Application of Blind Source Separation Algorithms and Ambient Vibration Testing to the Health Monitoring of Concrete Dams

Lin Cheng and Fei Tong

State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi’an University of Technology, Xi’an 710048, China

Correspondence should be addressed to Lin Cheng; wanmu2381@163.com

Received 13 August 2016; Revised 5 October 2016; Accepted 20 November 2016

Academic Editor: Yuri Vladimirovich Mikhlin

Copyright © 2016 L. Cheng and F. Tong. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this work, using AVT data, a health monitoring method for concrete dams based on two different blind source separation (BSS) methods, that is, second-order blind identification (SOBI) and independent component analysis (ICA), is proposed. A modal identification procedure, which integrates the SOBI algorithm and modal contribution, is first adopted to extract structural modal features using AVT data. The method to calculate the modal contribution index for SOBI-based modal identification methods is studied, and the calculated modal contribution index is used to determine the system order. The selected modes are then used to calculate modal features and are analysed using ICA to extract some independent components (ICs). The square prediction error (SPE) index and its control limits are then calculated to construct a control chart for the structural dynamic features. For new AVT data of a dam in an unknown health state, the newly calculated SPE is compared with the control limits to judge whether the dam is normal. With the simulated AVT data of the numerical model for a concrete gravity dam and the measured AVT data of a practical engineering project, the performance of the dam health monitoring method proposed in this paper is validated.

1. Introduction

Dam health monitoring [1] is a process based on the structural static/dynamic response and environmental variables measured by sensors installed on a dam body, foundation, and reservoir to evaluate structural health. It is an effective way to prevent the occurrence of catastrophic dam failure that would result in great loss of human life and property. Dam health monitoring is a special case of structural health monitoring (SHM) [2]. Among all SHM methods that have been proposed, the vibration-based SHM method [3] has obvious advantages. In conventional static dam health monitoring methods, the commonly measured dam responses are displacement, seepage flow, temperature, and so on, which can indicate only the local damage of a structure where the instruments are installed, and it is usually very difficult to evaluate the global state of a structure. Unlike static dam response quantities, many structural dynamic features are global, for example, changes in the natural frequencies of the structure. The advantage of global features is that damage can be detected remotely from a sensor [4]. Other dynamic features, such as modal shape vectors, can also provide information about the damage location. Thus, using the vibration-based method to monitor the health of large concrete dams is feasible and meaningful. Moreover, with the construction of some vibration monitoring systems, such as the dam strong earthquake monitoring system, the ambient vibration testing (AVT) [5–8] data of concrete dams under certain ambient excitation can be acquired to diagnose structural health conveniently. Therefore, research on dam health monitoring methods for concrete dams based on ambient AVT at full scale has received much attention in the last few years.

To diagnose dam health using AVT data, two key steps, that is, feature extraction and feature classification, are necessary. Feature extraction is the process of identifying damage-sensitive properties, derived from the measured dynamic response, which allows one to distinguish between undamaged and damaged structures. The area of SHM that receives the most attention in the technical literature is feature extraction [9]. Although many different types of features have been reported to be used as structural damage indictors, modal features, for example, natural frequencies,
modal shape vectors, and coordinate modal assurance criteria (COMAC) [10], are still the most commonly used features in SHM. The advantage of using modal features is that they are independent of external ambient excitation and reflect the structural state directly. Using AVT data to identify the modal parameters of concrete dams is the operation modal analysis (OMA) [11] problem of structures. Recently, one area of research focused on OMA methods is the use of the blind source separation (BSS) technique to implement output-only modal identification. Among all proposed BSS-based OMA methods, the SOBI-based method, which has been studied in our previous work [12], shows many advantages, such as high identification accuracy, high computation efficiency, and strong robustness. The main difficulty in applying system identification for OMA of large-scale structures is the selection of the model order and the corresponding system poles. The most commonly used method is the stabilization diagram method. A conventional stabilization diagram using identified natural frequencies, damping ratios, and modal assurance criterion (MAC) is used as an index to select stable modes. Recently, a new index, modal contribution [13], has been proposed and shows better performance in discriminating structural models from mathematical poles. Whereas Cara’s work [13] shows us how to calculate the modal contribution index for the stochastic subspace identification (SSI) method [14], the procedure to calculate the modal contribution index for the SOBI-based method has not been studied.

