Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 310 (2008) 134–148
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Leak location using blind system identification in water distribution pipelines

Jin Yang, Yumei Wen*, Ping Li

Department of Optoelectronic Engineering, Key Laboratory of Optoelectronic Technology and System of the Education Ministry of China, Chongqing University, Chongqing 400044, China

Received 15 July 2006; received in revised form 30 May 2007; accepted 28 July 2007
Available online 1 November 2007

Abstract

This paper presents a new leak location method without knowledge of pipe length in water distribution pipes. In conventional leak location surveys using acoustic signals in buried water distribution pipelines, the most widely used techniques are based on the correlation technique. The prerequisite of the method is that the accurate length of pipeline between two detection points is known. However, in practice, this prerequisite is not always satisfied. Based on the fact that the acquired acoustic signals contain the characteristics related to the two propagation channels, the blind system identification strategy is applied to estimate the transmission performances of the two acoustic channels. Due to the long impulse response of acoustic channel, the overlap-save and cross-correlation fitting technique are utilized in blind system identification to estimate the acoustic channels under a built constraint condition. In order to avoid converging to the local minima, the genetic algorithm is adopted to optimize the channel identification. The times due to the propagation of the leak source signal traveling from the leak point to either sensor are extracted from the identified acoustic channels. Consequently, for leak location, the pipe length is no longer a prerequisite. The practical location results show the validness of the proposed scheme.

1. Introduction

Leakage from buried water pipes is a major issue facing all water distribution companies. Water loss can be reduced by a program of locating and repairing leaking pipes. A popular technique described for example by Fuchs and Riehle [1] developed over the last 20 years is to locate leaks by acoustic/vibration signal analysis. The vibration or acoustic signals are acquired using either accelerometers or hydrophones at two access points, on either side of the location of a suspected leak. The correlation technique, which is used to estimate the time delay between the two acquired signals, is common process used to determine the leak location of buried pipelines [1–4]. For the established techniques to be effective, the acoustic signal propagation velocity and the exact pipe length must be known a priori. The studies of acoustic wave propagation along metal or plastic pipes have received attention for a number of years. A dispersive propagation model in buried iron

*Corresponding author. Tel./fax: +86 23 65112797.
E-mail address: ymwen@cqu.edu.cn (Y. Wen).

0022-460X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2007.07.067

Nomenclature

D, $\hat{D}$	time delay and its estimate
f_s	sampling frequency
$h_1(n)$, $h_2(n)$	transmission performances of the paths
$H(z)$	z-transform of $h(n)$
l	distance between two sensors
l_1, l_2	relative distance between the leak and sensors 1 and 2
$\hat{l}_1$, $\hat{l}_2$	location results
L	length of the longest channel impulse response
$n_1(n)$, $n_2(n)$	background noises
r	pipe size in diameter
$s(n)$	original leak source
$\hat{t}$	estimated absolute time delay of single channel
v	propagation speed
$x_1(n)$, $x_2(n)$	acquired acoustic/vibration signals
$Z\{H(z)\}$	roots of $H(z)$
∂	magnitude or power attenuation factor
Δl	location error
Φ	null set

water pipes was investigated (see for example Refs. [5,6]), and physical interpretations of results have been given. A low-frequency theoretical model of a buried fluid-filled pipe to predict wave speed was developed by Muggleton et al. [7], and validated experimentally for the case of an in-vacuum pipe, and a pipe buried in soil or water [8,9]. According to these investigations, it can be seen that, if the pipe conditions (such as the pipe thickness, material properties and surrounding media) are known, the propagation velocity can be calculated using corresponding theoretical model. In practice, if some pipe conditions are unknown, the propagation velocity can be measured onsite using a known simulated leak, for example, by releasing water at a fire hydrant [10]. All these studies have shown some promise in the leak location.

In leak location practice, the pipe length is usually read from the pipe distribution maps or measured manually, typically by walking with a measurement wheel [3,4]. But the records about the pipes are often out of date, or incomplete and inaccurate with errors as high as 15–30% [11,12]. For example, pipeline maps are not always accurate enough because of the change in surroundings, e.g. expansions of streets and reconstruction of houses. In addition, the actual length of a buried pipeline, with unknown turns and changes in depth, may deviate from the distance walked with the wheel at ground level. An imprecise knowledge of the pipeline length between sensors adds an error to the leak location. In these cases, a pipe locator is often adopted to measure the buried pipe length [12,13]. Then the obtained length is used to locate the leak point by correlation technique. Because more information about the pipe is unavailable except the measured pipe length by a pipe locator, it is often time-consuming to confirm the leak point, and the leak location is incorrect as usual. In order to expand the applications of the conventional leak location methods, the approach was proposed for automatic measurement of the pipe length [14]. The length is measured by an impulse signal or simulated leak signal, not by a wheel or pipe locator. According to the calculated result of leak point in the pipe by correlation technique, the operators will find the leak position relative to the ground surface by combining the information provided by the length measurement and pipe locator. Because the method takes advantage of the automatic length measurement and a pipe locator to pinpoint the leak relative to the ground surface, the performance of the technique is better than that of the method only using a pipe locator. In practice, although the method has shown some promise, it is usually more complex and time-consuming for measuring the pipe length, and may impose additional restriction for practical detection. So a quicker, cheaper, and more comprehensive method is needed for leak location without knowledge of pipe length.

In fact, the acquired signals contain not only the time lag information between the singles, but also the characteristics related to the propagation paths, such as frequency alternative, energy decay and propagation time [1]. This way, by appropriate signal analysis techniques, it is possible to identify the absolute propagation time from the acoustic source to either monitoring point instead of the time lag between distinct channels. Therefore, a new leak location method exclusive of the prior knowledge of pipe length is proposed, which is based on the identification of the acoustic propagation channels.

Since pure leak sources are inaccessible, and only the signals blurred with noises at the sensing points are available, the conventional system identification methods cannot be applied [15]. Blind channel identification techniques have gained extensive attention since the innovative idea was first proposed by Sato [16]. Early studies [17–19] of blind channel identification focused primarily on higher order statistics (HOS)-based schemes. These major problems of the schemes are their slow convergence and local minima. In 1991, Tong et al. [20] demonstrated the possibility of using only second-order statistics (SOS) of multi-channel system outputs to solve the channel identification problem. SOS-based methods have potentially fast convergence. Therefore, the focus of blind channel identification shifted to SOS methods [21,22]. Since then, many SOS-based approaches have been proposed, such as the subspace (SS) algorithm [23], the least-squares algorithm [24,25], the linear prediction method [26], and the method robust to order overestimation [27].

