spectral analyses of the sound signals of running devices show that the source signal frequencies of the devices and the background noises are overlapping, and accompany by reverberations and scattering normally, make the conventional Butterworth filter difficult to achieve the ideal effect. So a kind of Butterworth and wavelet combined filter is designed in this paper to solve the above-mentioned problems. The wavelet transform has good effect in denoising, it can extract the weak characteristic in source signals, and meanwhile using the frequency cut-off characteristic of Butterworth filter suppresses the frequency aliasing phenomenon in wavelet transform [4].

Set $\Psi(x)$ as wavelet function, and its scale function is $\Phi(x)$. The low pass filter related to $\Phi(x)$ is represented with $H_0(z)$, while the high pass filter related to $\Psi(x)$ is represented with $H_1(z)$, if $H_0(z)$ and $H_1(z)$ meet the formula below,

$$H_0(z)\,\mathrm{H}_0(z^{-1}) + H_0(-z)\,H_0(-z^{-1}) = 1 \tag{1}$$

$$H_1(z) = z^{-(N-1)} H_0(-Z^{-1}) \tag{2}$$

The wavelet function $\Psi(x)$ is the orthogonal wavelet. Namely constitute a pair of conjugate quadrature mirror filter banks (CQMFB)[5].

So the orthogonal wavelet design is the constructive process of a pair of orthogonal wavelet filter banks. Its core lies in design of $H_0(z)$ in formula(1). Once $H_0(z)$ is established, $H_1(z)$ can be got according to equation (2). And plug $z=e^{iw}$ into equation (1), it is expressed as

$$|H_0(e^{iw})|^2 + |H_0(e^{i(w+\pi)})|^2 = 1 \tag{3}$$

And the orthogonal wavelet filter banks can be constructed by the filters meet formula (3).

In the field of mechanical fault diagnosis, when extract fault characteristics, DB wavelet series belong to FIR filter banks are generally chosen[6, 7], but the amplitude-frequency characteristics of DB wavelet series are not ideal enough. After the transformation, there are considerable overlaps among each sub-band spectrum, it will bring more noise to each sub-band, and plus interferences among the fault features. So a pair of orthogonal wavelet filter banks based on Butterworth filter herein, and a kind of orthogonal wavelet is given, the specific processes are described as follows[8],

1) Design a Butterworth filter $H_0(z)$ which meet equation (3), the analog N-rank Butterworth low-pass filter of the amplitude square function is expressed as

$$|H_0(\Omega j)|^2 = 1/[1+(\Omega/\Omega_c)^{2N}] \tag{4}$$

Plug the bilinear mapping relation $\Omega=2\tan(w/2)/T_s$ into equation(4), and it is expressed as

$$|H_0(e^{jw})|^2 = 1/\{1+[\tan(w/2)\tan(w_c/2)]^{2N} \tag{5}$$

If $w_c=0.5\pi$, then

$$|H_0(e^{jw})|^2 = 1/\{1+[\tan(w/2)]^{2N} \text{ and } |H_0(e^{j(w+\pi)})|^2 = 1/\{1+[\tan(w/2)]^{-2N} \tag{6}$$

The following equation is obtained as

$$|H_0(e^{jw})|^2 + |H_0(e^{j(w+\pi)})|^2 = 1 \tag{7}$$

So for meeting the equation (7), the cut-off frequency is 0.5π.

2) In formula (2), for N=2, so $H_1(z) = z^{-1}H_0(-Z^{-1})$. Because of that $H_1(z)$ is the Butterworth filter, and its cut-off frequency is 0.5π. So $H_1(z)$ and $H_0(z)$ constitute a pair of orthogonal wavelet filter banks.