Concrete dams are large-scale engineering structures with complicated operation environments and are characterized by the structure-water interaction effect. Environmental variables, such as temperature, water level, and rainfall, cannot be ignored directly in practical engineering as in numerical and experimental investigations. Therefore, for dam health monitoring based on AVT, how to remove the variability due to environmental changes is an important problem. To monitor the health of a structure under a varying environment, a commonly used method that seeks to remove the variability due to environment without measuring the environmental variables is the latent variable-based method. For example, Sohn et al. [15], Yan et al. [16], Deraemaeker et al. [17], and Ni et al. [18] have used the principle component analysis (PCA) method for bridge structure health monitoring. In our previous work [19], the kernel component analysis (KPCA) method is adopted to monitor the health state of concrete dams. Recently, some literature has reported on the good performance of independent component analysis (ICA) [20–22], which is a typical BSS algorithm and can be deemed an extension of PCA, in state monitoring of complex systems. PCA can impose independence only up to second-order statistical information (mean and variance), whereas ICA involves higher-order statistics; that is, it not only decorrelates the data (second-order statistics) but also reduces higher-order statistical dependencies. Therefore, an ICA-based SHM method can give more sophisticated results than a PCA-based one because ICA can extract the essential independent components that drive a process.

In this work, we first present a health monitoring method for concrete dams based on AVT data. Using the measured AVT data, the SOBI-based modal identification method is adopted to extract the basic modal parameters of a dam. For the SOBI-based modal identification method, the method to calculate the modal contribution index is proposed for the first time. The modal contribution index is then introduced into the stabilization diagram method to remove spurious modes of ambient excitation. The identified or calculated modal features are then analyzed using ICA, another BSS method, to extract some ICs, which may represent the effect of environmental variables. The square prediction error (SPE) control charts and the contribution plots are then used to detect structural abnormalities and the locations of such abnormalities. At the end of this work, a numerical example and an engineering example are used to verify the proposed dam health monitoring method.

2. Modal Identification of Concrete Dams Based on SOBI and Modal Contribution

2.1. SOBI-Based Modal Identification. Based on McNeill’s derivation [23], an analytic expression of the structural acceleration response can be constructed using Hilbert transformation (HT). After considering observation noise \( v(t) \), the expression of an analytic signal is as follows:

\[
y(t) = \begin{bmatrix} y^R(t) \\ y^I(t) \end{bmatrix} = \begin{bmatrix} C\Phi^R & C\Phi^I \\ C\Phi^I & C\Phi^R \end{bmatrix} \begin{bmatrix} q^R(t) \\ q^I(t) \end{bmatrix} + v(t)
\]

(1)

where \( y^R(t) \) is the physical acceleration of \( l \) channels and \( y^I(t) \) is its HT. The HT is performed to identify imaginary parts of modal shape vectors. \( C \) is the acceleration response selection matrix; \( \Phi^R \) and \( \Phi^I \) are the real and imaginary parts of complex mode shapes, respectively. \( q^R(t) \) and \( q^I(t) \) are the real and imaginary parts of the modal response.

If all structural modal responses are independent of each other, (1) shows a BSS problem indeed. The structural vibration response under ambient excitation can be regarded as a linear mixture of each modal response. Each modal response can then be regarded as a virtual source signal [24], and the mixing matrix is \( \Theta \). If the BSS problem can be solved, the modal parameters can be extracted from the estimated mixing matrix \( \Theta \) and modal response \( \hat{q}(t) \). To solve the BSS problem shown in (1), the SOBI algorithm [25, 26] is adopted and a modal identification is proposed in our previous work.