The leak acoustic channels are supposed to be linear time-invariant FIR systems. It is remarkable that the leak acoustic channel system has an impulse response with quite a long sequence, especially while the transmission time along quite a long pipeline is desirable to be estimated according to the impulse response. The SS and least-squares techniques are robust when the channel order is low [22–25], but their performances degrade when the order is high. The normalized multi-channel frequency-domain LMS algorithm [28] and the normalized blind frequency-domain least mean square method [29] are possible to adaptively estimate the room acoustic impulse response with large length. However a block delay will be unavoidably introduced in a frequency-domain implementation. At the same time, in practice, due to the long impulse response of the two channels, the "almost common zero" situations may occur. In such cases, the channels would be ill-conditioned. The multi-channel Newton (MCN) algorithm [30] offers good performance when the channels are ill-conditioned. But the MCN method needs to invert a nondiagonal Hessian matrix, which involves extensive computation. The QR-based algorithms [31] are not good enough for the long impulse response. Because the overlap-save technique [32] is a block processing where the identifying channel coefficients are updated once per block of input data. It is accurate for the identification of long impulse response. The cross-correlation fitting technique [33] shows superior performance when the channel is close to unidentifiable. So in this paper the overlap-save and cross-correlation fitting technique are combined to identify the long leak acoustic channels. Further, a common assumption in the blind system identification is that the order of the channel is known. However, such information is not available in the leak detection practice. In this case, a constraint condition is built to confirm that there is an unique solution of the algorithm based on the overlap-save and correlation fitting techniques. At the same time, in order to avoid converging to the local minima, the genetic algorithm [34] is used to minimize the cost functions.

This paper is organized as follows. Section 2 reviews the traditional leak location scheme based on correlation technique. The details of the principle of leak location based on the blind system identification are given in Section 3. Computer simulation examples are presented in Section 4 and the practical leak locations without the knowledge of the detection length are also presented to be compared with those of the traditional scheme.

2. Leak location using time delay estimation

The vibration or acoustic signals are generated by friction and cavitations of the fluid leak. To perform leak detection, vibration or acoustic signals are acquired at two spatially separate monitoring points using sensors such as accelerometers or hydrophones, on either side of a suspected leak, as shown in Fig. 1.

The two acquired signals are mathematically modeled as

$$\begin{cases} x_1(n) = s(n) + n_1(n), \\ x_2(n) = \partial s(n - D) + n_2(n), \end{cases} \tag{1}$$

where $x_1(n)$ and $x_2(n)$ denote the sensor outputs, $s(n)$ is the original leak source and unknown, $n_1(n)$ and $n_2(n)$ are zero-mean random noises representing the ambient interferences and noises, ∂ is a magnitude or power attenuation factor caused by the acoustic path difference and D is the time delay. $s(n)$, $n_1(n)$ and $n_2(n)$ are supposed to be mutually uncorrelated. The leak position is located with respect to the delay of the leak noise

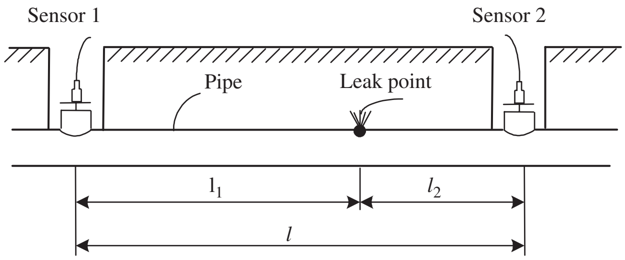

Fig. 1. A typical leak detection in a buried water pipe.

picked up by the two sensors. The leak location is determined by

$$\hat{l}_1 = \frac{l + v\hat{D}}{2}, \tag{2}$$

where v is the propagation speed of the acoustic vibration along the monitored pipeline, l is the length of pipeline between sensing points, $\hat{D}$ is the estimate of D by the correlation methods. In Eq. (2), in order to locate the leak, it can be seen that the pipeline length between sensing points is a necessity. This is to say that the accurate length of a pipeline has to be known. However, in practice, this prerequisite is not always satisfied. So a new location scheme without a priori knowledge of the pipe length is desirable.

3. Principle of leak location based on the blind system identification

Actually, the detection model expressed as Eq. (1) is a simplified model to detect leak signature. It supposes that the leak noises from an identical source but collected spatially separately are correlated and have not been modified by distinct acoustic paths. The more general and realistic model is

$$\begin{cases} x_1(n) = h_1(n) \otimes s(n) + n_1(n), \\ x_2(n) = h_2(n) \otimes s(n) + n_2(n), \end{cases} \tag{3}$$

where $h_i(n)$ $(i = 1{,}2)$ represents the transmission system of the leak acoustic channel from the source to the ith acquisition point along the pipeline. $h_i(n)$ $(i = 1{,}2)$ contains not only the time delay information of the two sensor outputs, but also the information of the time lag of the corresponding channel. In order to extract the propagation times from the acoustic source to either monitoring point, the blind system identification technique is adopted to estimate the leak acoustic channels according to the acquired leak signals.

3.1. Principle of the blind leak acoustic channel identification

3.1.1. Principle of blind system identification

The relationships between the input and the observations in a single-input two-output (SITO) system and the relative blind system identification scheme are depicted together in Fig. 2.

In a vector/matrix form, the relationship (3) becomes

$$\mathbf{x}_i(n) = \mathbf{s}(n) \cdot \mathbf{h}_i + \mathbf{n}_i(n), \quad i = 1, 2, \tag{4}$$

where

$$\mathbf{x}_i(n) = [\, x_i(n), \quad x_i(n-1), \quad \ldots, \quad x_i(n-L+1)\,]^{\mathrm{T}},$$

$$\mathbf{h}_i = [\, h_{i,0}, \quad h_{i,1}, \quad \ldots, \quad h_{i,L-1}\,]^{\mathrm{T}},$$

$$\mathbf{n}_i(n) = [\, n_i(n), \quad n_i(n-1), \quad \ldots, \quad n_i(n-L+1)\,]^{\mathrm{T}},$$