3) Butterworth function $\Psi(x)$ and its scaling function $\Phi(x)$ can be iteratively calculated [9] by the follow formulas as

$$\begin{cases} \Psi(z) = \frac{1}{\sqrt{2}}\left(\frac{z}{2}\right)\prod_{j=2}^{+\infty}\frac{H_0\left(z/_{2j}\right)}{\sqrt{2}} \\ \Phi(z) = \prod_{j=1}^{+\infty}\frac{H_0\left(z/_{2j}\right)}{\sqrt{2}} \end{cases} \tag{8}$$

But the analytical solutions of Wavelet function and its scaling function can't be obtained normally, so the approximate solution is got by iterative numerical convolution of unit impulsion response $h_0(n)$ of $H_0(z)$. In order to ensure convergence of the iterative process, the low-pass filter needs some null points when z=-1, and according to formula (7), there are N null points. So the Butterworth wavelet function and its scaling function can be obtained by iterative operation of $h_0(n)$ [10].

3. The Generalized Cross-correlation Based Time Delay Estimation

Steered beam former, high resolution spectrum estimation, sound pressure amplitude ratio based positioning and time delay estimation (TDE) is the common methods of sound source localization. Each method has its advantages and disadvantages, so it should be chosen flexibly in accordance with different applications. Electromechanical device condition monitoring needs conduct real-timely, and the failures always occur serially. So TDE based method is chosen herein because its small calculating amount and superiority for single near field single sound source.

The distance between the any two microphones in microphone array is very short, and the time delay is measured in the microsecond range. So it's difficult to get the accurate calculation, but is estimated in some certain ways [11]. In sound source positioning, Time-delay estimation (TDE) is extremely important, and its value is positively correlation with the location result. In the device run-time, it's generally believed sound signal propagates in the form of spherical wave, aimed to simplify the algorithm, the sound source and the microphone array are perceived in the same plane, so the model can be transformed to two-dimensional surface shown in Figure 1.

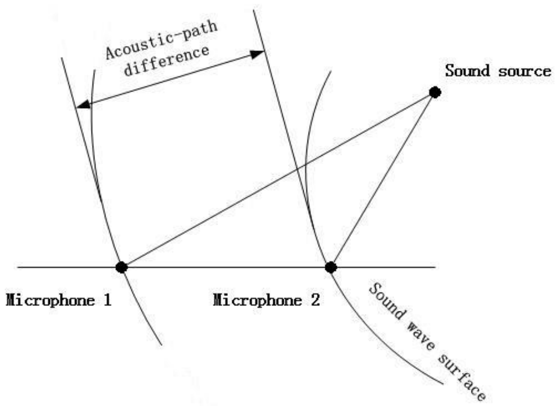

Figure 1. The Schematic of Time Delay Estimation

In the two-dimensional plane model, the surface of sound vibration wave has certain acoustic path difference between the two microphones, and the time delay is generated. Set the time delay between the two microphones is τ_{12}, and the ideal formula is expressed as

$$x_1(t) = \frac{1}{d_1}s(t) + v_1(t) \tag{9}$$

$$x_2(t) = \frac{1}{d_2}s(t-\tau_{12}) + v_2(t) \tag{10}$$

Where, the signals that received by different microphones are represented with $x_1(t)$ and $x_2(t)$, $s(t)$ is the signal from the sound source, d_1 *and* d_2 are the attenuation coefficients of the sound wave, $v_1(t)$ and $v_2(t)$ are two irrelevant noises, assuming that the sound source and the noises are independent of each other, according to equation(9) and equation(10), the cross correlation function can be expressed as

$$R_{12}(\tau_{12}) = E[x_1(t)x_2(t)] = E\{\left[\frac{1}{d_1}s(t) + v_1(t)\right]\left[\frac{1}{d_2}s(t-\tau_{12}) + v_2(t)\right]\}$$

$$= E\left[\frac{1}{d_1}s(t)\frac{1}{d_2}s(t-\tau_{12})\right] + E\left[\frac{1}{d_2}s(t-\tau_{12})v_1(t)\right] + E\left[\frac{1}{d_1}s(t)v_2(t)\right] + E[v_1(t)v_2(t)] \quad (11)$$

Because $s(t)$, $v_1(t)$ and $v_2(t)$ are irrelevant, the values of the last three terms of equation(11) are all zero, then

$$R_{12}(\tau_{12}) = \frac{1}{d_1 d_2}E[s(t)s(t-\tau_{12})] \quad (12)$$

According to the nature of cross-correlation function, when $R_{12}(\tau_{12})$ has the maximal value, $t=\tau_{12}$. And it is the time delay value between the two microphones, this is the method of basic cross correlation, it can fast calculate the value of time lag conveniently. However, it is sensitive to sound reflection and reverberation, so in the industrial environment, it is difficult to ensure the estimated accuracy. Aiming at the problems, a developed method based on generalized cross correlation function is proposed to reduce the noisy impact and inhibit the reverberation phenomenon.