After the discrete sampling of the physical acceleration of \( l \) channels at \( N \) identical time intervals, the SOBI-based modal identification method forms a vector \( Y_k \in \mathbb{R}^{2l \times 1} \), which consists of the measured seismic response of \( l \) channels \( y_k \in \mathbb{R}^{n \times 1} \) and its time-lagged data.

\[
Y_k = \left( y_k^T y_{k+1}^T \cdots y_{k+p-1}^T \right)^T, \quad (k = 1, 2, \ldots, L),
\]

(2)

where \( y_k = \begin{bmatrix} y_k^R \\ y_k^I \end{bmatrix} \) is the measured physical acceleration of \( l \) channels at time interval \( k \), \( y_k^R \) is its Hilbert transformation, and \( L = N - p + 1 \). If \( l \geq n \), then parameter \( p = 1 \), and no time-lagged data are needed; otherwise, the time delay should
satisfy \( 2p \times l > n \). \( n \) is the system order, which is equal to the number of structural active modal orders multiplied by 2.

The Hankel matrix \( \mathbf{H}(\tau_i) \in \mathbb{R}^{2p \times l} \) with time delay \( \tau_i \) can then be defined as

\[
\mathbf{H}(\tau_i) = E \begin{bmatrix} \mathbf{Y}_i \mathbf{Y}_i^T \end{bmatrix}
\]

\[
= \begin{bmatrix}
\mathbf{R}_{YY}(\tau_i) & \mathbf{R}_{YY}(\tau_{i+1}) & \cdots & \mathbf{R}_{YY}(\tau_{i+p-1}) \\
\mathbf{R}_{YY}(\tau_{i+1}) & \mathbf{R}_{YY}(\tau_{i+2}) & \cdots & \mathbf{R}_{YY}(\tau_{i+p}) \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{R}_{YY}(\tau_{i+p-1}) & \mathbf{R}_{YY}(\tau_{i+p}) & \cdots & \mathbf{R}_{YY}(\tau_{i+2p-2}) 
\end{bmatrix},
\]

(3)

If \( l \geq n \), \( p = 1 \) and the Hankel matrix becomes the covariance function matrix \( \mathbf{R}_{YY}(\tau_i) = E[\mathbf{y}_i \mathbf{y}_i^T] \). \( E[\cdot] \) is the expectation operator. For a structure under white-noise-like support excitation, the covariance function matrix of the measured acceleration \( \mathbf{R}_{yy}(\tau_i) \) is related only to the initial state and system parameters (modal parameters) of the structure, which is similar to that of the structural free vibration response and the impulse response when time \( \tau_i \) is big enough.

By further deduction, the Hankel matrix \( \mathbf{H}(\tau_i) \) can be expressed as

\[
\mathbf{H}(\tau_i) = \begin{bmatrix}
\mathbf{\Theta} \\
\mathbf{\Theta}\Sigma(T_i) \\
\vdots \\
\mathbf{\Theta}\Sigma^{p-1}(T_i) \\
\end{bmatrix}
\]

\[
\cdot \begin{bmatrix}
\mathbf{\Theta}^H \\
\mathbf{\Sigma}(T_i)^H \\
\vdots \\
\mathbf{\Sigma}^{p-1}(T_i)^H \\
\end{bmatrix} = \mathbf{\Theta}\mathbf{R}_{qq}(\tau_i)^{-1}
\]

(4)

where \( \Sigma(T_i) = \exp(\Lambda T_i), \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \), and \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of the system matrix \( \mathbf{A} \).

\[
\mathbf{R}_{qq}(\tau_i) = E[\mathbf{q}_i \mathbf{q}_i^T] = \begin{bmatrix} \mathbf{R}_{q_i q_i}(\tau_i) & 0 \\
0 & \mathbf{R}_{q_i q_i}(\tau_i) \end{bmatrix}.
\]

\( \mathbf{C}_u = \mathbf{C}_{2u} \) is the output selection matrix of acceleration. Superscript \( H \) represents Hermitian (complex conjugate and transpose).