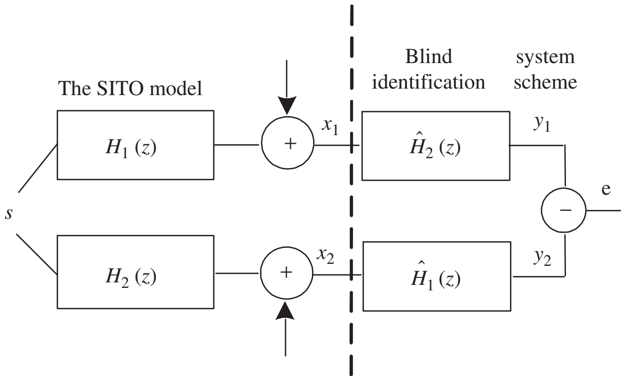

Fig. 2. Illustration of the single-input two-output model and blind system identification scheme.

$$\mathbf{s}(n) = \begin{bmatrix} s(n) & s(n-1) & \cdots & s(n-L+1) \\ s(n-1) & s(n-2) & \cdots & s(n-L) \\ \vdots & \vdots & \cdots & \vdots \\ s(n-L+1) & s(n-L) & \cdots & s(n-2L+2) \end{bmatrix}_{L\times L}$$

and $(\cdot)^{\mathrm{T}}$ denotes a vector/matrix transpose, L is set to the length of the longest channel impulse response by assumption.

In the absence of noise, the relationship between the input and the observations at time n is written as

$$\mathbf{x}_1^{\mathrm{T}}(n)\mathbf{h}_2 = [\mathbf{s}(n)\mathbf{h}_1]^{\mathrm{T}}\mathbf{h}_2. \tag{5}$$

Then

$$[\mathbf{s}(n)\mathbf{h}_1]^{\mathrm{T}}\mathbf{h}_2 = \mathbf{h}_1^{\mathrm{T}}[\mathbf{s}^{\mathrm{T}}(n)\mathbf{h}_2] = \mathbf{h}_1^{\mathrm{T}}\mathbf{x}_2(n). \tag{6}$$

In a matrix form, this set of equations is written as

$$\mathbf{X}(n)\mathbf{h} = 0, \tag{7}$$

where

$$\mathbf{X}(n) = [\,\mathbf{x}(n)_2^{\mathrm{T}} \quad -\mathbf{x}(n)_1^{\mathrm{T}}\,],$$

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}.$$

When observation noise is present, the right-hand side of Eq. (7) is no longer zero and an error is produced

$$e(n) = \mathbf{X}(n)\mathbf{h}. \tag{8}$$

This error can be used to define a cost function

$$J(n) = \frac{e^2(n)}{||\mathbf{h}(n)||^2}. \tag{9}$$

According to Huang and Benesty [30], a vector $\hat{\mathbf{h}}$ as the solution can be determined by minimizing the cost function (9) in the adaptive multi-channel LMS algorithm

$$\hat{\mathbf{h}} = \arg \min_{||\hat{\mathbf{h}}||=1} E[J(n)]. \tag{10}$$

In order to avoid a trivial estimate with all zero elements, the unit-norm constraint is proposed, i.e. $||\hat{\mathbf{h}}|| = 1$.

3.1.2. Cost function construction using overlap-save and cross-correlation fitting techniques

The adaptive multi-channel LMS algorithm based on Eq. (9) exploits the channel diversity and minimizes an error criterion in a novel and systematic way. The ability to adapt makes it possible to apply blind multi-channel identification techniques in many practical applications. However, the utilization of the blind algorithm to leak acoustic channels poses a practical problem. In general, the leak acoustic channels have significantly different characteristics from RF channels due to the speed of sound and sampling rates. This introduces an estimation problem for realistic channel lengths of hundreds of samples. The large channel order prohibits the implementation of the blind algorithm. So the overlap-save technique is used to establish a different cost function to estimate the long length channels.

According to the overlap-save technique, the relationship between the input series and output is defined as

$$\hat{\mathbf{Y}}_{ij}(n) = \hat{\mathbf{X}}_i(n)\hat{\mathbf{h}}_j(n), \quad i,j = 1,2, \quad i \neq j, \tag{11}$$

where

$$\hat{\mathbf{h}}_j(n) = [\,\hat{h}_{j,0}(n), \quad \hat{h}_{j,1}(n), \quad \ldots, \quad \hat{h}_{j,L-1}(n)\,]^{\mathrm{T}},$$

$$\hat{\mathbf{Y}}_{ij}(n) = [\hat{y}(nL-L), \hat{y}(nL-L+1), \ldots, \hat{y}(nL-L+N)]^{\mathrm{T}},$$

$$\hat{\mathbf{X}}_i(n) = \begin{bmatrix} x_i(nL-L) & x_i(nL-L-1) & \cdots & x_i(nL-2L+1) \\ x_i(nL-L+1) & x_i(nL-L) & \cdots & x_i(nL-2L+2) \\ \vdots & \vdots & \vdots & \vdots \\ x_i(nL-L+N) & x_i(nL-L+N-1) & \cdots & x_i(nL-2L+N+1) \end{bmatrix}.$$

Then

$$\mathbf{e}(n) = \hat{\mathbf{Y}}_{12}(n) - \hat{\mathbf{Y}}_{21}(n), \tag{12}$$

where, $\mathbf{e}(n) = [\,e_1(n), \quad e_2(n), \quad \ldots, \quad e_{N+1}(n)\,]$

And the corresponding cost function is given by

$$J_1(n) = \frac{\sum_{i=1}^{N+1} e_i^2(n)}{||\hat{\mathbf{h}}(n)||^2}. \tag{13}$$

Therefore, the desired solution for $\mathbf{h}$ is determined by minimizing the mean value of the cost function $J_1(n)$:

$$\hat{\mathbf{h}} = \arg \min_{||\hat{\mathbf{h}}||=1} E[J_1(n)]. \tag{14}$$

At the same time, because of the long impulse responses of the channels, it is unavoidable that some zeros of the two acoustic channels may be very close. In this case, the performance of the algorithm based on the cost function (14) will degrade. So another cost function based on the cross-correlation fitting principle is defined as follows:

The estimated correlation of the channel output is given by

$$r_{ij}(\tau) = E[x_i(n)x_j(n+\tau)], \quad i,j = 1,2, \quad i \neq j, \quad \tau \geqslant 0. \tag{15}$$