The process of generalized cross correlation method can be described as follows [12]. Firstly, the two sound signals is pre-filtered. Then mutual power spectrum function is calculated and weighted. Finally, inverse Fourier transform is used to the above results and the generalized cross-correlation function will be obtained. Among them, weighting can suppress the noise and reverberation, the time difference value of the two signals is got at the peak of cross-correlation function.

According to equation (12), the generalized cross-correlation function is expressed as

$$R_{12}(\tau_{12}) = \int_{-\infty}^{\infty} \Psi_{12}(\omega)\phi_{x_1x_2}(\omega)e^{-j\omega\tau_{12}}d\omega \quad (13)$$

Where, $\Psi_{12}(\omega)$ is mutual power spectrum function, $\Phi_{x1x2}(\omega)$ is the power spectrum of the both two signals.

Generalized cross-correlation method can reduce the calculation, if the prior knowledge of background noise was known, the noise and reverberation can be suppressed through the weighed function of mutual power spectrum [13]. Accuracy can be ensured in the case of high signal noise ratio (SNR), and it can also be improved through conditioning the weighted function if the SNR is low. In general, the method can satisfy the request of real-time computation for acquisition of electromechanical device sound signal.

4. Coordinates Calculation of Sound Source Position

The research is based on microphone quaternion array mode herein, and it is shown in Figure 2:

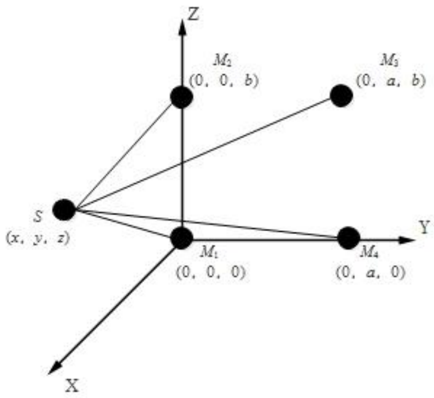

Figure 2. A Kind of Microphone Quaternion Array Model

In the microphone array coordinate system, S is the acoustic source , and its coordinate is $S(x, y, z)$, while the four microphones are $M_1(0, 0, 0)$, $M_2(0, 0, b)$, $M_3(0, a, b)$, $M_4(0, a, 0)$. Set M_1 is the standard microphone, and the time delays between other three microphones and M_1 are τ_{21}, τ_{31} and τ_{41}. The distances between the sound source S and the four microphones are r, d_2, d_3, d_4, and the sound source of running device is considered as a point source in near field, so sound wave is spherical wave, set the speed of sound in air is c, and the following equations can be obtained,

$$\begin{cases} d_{21} = d_2 - r = \tau_{21}c \\ d_{31} = d_3 - r = \tau_{31}c \\ d_{41} = d_4 - r = \tau_{41}c \end{cases} \tag{14}$$

$$\begin{cases} x^2 + y^2 + z^2 = r^2 \\ x^2 + y^2 + (z-b)^2 = (r + d_{21})^2 \\ x^2 + (y-a)^2 + (z-b)^2 = (r + d_{31})^2 \\ x^2 + (y-a)^2 + z^2 = (r + d_{41})^2 \end{cases} \tag{15}$$

Solve the equation sets, it is expressed as

$$\begin{cases} -2bz - 2\tau_{21}cr = \tau_{21}{}^2c^2 - b^2 \\ -2ay - 2bz - 2\tau_{31}cr = \tau_{31}{}^2c^2 - b^2 - a^2 \\ -2ay - 2\tau_{41}cr = \tau_{41}{}^2c^2 - a^2 \end{cases} \tag{16}$$