In order to realize the diagonalization of the Hankel matrix \( \mathbf{H}(\tau_i) \), a robust algorithm called the joint approximate diagonalization (JAD) [27] technique is adopted. Instead of diagonalizing only one Hankel matrix, a group of Hankel matrices with different time delays \( \mathbf{H}(\tau_i), \mathbf{H}(\tau_j), \ldots, \mathbf{H}(\tau_k) \) are diagonalized simultaneously in the JAD technique. For the JAD technique, the corresponding optimization problem is expressed as follows:

\[
\min \sum_{i=1}^{k} \left\| \mathbf{U}^H \mathbf{W} \mathbf{H}(\tau_i) \mathbf{W}^H \mathbf{U} \right\|^2 \quad (\tau_i > \tau_0)
\]

(5)

The whitening matrix \( \mathbf{W} \) is obtained using the PCA method.

After solving the optimization problem shown above, the generalized diagonalization matrix \( \mathbf{U} \) can be obtained. The mixing matrix \( \mathbf{\Theta} \), separation matrix \( \mathbf{\Theta}^+ \), and source signal, that is, the modal response \( \mathbf{q}(t) \), can then be expressed as

\[
\mathbf{\Theta} = \mathbf{W}^H \mathbf{U};
\]

\[
\mathbf{\Theta}^+ = \mathbf{U}^H \mathbf{W};
\]

(6)

The submatrix \( \mathbf{\Theta} \) of \( \mathbf{\Theta}^+ \) contains the real parts \( \mathbf{\Psi}^R = \mathbf{C}_\mathbf{\Psi}^R \) and the imaginary parts \( \mathbf{\Psi}^I = \mathbf{C}_\mathbf{\Psi}^I \) of the observed complex mode shape. Using the identified real modal response \( \mathbf{\dot{q}}^R(t) \) and the modal identification method of the single-degree-of-freedom (SDOF) system, the natural frequencies \( \nu_i \) \( (i = 1, \ldots, n) \) and damping ratios \( \xi_i \) \( (i = 1, \ldots, n) \) can be obtained. Some details of the above modal identification method can be found in our previous work.

2.2. Modal Contribution and System Order Determination. In the modal identification process presented above, activated system order \( n \) is an important parameter. However, it is usually unknown previously. Moreover, practical ambient excitation may have some dominant frequency components and may not be like white noise support excitation at all. Therefore, it is of great importance to determine the system order and discard the dominant frequencies of ambient excitation for the OMA of concrete dams. Recently, Cara et al. proposed a new index, that is, modal contribution, to the system order and discard the dominant frequencies of ambient excitation for the OMA of concrete dams. Here, we extend this method to the SOBI-based modal identification procedure and propose an automated operation modal analysis procedure for concrete dams.

Let \( \mathbf{Y}' = [\mathbf{y}_1', \mathbf{y}_2', \ldots, \mathbf{y}_N'] \in \mathbb{R}^{k \times N} \) be the measured physical acceleration response of a dam at \( N \) different times; \( \mathbf{Y}' \) can then be expressed as

\[
\mathbf{Y}' = \mathbf{Y}^m + \mathbf{Y}^e = \sum_{k=1}^{n} \mathbf{Y}^{m_k} + \mathbf{Y}^e,
\]

(7)

where \( \mathbf{Y}^m \) is the acceleration due to identified modes, \( \mathbf{Y}^{m_k} = [\mathbf{y}_{1k}, \mathbf{y}_{2k}, \ldots, \mathbf{y}_{Nk}] \) is the error term of the measured acceleration.

According to (1) and (2), the observed acceleration response \( \mathbf{y}_k = \mathbf{y}_k'^R \) of a structure can be expressed by

\[
\mathbf{y}_k = \sum_{r=1}^{n} \mathbf{\Psi}_r^R \mathbf{\dot{q}}_{rk}^R + \mathbf{\Psi}_r^I \mathbf{\dot{q}}_{rk}^I + \mathbf{v}_k = \sum_{r=1}^{n} \mathbf{y}_k^{m_r} + \mathbf{v}_k,
\]

(8)

where \( \mathbf{y}_k^{m_r} \) is the acceleration due to the \( r \)th mode; \( \mathbf{\Psi}_r^R \) and \( \mathbf{\Psi}_r^I \) are the \( r \)th columns of matrices \( \mathbf{\Psi}^R \) and \( \mathbf{\Psi}^I \), respectively; and \( \mathbf{\dot{q}}_{rk}^R \) and \( \mathbf{\dot{q}}_{rk}^I \) are the \( r \)th real and imaginary modal response coordinates, respectively, at the \( k \)th time interval.