By the relationship between the correlation and the convolution, Eq. (15) can be simplified as follows:

$$r_{ij}(\tau) = r_{ss}(\tau) \otimes r_{h_i h_j}(-\tau), \tag{16}$$

where

$$r_{ss}(\tau) = \sum_{n=-\infty}^{\infty} s(n)s(n+\tau),$$

$$r_{h_i h_j}(\tau) = \sum_{n=1}^{L'} h_i(n) h_j(n+\tau), \quad L' > L.$$

Thus the correlations in the vector form is written as

$$\mathbf{r}_{ij} = \left[r_{ij}(0), \quad r_{ij}(1), \quad \ldots, \quad r_{ij}(L'-1) \right]. \tag{17}$$

When $\tau > L_2$, $r_{h_1 h_2}(\tau) = 0$ then

$$\mathbf{r}_{12} = \left[r_{12}(0), r_{12}(1), \ldots, r_{12}(L_2 - 1), \underbrace{0, \ldots, 0}_{L'-L_2} \right]. \tag{18}$$

When $\tau > L_1$, $r_{h_2 h_1}(\tau) = 0$ then

$$\mathbf{r}_{21} = \left[r_{21}(0), r_{21}(1), \ldots, r_{21}(L_1 - 1), \underbrace{0, \ldots, 0}_{L'-L_1} \right]. \tag{19}$$

To combine the two vectors, define the vector of correlation function as

$$\mathbf{r} = [\, \mathbf{r}_{12} \quad \mathbf{r}_{21} \,]. \tag{20}$$

From Eq. (20), it can be seen that the estimated correlation $\mathbf{r}$ contains all the second-order statistical information of the receive signals. Similar to the definition of $\mathbf{r}_{ij}$, the vectors of $\mathbf{r}_{\hat{h}_{ij}}$ about the estimated impulse response $\hat{h}_i(n)$ are defined as

$$\mathbf{r}_{\hat{h}_{ij}(n)} = \left[r_{\hat{h}_{ij}(n)}(0), \quad r_{\hat{h}_{ij}(n)}(1), \quad \ldots, \quad r_{\hat{h}_{ij}(n)}(L'-1) \right], \tag{21}$$

where $r_{\hat{h}_{ij}(n)}(\tau) = E\{\hat{h}_i(n)\hat{h}_j(n+\tau)\}$.

Then

$$\mathbf{r}_{\hat{h}(n)} = \left[\mathbf{r}_{\hat{h}_{12}(n)} \quad \mathbf{r}_{\hat{h}_{21}(n)} \right]. \tag{22}$$

The width of the first lobe of auto-correlation function of $s(n)$ is considered to be narrower than that of the correlation function of $h_1(n)$ and $h_2(n)$. According to Eqs. (20) and (22), an error vector is obtained:

$$\mathbf{E}(n) = \frac{\mathbf{r}}{||\mathbf{r}||} - \frac{\mathbf{r}_{\hat{h}(n)}}{\left\| \mathbf{r}_{\hat{h}(n)} \right\|}, \tag{23}$$

where $||\cdot||$ stands for a matrix norm.

The corresponding cost function is defined as

$$J_2(n) = \mathbf{E}(n)\mathbf{E}^{\mathrm{T}}(n). \tag{24}$$

Hence, the desired solution for $\mathbf{h}$ is given by

$$\hat{\mathbf{h}} = \arg \min_{||\hat{\mathbf{h}}||=1} E[J_2(n)]. \tag{25}$$

3.1.3. Constraint condition

In the above, the maximum order of multi-channels L is considered to be known. In practice, it is not necessarily true. Now, the case will be discussed that whether the channels can be uniquely identified from Eqs. (14) and (25) when the channel order L is perfectly estimated, underestimated or overestimated.

By the principle of blind system identification, for the cost function $J_1(n)$, it is not difficult to see

$$H_1(z)\hat{H}_2(z) = H_2(z)\hat{H}_1(z). \tag{26}$$

Let us assume that $H_1(z)$ has the maximum order L. Then

$$Z\{H_1(z)\} \in Z\{H_2(z)\} \cup Z\{\hat{H}_1(z)\}, \tag{27}$$

where $Z\{H(z)\}$ denotes the roots of $H(z)$. Since $H_i(z)$ does not share some common roots, by Eq. (27),

$$Z\{H_1(z)\} \in Z\{\hat{H}_1(z)\} \tag{28}$$

(1) When the maximum order of $\hat{H}_1(z)$ is L, $Z\{H_1(z)\} = Z\{\hat{H}_1(z)\}$, and therefore, $\hat{H}_1(z) = gH_1(z)$, where g is any nonzero constant. This means that the solution is unique up to a scalar multiple.
(2) If the maximum order of $\hat{H}_1(z)$ is underestimated, then Eq. (26) will only have zero solution.
(3) If the order is overestimated, then $g = g(z)$, that is to say g is not a nonzero constant. Eq. (26) will have more than one solution.

By the principle of correlation fitting technique, for the cost function $J_2(n)$, it is not difficult to see

$$H_1(z)H_2(z^{-1}) = \hat{H}_1(z)\hat{H}_2(z^{-1}). \tag{29}$$

(1) When the maximum order of $\hat{H}_1(z)$ is L, the solution is unique up to a scalar multiple.
(2) If the maximum order of $\hat{H}_1(z)$ is underestimated, then Eq. (29) will only have zero solution.
(3) If the order is overestimated, Two functions $G_1(z)$ and $G_2(z)$ are defined as

$$\hat{H}_1(z) = G_1(z)H_1(z), \tag{30}$$

$$\hat{H}_2(z) = G_2(z)H_2(z). \tag{31}$$

Then

$$\hat{H}_1(z)\hat{H}_2(z^{-1}) = G_1(z)H_1(z)G_2(z^{-1})H_2(z^{-1}). \tag{32}$$

When $G_1(z) = G_2(z)$the $\hat{H}_i(z)$ in Eqs. (30) and (31) can satisfy Eq. (29). It can be concluded that if the order is overestimated, Eq. (29) have more than one solution.

In order to identify the channel impulse responses correctly, Eqs. (14) and (25) are modified to include the order and parameters

$$\{\hat{\mathbf{h}}, L\} = \arg \min_{||\hat{\mathbf{h}}||=1} E[J_1(n)], \tag{33}$$

$$\{\hat{\mathbf{h}}, L\} = \arg \min_{||\hat{\mathbf{h}}||=1} E[J_2(n)]. \tag{34}$$

At the same time, as discussed in the foregoing section, when the order is overestimated, there is more than one vector $\hat{\mathbf{h}}$ as the solution that can minimize the cost functions. That is to say, in this case, Eqs. (33) and (34) will also have more than one solution. Then a constraint condition is defined as

$$Z\{\hat{H}_1(z)\} \cap Z\{\hat{H}_2(z)\} = \Phi, \tag{35}$$

where Φ denotes null set.