The equation set(16) can be transformed to up the form triangular matrix,

$$\begin{bmatrix} 2a & 0 & 2(\tau_{31} - \tau_{21})c \\ 0 & -2b & -2\tau_{21}c \\ 0 & 0 & 2(\tau_{31} - \tau_{21} - \tau_{41})c \end{bmatrix} \begin{bmatrix} y \\ z \\ r \end{bmatrix} = \begin{bmatrix} (\tau_{21}{}^2 - \tau_{31}{}^2)c^2 + a^2 \\ \tau_{21}{}^2c^2 - b^2 \\ (\tau_{21}{}^2 + \tau_{41}{}^2 - \tau_{31}{}^2)c^2 \end{bmatrix} \tag{17}$$

For presentation purpose, equation (17) can be expressed as follows,

$$\begin{bmatrix} L_{11} & L_{12} & L_{13} \\ 0 & L_{22} & L_{23} \\ 0 & 0 & L_{33} \end{bmatrix} \begin{bmatrix} y \\ z \\ r \end{bmatrix} = \begin{bmatrix} U_{11} \\ U_{21} \\ U_{31} \end{bmatrix} \tag{18}$$

Equation(18) can be transformed to algebraic form,

$$\begin{cases} L_{11}y + L_{12}z + L_{13}r = U_{11} \\ L_{22}z + L_{23}r = U_{21} \\ L_{33}r = U_{31} \end{cases} \tag{19}$$

Solve the equations above,

$$\begin{cases} y = \frac{U_{11}}{L_{11}} - \frac{L_{13}U_{31}}{L_{11}L_{33}} \\ z = \frac{U_{21}}{L_{22}} - \frac{L_{23}U_{31}}{L_{22}L_{33}} \\ r = \frac{U_{31}}{L_{33}} \end{cases} \tag{20}$$

The value of x is obtained by equation (15),

$$x = \pm\sqrt{r^2 - y^2 - z^2} \tag{21}$$

The sound source points are all in the X-axis positive direction of the microphone array coordinate system, so the negative parts can be gave up, the terms of equation (18) are put in the corresponding term of equation (20) and equation (21), and the coordinate value of the sound source in microphone array coordinate system can be obtained.

5. Experimental Testing and Analysis

A vertical milling machine in a workshop was chosen as experimental subject, and the fault sound signals were from work pieces processing with a counter rotating cut. The sound source location is stationary, so aimed to simulate the positioning of different points, the experiments was operated in the form of moving the microphone array.

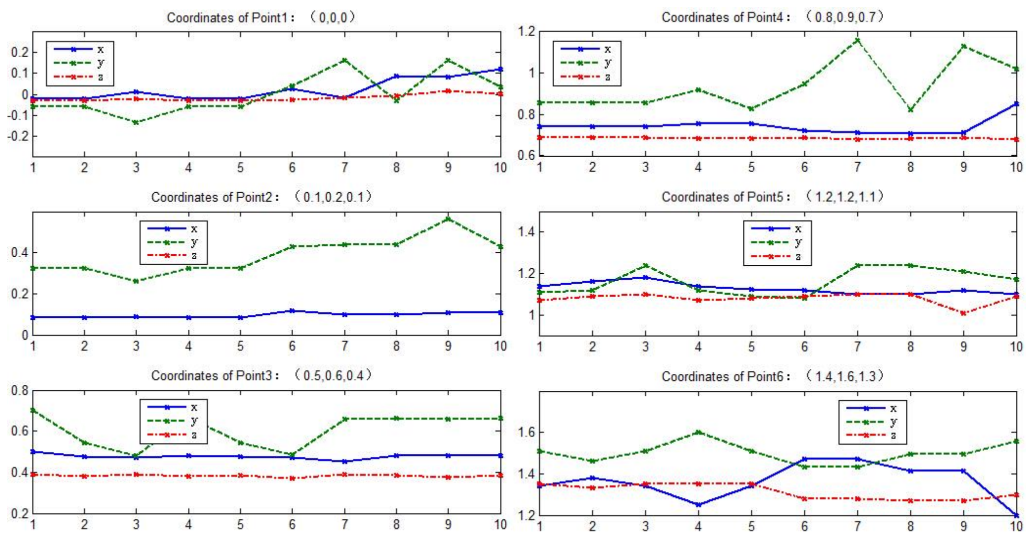

Figure 3. Coordinates Traces of the Measured Results (Unit: M)