For the SOBI-based modal identification method mentioned in Section 2.1, \( \mathbf{\Psi}_r^R \), \( \mathbf{\Psi}_r^I \), \( \mathbf{\dot{q}}(t) \), and \( \mathbf{\dot{q}}(t) \) can be calculated using (6). The acceleration due to each mode \( \mathbf{y}_k^{m_r} \) can then be calculated as shown in (8).
The covariance matrix can be calculated as

\[
(Y'Y'^T)_D = (Y^mY'^T)_D + (Y^eY'^T)_D,
\]

(9)
The operator \((\cdot)_D\) means that \((M)_D\) is a matrix consisting of the diagonal elements of matrix \(M\) and zeros elsewhere. Normalizing the above equation,

\[
(Y'Y'^T)_D^{-1} (Y'Y'^T)_D = \sum_{k=1}^{n} (Y'^mY'^T)_D^{-1} (Y'^mY'^T)_D + (Y'^eY'^T)_D^{-1} (Y'^eY'^T)_D.
\]

(10)

For the diagonal elements of the diagonal matrices, the following expression can be obtained:

\[
[1]_m = \sum_{k=1}^{n} \Delta_{m_k} + \Delta_s,
\]

(11)

where \(\Delta_{m_k} = \text{diag}(\Delta_{m_k}(1), \Delta_{m_k}(2), \ldots, \Delta_{m_k}(l))\) and \(\Delta_s = \text{diag}(\Delta_s(1), \Delta_s(2), \ldots, \Delta_s(l))\) are the diagonal elements of the diagonal matrices \((YY'^T)_D^{-1} (Y^mY'^T)_D\) and \((YY'^T)_D^{-1} (Y^eY'^T)_D\), respectively.

The two metrics \(\delta_{m_k}\) and \(\delta_s\) can then be calculated using the diagonal elements of the above two matrices as

\[
\delta_{m_k} = \frac{1}{l} \sum_{j=1}^{l} \Delta_{m_k}(j);
\]

\[
\delta_s = \frac{1}{l} \sum_{j=1}^{l} \Delta_s(j).
\]

(12)

As noted by Cara et al., \(\delta_{m_k}\) indicates the importance of the \(k\)th identified mode; the modal contributions of these spurious modes and weak modes are almost zero. In the vector \(\delta_{m} = (\delta_{m_1}, \delta_{m_2}, \ldots, \delta_{m_n})\), the greater the number of nonzero elements is, the more the activated modes are included in the vibration response and the more abundant the information therein is.

2.3. The Flowchart of Modal Identification. After extracting the modal parameters, the MAC and COMAC could be calculated using the following equations:

\[
\text{MAC}(j) = \frac{\Phi^b_j \Phi^e_j}{\left(\Phi^b_j \Phi^b_j\right)\left(\Phi^e_j \Phi^e_j\right)},
\]

\[
(j = 1, 2, \ldots, n),
\]

(13)

\[
\text{COMAC}(k) = \frac{\sum_{j=1}^{n} |\Phi^b_{jk} \cdot \Phi^c_{jk}|^2}{\sum_{j=1}^{n} \Phi^b_{jk}^4 + \sum_{r=1}^{n} \Phi^e_{jk}^2},
\]

\[(k = 1, 2, \ldots, l),
\]

(14)

where \(\Phi^b_j\) is the \(j\)th modal shape vector of an undamaged structure; \(\Phi^e_j\) is the \(j\)th modal shape vector of a structure with an unknown state; and \(\Phi^b_{jk}\) and \(\Phi^c_{jk}\) are the \(k\)th components of modal shape vectors \(\Phi^b_j\) and \(\Phi^c_j\), respectively.

Many researchers have proved that MAC and COMAC are more sensitive to structural damage than natural frequencies and modal shape vectors. In addition, COMAC can also provide damage location information of a structure.

According to the SOBI-based modal identification methods, stabilization diagram, and proposed modal contribution index, an OMA procedure for concrete dams using AVT data is proposed. The flowchart of the integrated OMA method is shown in Figure 1.