When the order is overestimated, Eqs. (30) and (31) as the solution share common zeros due to $G_1(z) = G_2(z)$. Then the constraint condition is adopted to confirm that the situation ($G_1(z) = G_2(z)$) cannot be satisfied. Thus, there is no vector $\hat{\mathbf{h}}$ as the solution that can minimize the cost functions when the order is overestimated. That is to say, only the correct order and impulse responses parameters can minimize the cost functions.

There are two cost functions that include unknown order and parameters and a constraint condition. The conventional gradient optimization techniques may converge to “wrong” solutions unless a good initial value. To overcome the problem of local minima, here the genetic algorithm is adapted to optimize the solution.

3.2. The GA procedure

In this paper, the GA proposed in Ref. [34] is modified for minimizing the cost functions formulated in Section 3.1. Since there are two cost functions and a constraint condition, the optimal channels can be obtained satisfying the following expansion:

$$\begin{cases} \min f(\hat{\mathbf{h}}, L) = \left[f_1(\hat{\mathbf{h}}, L) \quad f_2(\hat{\mathbf{h}}, L) \right] \\ \qquad\qquad \text{s.t.} \quad g(\hat{\mathbf{h}}), \end{cases} \tag{36}$$

where $f(\hat{\mathbf{h}}, L)$ denotes the multi-object function, $f_1(\hat{\mathbf{h}}, L)$ denotes the function (33), $f_2(\hat{\mathbf{h}}, L)$ denotes the function (34), $g(\hat{\mathbf{h}})$ denotes the constraint condition.

In Section 3.1, it can be seen that the two cost functions have the same optimal solution under the constraint condition. So a resultant objective function is constructed, which is a weighted combination of the two objectives through a weighted approach. The resultant objective function is given by

$$\xi(\hat{\mathbf{h}}, L) = \sum_{i=1}^{2} w_i f_i(\hat{\mathbf{h}}, L). \tag{37}$$

As a result, the value of the fitness function can be defined as

$$\text{fit}(\xi(\hat{\mathbf{h}}, L)) = -\xi(\hat{\mathbf{h}}, L). \tag{38}$$

3.3. Leak location based on the time lag of the signal channel

For a linear phase FIR filter, the tap of the maximum coefficient of the filter is the time lag of the filter system. In practice, the leak acoustic channel system will not have an absolutely linear phase as usual. However, the available spectrum is mainly considered, which has the relation with the maximum value of the filter coefficients. Thus it is reasonable to use the tap of the maximum coefficient indicating the time delay of the channel. Thus the leak location based on the time lag of the channel is obtained as follows:

$$\hat{l}_1 = \frac{\hat{D}_1(\text{Tap}) \cdot v}{f_s}, \tag{39}$$

where $\hat{D}_1$ is the tap of the maximum value of $\hat{h}_1(n)$.

4. Applications of the proposed scheme

4.1. Application to known sources and channels

To verify the validity and performance of the algorithm, the source (a white noise) and the channels are given as known. All the tests are concerned with two-channel FIR system. The white noises are added to the channel outputs, and the SNR is 8 dB.

Example 1. The algorithm for low order and well-conditional channels.

The true channel coefficients are presented in Fig. 3. The order of each channel in Fig. 3 is 8. In order to compare with the estimated channels, the tap number in Fig. 3 is the same as that in Fig. 4.

The initial value for the channel order is set $L' = L + 8 = 16$ that is larger than the true channel order $L = 8$. After genetic generations, the channel order (order = 8) is found correctly. Fig. 4 plots the estimated results. And the tap number in Fig. 4 is the same as the estimated channel order $L' = 16$. The coefficients of the two estimated filters before the eighth tap are very close to the true values. The rest are not real filter coefficients, so they converge to small values by the scheme. It can be seen that the proposed algorithm is able to estimate the channel response when the channel order is unavailable.

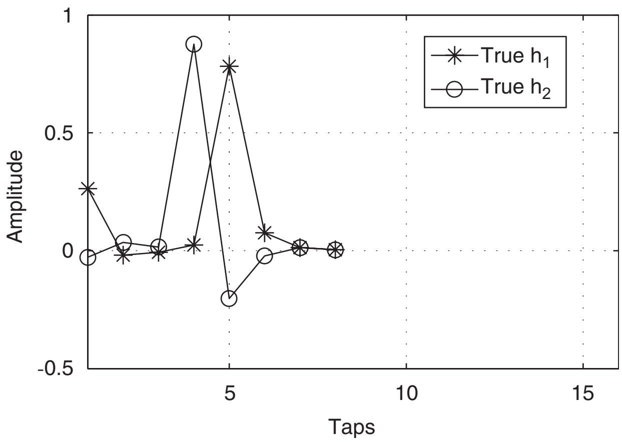

Fig. 3. Low order and well-conditional channels.

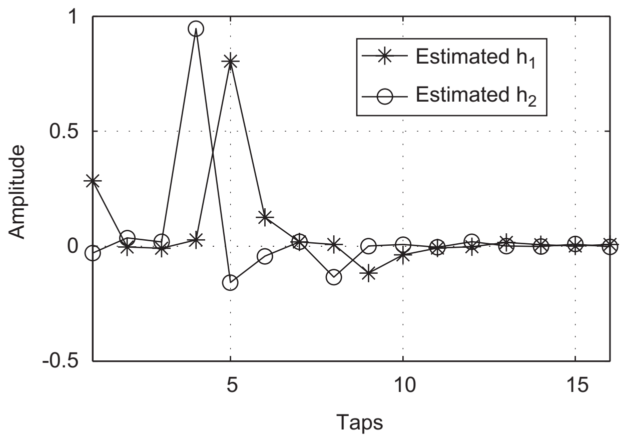

Fig. 4. Identification results.

Example 2. The algorithm for high order and ill-conditional channels.

In the second experiment, the channel order is set to be 50. In order to compare with the estimated channels, the tap number in Fig. 5 is the same as that in Fig. 7. The two channels are presented in Fig. 5. The zeros of the two impulse responses are shown in Fig. 6. It is obvious that some zeros of the two channels are very close.