Refer to the coordinate system of Figure 3, fifty groups of experiments was done to different positions, and six typical coordinate points as (0, 0, 0), (0.1, 0.2, 0.1), (0.5, 0.6, 0.4), (0.8, 0.9, 0.7), (1.2, 1.2, 1.1), (1.4, 1.6, 1.3) were chosen to present. The acquisition time was ten seconds at each points, the system calculated the position coordinate per second, so there are ten groups data at each point. The result is shown in Figure3.

In Figure 3, the x, y and z position traces of estimated locations are described with full line, imaginary line and chain line. Calculate the root mean square error and standard deviation separately of the data. Where, the root mean square error is expressed as formula (24), it represent the deviation value of the experimental result; The standard deviation is expressed as formula(25), it shows the discrete degree of the experimental data itself.

$$MSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - x_0)^2} \tag{24}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{25}$$

Where, x_i is the successive measurement of the sound source, x_0 is the real coordinate value, and $\bar{x}$ is the average value of x_i.

When the sound source position moves relative to the microphone array, it shows that when the distance of them is in a certain range, the stability and the accuracy of the location could maintain at a high level relatively. Meanwhile, the stability and accuracy both declines slowly with the increase of the distance. But when the distance is out of the range, the location accuracy and stability begin to decrease significantly, until both of them can't be identified. Through a large amount of tests, the critical value of the distance is obtained as 1.7m.

The root mean square errors of the six groups of data were revealed in Table 1.

Table 1. The Values of Root Mean Square Errors (Unit: M)

Sound Sources	Error of x	Error of y	Error of z
(0, 0, 0)	0.0112	0.0103	0.0130
(0.1, 0.2, 0.1)	0.0127	0.0109	0.0102
(0.5, 0.6, 0.2)	0.0212	0.0167	0.0239
(0.8, 0.9, 0.3)	0.0335	0.0396	0.0433
(1.2, 1.2, -0.2)	0.0216	0.0471	0.0307
(1.4, 1.6, -0.3)	0.0506	0.0527	0.4597

The standard deviations of the six groups of data were listed in Table 2.

Table 2. The Standard Deviations (Unit: m)

Sound Sources	Error of x	Error of y	Error of z
(0, 0, 0)	0.0157	0.0281	0.0036
(0.1, 0.2, 0.1)	0.0172	0.0299	0.0165
(0.5, 0.6, 0.2)	0.0219	0.0213	0.0179
(0.8, 0.9, 0.3)	0.0232	0.0331	0.0289
(1.2, 1.2, -0.2)	0.0322	0.0317	0.0372
(1.4, 1.6, -0.3)	0.0385	0.0395	0.0406

According to the results of the experiments, it's concluded that the accuracy and the stability can only be guaranteed in a certain distance of the sound source and the pickup, their values are also discrepant with different localization algorithms. Meanwhile, the localization accuracy of sound source is also contained of hardware conditions just like the number of the microphone, the array form, the acoustic field environment, and so on. In the experiment, if the distance between spot S and the origin is less than 1.7m, the value of root mean square error can be controlled within 5cm, and the value of standard deviation can be controlled within 4cm, the location accuracy and stability are both satisfactory, it can meet the needs of practical application.

6. Conclusion

This paper mainly put forward a method of sound source positioning for device fault diagnosis. Focus on the sound signals aliasing problem, the Butterworth-wavelet filter was designed to separate the background noise of source signals. Meanwhile, the quaternion microphone array model was established, and the developed algorithm based on generalized cross-correlation was proposed. The trial data demonstrate that in the industrial environment, within proper range in acoustic field, the localization accuracy error of the proposed method is less than 5cm for the abnormal sound of device, and the value of standard deviation is less than 4cm, it can realize the accurate positioning of the fault device sound source, and meanwhile the reliability and real-time performance are guaranteed. Thus, the method is feasible and worth popularizing.