3. Dam Health Monitoring Using an ICA-Based Method

The preceding section described a procedure to obtain an \(n\)-dimensional feature space of modal features. In a situation where multidimensional feature vectors exist, several monitoring methods may be employed for feature vector discrimination [28]. As a result of environmental variables and some other stochastic factors, the \(n\)-dimensional feature vector is random. For a multidimensional random variable, the statistical process control (SPC) based method can be adopted to evaluate the structural health based on identified modal features. SPC is a collection of tools useful in process monitoring. The control chart is the most commonly used one and is very suitable for automated continuous system monitoring.
Hereafter, a newly developed ICA-based SPC method is introduced into the dam health monitoring process.

3.1. Perform ICA on the Extracted Modal Features. Assume that the identified modal features (natural frequencies, modal shape vectors, MAC or COMAC, etc.) are arranged in a zero-mean vector $x(t)$ with $m$ components, $x_1, x_2, \ldots, x_m$, which can be expressed as [1, 16]

$$x(t) = f(H, T, \theta, \ldots) + g(\eta) = x(t) + \varepsilon(t), \quad (15)$$

where $f(\cdot)$ is a vector nonlinear function that projects environmental variables, such as water level $H$, temperature $T$, and time effect $\theta$, into the determined components $x(t)$ of structural dynamic features. For a concrete dam without damage, $x(t)$ preserves the major information of the calculated modal features $\tilde{x}(t)$, which is mainly the effect of some environmental variables; $\varepsilon(t) = g(\eta)$ is the random component, which contains the influence of some other random factors $\eta$, such as the identification error of modal features.

The nonlinear mapping $f(\cdot)$ is expressed as another nonlinear mapping $\Theta(\cdot)$ multiplied by a linear mapping $\Lambda$, that is,

$$\tilde{x}(t) = f(H, T, \theta, \ldots) = \Lambda [\Theta(H, T, \theta, \ldots)] = \Lambda \xi(t), \quad (16)$$

where $\xi(t) = [\xi_1, \ldots, \xi_n]^T$ is called the latent variables vector. Latent variables $\xi_1, \ldots, \xi_n$ are independent of each other. The number of latent variables $s \leq m$.

For a concrete dam that has worked a long time or has recently undergone a large earthquake or flood, some damage may be found, and the damage will change the probability distribution of random vector $\varepsilon(t)$. Therefore, the structural damage can be detected by comparison of the probability distribution of $\varepsilon(t)$ after the determined components have been extracted. Hereafter, the ICA-based method is adopted to identify the linear mapping matrix $\Lambda$ and the latent variables $\xi_1, \ldots, \xi_n$ also called independent components (ICs), to calculate the two parts, that is, $\tilde{x}(t)$ and $\varepsilon(t)$.

The principle of ICA estimation is based on the central limit theorem (CLT), which states that the sum of independent random variables tends to be distributed towards Gaussian, regardless of the underlying distribution. This means that a mixture of some independent random variables is always Gaussian for the original variables. Usually, in ICA, the original feature samples must be whitened first. The whitened variables $z(t)$ are obtained by

$$z(t) = Q x(t) = Q \Lambda \xi(t) + Q \varepsilon(t), \quad (17)$$

where the whitening matrix $Q = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)^{-1/2} \beta_1, \beta_2, \ldots , \beta_m^T$ is obtained using PCA. $\lambda_i, i = 1, \ldots, m$, are eigenvalues of the covariance matrix $C_x = E(xx^T)$; $\beta_i, i = 1, \ldots, m$, are its eigenvectors.