The initial value for the channel order is set $L' = L + 29 = 79$ that is larger than the true channel order $L = 50$. Fig. 7 shows the estimated results by the cost function (13). It can be seen that the identification performance degrades drastically when the two channels are ill-conditioned. Fig. 8 shows the results by the proposed scheme. The taps of the maximum values of the coefficients are the same to the true. Identification results show that the proposed method works effectively for both high order and ill-conditioned channels.

4.2. Application in practical water pipeline leak location

The technique presented in this paper has been used to locate leak in practical water distribution systems. The two accelerometers are put on either side of the suspected leak, as shown in Fig. 1.

To check the validity of the proposed scheme, the leak position in a buried pipe is located by the proposed blind identification approach. The detection pipe is a 200 mm diameter cast iron pipe, which is buried in soil at a depth of 1.5 m. In blind system leak location process, the initial maximum order L' of the two channels is

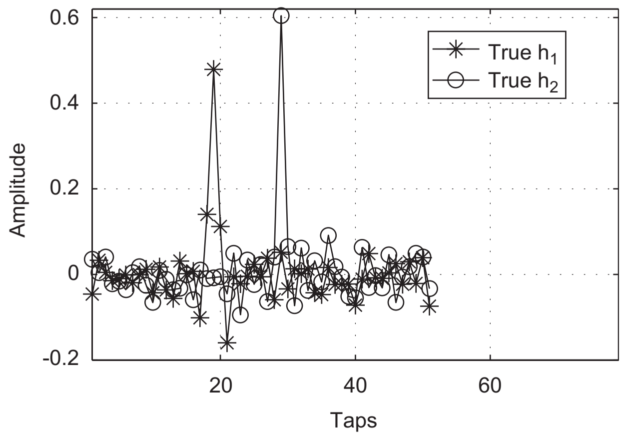

Fig. 5. High order and ill-conditional channels.

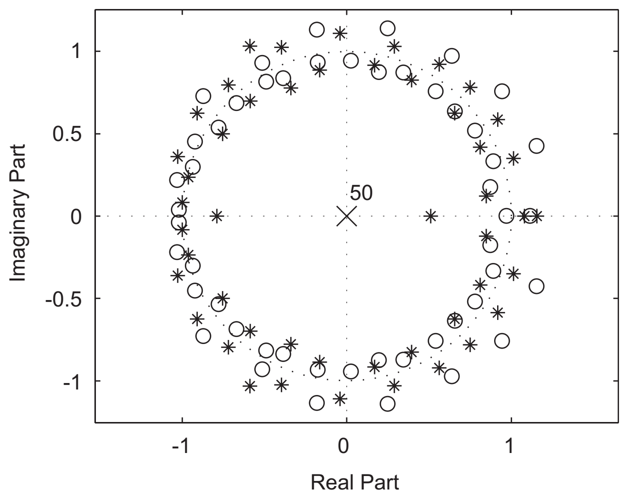

Fig. 6. Zero distribution of the two channels.

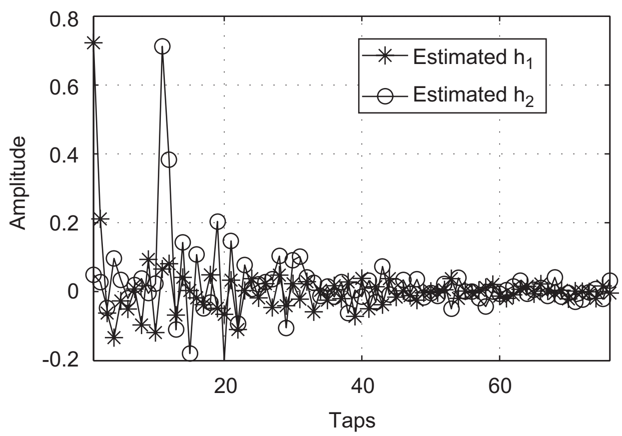

Fig. 7. Identification results by the cost function (13).

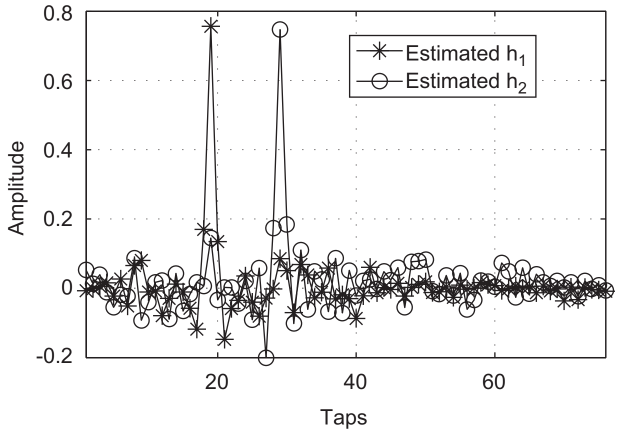

Fig. 8. Identification results by the scheme presented in this paper.

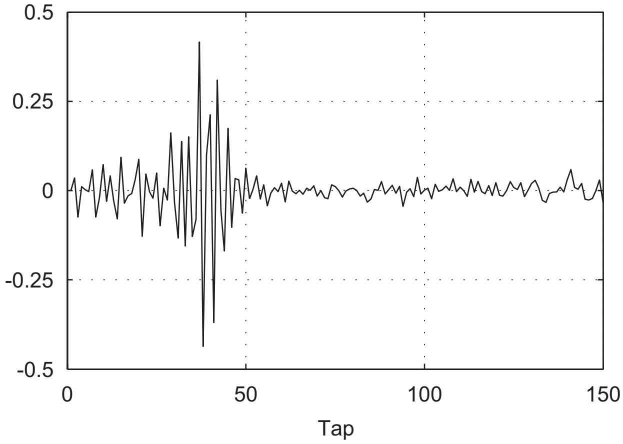

Fig. 9. Leak acoustic channel identification result for the short sensing distance.

chosen such that $L'v/f_s > l$, where $f_s = 4.861$ ks/s is the sampling frequency; $v = 1.17$ m/ms. Figs. 9 and 10 are the leak acoustic channel identification results of the pipe with one leak. The associated frequency responses of the estimated channels are plotted in Figs. 11 and 12.

The tap of the maximum value of the filter coefficients in Fig. 9 is 36, and the tap of the maximum value in Fig. 10 is 103. The corresponding distances of the leak locations are calculated by the form (39) as follows: $\hat{l}_1 = 36 \times 1170/4861 = 8.66$ m, $\hat{l}_2 = 103 \times 1170/4861 = 24.8$ m. Compared with the actual values $l_1 = 9.5$ m and $l_2 = 25$ m, it is obvious that the location results are very close to the true. From Figs. 11 and 12, it can be seen that the longer the leak acoustic channel, the lower the predominant frequencies. The conclusion is the same as in Refs. [1–4]. That means the frequency feature of the channel can also be identified correctly.