Acknowledgements

The work is supported by the National Nature Science Foundation of China (No. 51175145).

References

[1] Z. Liu and R. L. Chen, "Sound source localization system based on planar microphone array", Technical Acoustics, vol. 30, **(2011)**, pp. 123-128.
[2] T. L. Ju, "Research of sound source localization algorithm based on microphone array", University of Electronic Science and Technology of China, **(2006)**.
[3] H. Cormac and V. Martin, "Wavelets and recursive filter banks, in: IEEE Trans on Signal Processing", vol. 41, **(1993)**, pp. 2536-2556.
[4] L. S. He and J. F. Wang, "Butterworth based on discrete wavelet transform and Its application in machine fault diagnosis", Huazhong Univ. of Sci. & Tech., vol. 28, **(2000)**, pp. 31-33.
[5] V. Martin and H. Cormac, "Wavelets and filter banks: theory and design. In: IEEE Trans on Signal Processing", vol. 40, **(1992)**, pp. 2207-2232.
[6] S. Prabhakar, A. R. Mohanty and A. S. Sekhar, "Application of discrete wavelet transform for detection of ball bearing race faults", Tribology International, vol. 25, **(2002)**, pp. 793-800.
[7] Z. K. Peng, M. R. Jackson and J. A. Rongon, "On the energy leakage of discrete wavelet transform", Mechanical Systems and Signal Processing, vol. 23, **(2009)**, pp. 330-343.
[8] X. Z. Zhao and B. Y. Ye, "Convolution wavelet packet transform its applications to signal processing", Digital Signal Processing, vol. 20, **(2010)**, pp. 1352-1364.
[9] S. G. Mallat, "A theory for multiresolution signal decomposition: wavelet representation", IEEE Trans on Pattern Analysis and Machine Intelligence, vol. 11, **(1989)**, pp. 675-693.
[10] R. Luo, F. Q. Tian and F. Zhou, "Construction of Butterworth wavelet and its application in fault diagnosis", Huazhong Univ. of Sci. & Tech., vol. 41, **(2013)**, pp. 68-73.
[11] X. Wu, S. J. Ji and Y. B. Li, "Sound source localization based on microphone arrays", Chinese Sensors and Actuators, vol. 23, **(2010)**, pp. 682-686.
[12] H. M. Lu, W. Chen and S. B. Liu, "Design of acoustic source localization system based on microphone array", Transducer and Microsystem Technologies, vol. 31, **(2012)**, pp. 79-85.
[13] X. L. Yu and C. G. Lv, "Research on speech localization in near field", Intelligent Computer and Applications, vol. 51, **(2011)**, pp. 69-76.

Authors

Zhang Liang, received his M.S. degree in control engineering from Huazhong university of Science and Technology in Wuhan, China. He is currently a doctor candidate in Hebei University of Technology in Tianjin. His research interest is mainly in the areas of Machine Vision, Multisensor Information Fusion and Artificial Intelligence. He has published 9 research papers in scholarly journals in the above research areas.

Li Tiejun, received his Ph.D degree in electromechanical control and automation from Harbin Institute of Industry in Harbin, China. He is currently a professor and a doctoral supervisor in Hebei University of Technology in Tianjin. His research interest is mainly in the areas of Artificial Intelligence, Robotics and Intelligent Device.

Zong Yinxue, received her bachelor's degree in Mechanical engineering from Hebei University of Technology in Tianjin, China. She is currently a master candidate in the same University. Her research interest is mainly in the areas of Signal processing, Sound Source Localization and Artificial Neural Network. She has published 5 research papers in scholarly journals in the above research areas.

Tan Yang, received his M.S. degree in Mechanical engineering from Hebei University of Technology in Tianjin, China. He is currently a doctor candidate in the same University. His research interest is mainly in the areas of Signal processing, Transducer Technology and Internet of Things.