To compare the performance of the proposed approach with the conventional method [1], lots of leak signals acquired in confirmed leak pipelines are processed. Table 1 is the location results with the traditional correction method. In this case, the length of pipes is measured with the wheel rolling on the ground. Table 2 gives the location results with the blind system identification leak location, which does not need the knowledge of the pipe length. In Tables 1 and 2, compared with the actual value l_1, it can be seen that the leak location results in Table 2 are more accurate. The difference between the rolling direction of the measurement wheel at the ground level and the actual direction of the pipe under the ground adds error to the conventional leak location.

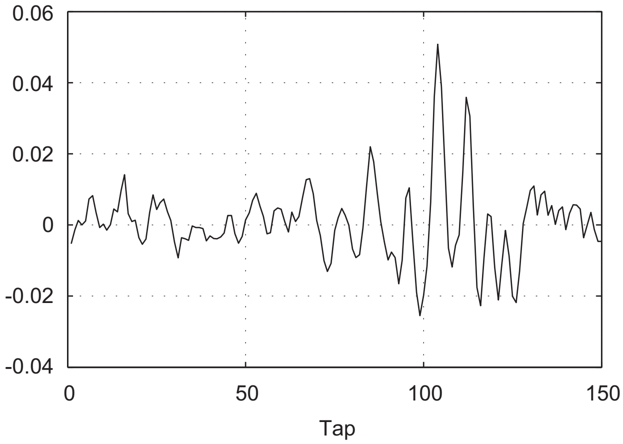

Fig. 10. Leak acoustic channel identification result for the long sensing distance.

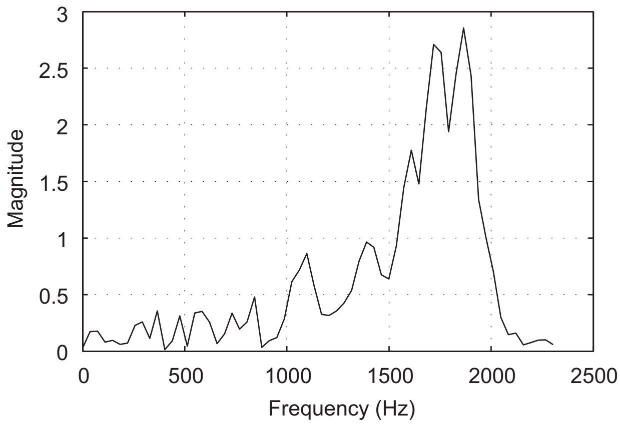

Fig. 11. Frequency response of the estimated channel in Fig. 9.

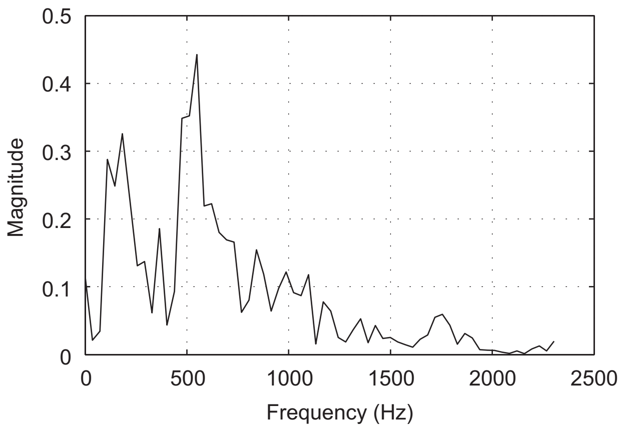

Fig. 12. Frequency response of the estimated channel in Fig. 10.

As can be seen from Table 2, in comparison with the conventional leak location methods based on correlation, the leak can be located using the blind system identification without the knowledge of the detection length of pipe. The average leak location errors are within 1 m.

Table 1
The leak location results of the traditional method

r (mm)	v (m/ms)	l_1 (m)	$\hat{l}_1$ (m)	Δl (m)
100	1.25	36.0	37.9	1.9
200	1.17	24.3	25.6	1.3
200	1.17	25.7	24.0	−1.7
200	1.17	34.3	31.2	−3.1
200	1.17	39.0	36.9	−2.1
200	1.17	53.5	55.5	2.0
200	1.17	67.1	69.6	2.5

Table 2
The leak location results without the knowledge of the detection length

r (mm)	v (m/ms)	$\hat{t}$ (ms)	l_1 (m)	$\hat{l}_1$ (m)	Δl (m)
100	1.25	29.1	36.0	36.4	0.4
200	1.17	20.5	24.3	24.0	−0.3
200	1.17	21.5	25.7	25.2	−0.5
200	1.17	30.0	34.3	35.1	0.8
200	1.17	32.3	39.0	37.8	−1.2
200	1.17	46.5	53.5	54.5	1.0
200	1.17	58.6	67.1	68.6	1.5

r: pipe size in diameter; $\hat{t}$: estimated absolute time delay of single channel; l_1: actual length from the leak to the sensor 1; $\hat{l}_1$: location result; Δl: location error.

5. Conclusion

In leak location practice, the information about the pipe length is often inaccurate, incomplete or out of date. In these cases, the conventional location method based on time delay estimation is no longer feasible. In this paper, the blind identification strategy is applied to locate leak point without knowledge of the length of pipeline in the water distribution system. The main conclusions of the investigation can be summarized as follows:

- In blind system identification, the overlap-save and correlation fitting techniques are adopted to build two cost functions to identify the long and ill-conditional leak acoustic channels under a built constraint condition.
- In order to avoid converging to the local minima, a resultant objective function is constructed, which is a weighted combination of the two objectives. Then the genetic algorithm is used to minimize the resultant objective function.
- The practical leak location results in water distribution system show that the transmission performances of the two acoustic paths can be identified by the scheme, and the absolute time from the leak source to either monitoring sensor is determined. Thus, for leak location, compared with the conventional leak location methods, the prior knowledge of pipe length is no longer a prerequisite.

References

[1] H.V. Fuchs, R. Riehle, Ten years of experience with leak detection by acoustic signal analysis, *Applied Acoustics* 33 (1991) 1–19.
[2] D.A. Liston, J.D. Liston, Leak detection techniques, *Journal of the New England Water Works Association* 106 (1992) 103–108.

[3] Y. Gao, M.J. Brennan, P.F. Joseph, J.M. Muggleton, O. Hunaidi, A model of the correlation function of leak noise in buried plastic pipes, *Journal of Sound and Vibration* 277 (2004) 133–148.

[4] Y. Gao, M.J. Brennan, P.F. Joseph, A comparison of time delay estimators for the detection of leak noise signals in plastic water distribution pipes, *Journal of Sound and Vibration* 292 (2006) 552–570.

[5] R. Long, K. Vine, M.J.S. Lowe, P. Cawley, Monitoring acoustic wave propagation in buried cast iron water pipes, *Review of Progress in Quantitative Nondestructive Evaluation* 20 (2001) 1202–1209.

[6] R. Long, P. Cawley, M.J.S. Lowe, Acoustic wave propagation in buried iron water pipes, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 459 (2003) 2749–2770.

[7] J.M. Muggleton, M.J. Brennan, R.J. Pinnington, Wavenumber prediction of waves in buried pipes for water leak detection, *Journal of Sound and Vibration* 249 (5) (2002) 939–954.

[8] J.M. Muggleton, M.J. Brennan, P.W. Linford, Axisymmetric wave propagation in fluid-filled pipes: wavenumber measurements in in vacuo and buried pipes, *Journal of Sound and Vibration* 270 (1–2) (2004) 171–190.

[9] J.M. Muggleton, M.J. Brennan, Leak noise propagation and attenuation in submerged plastic water pipes, *Journal of Sound and Vibration* 278 (3) (2004) 527–537.

[10] O. Hunaidi, A. Wang, M. Bracken, T. Gambino, C. Fricke, Acoustic methods for locating leaks in municipal water pipe networks, *International Conference on Water Demand Management*, Dead Sea, 2004, pp.1–14.

[11] R.E. Stevens, J.H. Anspach, New technology overcomes the problems of under ground system interferences on power projects, *Proceedings of the American Power Conference*, Vol. 55, 1993, pp. 323–326.

[12] M. Farley, R. Liemberger, Developing a non-revenue water reduction strategy: planning and implementing the strategy, *Water Science and Technology: Water Supply* 5 (1) (2005) 41–50.

[13] K. Takahashi, M. Sato, Estimation of a buried pipe location by borehole radar, *IEEE International Proceedings, Geoscience and Remote Sensing Symposium* 1 (2005) 304–307.

[14] P. Lander, W.E. Saltzstein, Method for detecting leaks in pipelines, US Patent No. 5974862, 1999.

[15] A. Krzyzak, Nonparametric estimation and classification using radial basis function nets and empirical risk minimization, *IEEE Transactions on Neural Networks* 7 (2) (1996) 475–487.

[16] Y. Sato, A method of self-recovering equalization for multilevel amplitude-modulation, *IEEE Transactions on Communication* 23 (6) (1975) 679–682.

[17] D.N. Godard, Self-recovering equalization and carrier tracking in two-dimensional data communication systems, *IEEE Transactions on Communication* 28 (1) (1980) 1867–1875.

[18] G.B. Giannakis, J. Mendel, Identification of non-minimum phase systems using higher-order statistics, *IEEE Transactions on Acoustic, Speech and Signal Processing* 37 (1989) 7360–7377.

[19] O. Shalvi, E. Weinstein, New criteria for blind deconvolution of nonminimum phase system (channels), *IEEE Transactions on Information Theory* 36 (3) (1990) 312–320.

[20] L. Tong, G. Xu, T. Kailath, A new approach to blind identification and equalization of multipath channels, *Asilomar Conference on Signal, System, and Computers*, Vol. 2, 1991, pp. 856–860.

[21] L. Tong, G. Xu, T. Kailath, Blind identification and equalization based on second-order statistics: a time domain approach, *IEEE Transactions on Information Theory* 40 (2) (1994) 340–349.

[22] H.H. Zeng, L. Tong, Blind channel estimation using the second-order statistics: algorithms, *IEEE Transactions on Signal Processing* 45 (8) (1997) 1919–1930.

[23] E. Moulines, P. Duhamel, J.F. Cardoso, S. Mayrarque, Subspace methods for blind identification of multichannel FIR filters, *IEEE Transactions on Signal Processing* 43 (2) (1995) 516–525.

[24] G. Xu, H. Liu, L. Tong, T. Kailath, A least-squares approach to blind channel equalization, *IEEE Transactions on Signal Processing* 43 (12) (1995) 2982–2993.

[25] L. Tong, Q. Zhao, Joint order detection and blind channel estimation by least squares smoothing, *IEEE Transactions on Signal Processing* 47 (9) (1999) 2345–2355.

[26] A.P. Liavas, P.A. Regalia, J.P. Delmas, On the robustness of the linear prediction method for blind channel identification with respect to effective undermodeling/overmodeling, *IEEE Transactions on Signal Processing* 48 (2001) 1477–1481.

[27] H. Gazzah, P.A. Regalia, J.P. Delmas, K. Abed-Meraim, A blind multichannel identification algorithm robust to order overestimation, *IEEE Transactions on Signal Processing* 50 (6) (2002) 1449–1458.

[28] Y. Huang, J. Benesty, A class of frequency-domain adaptive approaches to blind multichannel identification, *IEEE Transactions on Signal Processing* 51 (1) (2003) 11–24.

[29] Z.L. Yu, M.H. Er, A robust adaptive blind multichannel identification algorithm for acoustic applications, *IEEE International Conference on Acoustic, Speech, and Signal Processing*, Vol. 2, 2004, pp. 25–28.

[30] Y. Huang, J. Benesty, Adaptive multi-channel least mean square and Newton algorithms for blind channel identification, *Signal Processing* 82 (2002) 1127–1138.

[31] X.H. Li, H. Fan, QR factorization based blind channel identification and equalization with second-order statistics, *IEEE Transactions on Signal Processing* 48 (1) (2000) 60–69.

[32] A.V. Oppenheim, R.W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.

[33] G.B. Giannakis, S.D. Halford, Performance analysis of blind equalizers based on cyclostationary statistics, *IEEE International Conference on Information Sciences and Systems*, Princeton, NJ, 1994, pp. 711–716.

[34] K.F. Man, K.S. Tang, S. Kwong, Genetic algorithm: concepts and applications, *IEEE Transactions on Industrial Electronics* 43 (5) (1996) 519–534.