The goal of ICA is to find a de-mixing matrix $W = O^T Q$ from the original observed data $x(t)$ such that the estimated ICs, that is, latent variables, can be expressed as

$$\tilde{\xi}(t) = W x(t) = O^T z(t), \quad (18)$$

where $O$ is an orthonormal matrix and determined by maximizing the non-Gaussianity of $\tilde{\xi}(t)$. The non-Gaussianity of a random variable $v$ can be measured by negentropy. The negentropy can be approximated by the kurtosis $\text{kurt}(v)$

$$\text{kurt}(v) = E[v^4] - 3 \left( E[v^2] \right)^2. \quad (19)$$

For a Gaussian random variable $v$, negentropy equals zero, and $\text{kurt}(v) = 0$. With increasing non-Gaussianity of a random variable, negentropy increases. Therefore, the basic principle of the ICA estimation method can be summarized as finding the proper demixing matrix $W$, which can maximize the non-Gaussianity measured by the negentropy of each recovered IC.

Another measure of the non-Gaussianity of a random variable is the entropy-based negentropy, which is more statistically justified. The entropy of a discrete random variable can be calculated as

$$J(v) = H(v_{\text{gaus}}) - H(v), \quad (21)$$

where $v_{\text{gaus}}$ is a standard Gaussian random variable with zero mean and unit variance.

The proper demixing matrix $W = O^T Q$ is obtained through an optimization process which maximizes the non-Gaussianity of each latent variable and the non-Gaussianity is measured by negentropy calculated using (21). In order to realize this optimization object, the most widely used algorithm, that is, the FastICA algorithm [29], is adopted in this study. Then the ICs called latent variables for the process monitoring problem can be calculated using (18). In the error space, the square prediction error (SPE) can be calculated for use as an index of structural health.

$$\text{SPE}(t) = \left[ x(t) - Q^{-1} R_d^{-1} O_d Q x(t) \right]^T \cdot \left[ x(t) - Q^{-1} R_d^{-1} O_d Q x(t) \right], \quad (22)$$

where $Q_d$ is composed of the first $d$ row vectors of $O$.

Selecting an appropriate parameter $d$ is a key step of the above procedure. This parameter determines the division of the observation space into the dominant components space and the residual parts space. Based on Wang’s proof, ICs could be ranked and selected by the value of the $L_2$ norm of a mixing matrix $\Lambda$ column according to the minimum mean square error (MSE). The criterion can be expressed as

$$\arg \min_{\xi_i} / (\tilde{\xi}) = \arg \max_{\xi_i} \sum_{i=1}^{d} \|a_i\|^2, \quad (23)$$

where $a_i$ is the $i$th column of mixing matrix $\Lambda$. 
3.2. Setting Control Limits for SPE. When the majority of non-Gaussianity is removed by the ICs, the noise in the residuals $\varepsilon(t)$ will approximately follow a Gaussian distribution, so the control limit of SPE can be calculated from the following weighted $\chi^2$ distribution:

$$\text{SPE} \sim \mu x^2. \quad (24)$$

In addition, the kernel density estimation (KDE) method proposed by Martin and Morris [30] can also be used to determine the confidence limit of the SPE statistic because it does not require an assumption for the distribution of data. Using the KDE-based method, the probability function $\hat{D}(v)$ of the SPE index can be evaluated by

$$\hat{D}(v) = \frac{1}{N_\eta h} \sum_{j=1}^{n} \kappa \left( \frac{v - \mu_j}{h} \right), \quad (25)$$

where commonly used kernel function $\kappa(\cdot)$ is a Gaussian function, $N_j$ is the sample number, $h$ is the time window length, and $\mu_j$ is the $i$th monitoring data point.

Given a testing level $\alpha$, the upper control limit (UCL) can be determined using the estimated probability function as follows:

$$P(v < \text{UCL}) = 1 - \alpha. \quad (26)$$

The UCL is calculated to determine whether the SPE index is normal. When the calculated SPE index is greater than the UCL, the structure may be abnormal. If the analyzed features are modal shapes, COMAC, and other metrics that can provide information on the damage location, the distribution of SPE index for different sensors indicates the location of structural damage. Then, another analysis tool called contribution plots [31] is used to locate the position of sensors with abnormal information.

The SPE shown in (22) can be rewritten in the following expression:

$$\text{SPE} = \frac{1}{j=1} \left\| \mathbf{x}^{j} - \mathbf{X} \right\|^2 = \frac{1}{j=1} \text{Cspe}_{ej}, \quad (27)$$

where $\mathbf{x}^{j}$ and $\mathbf{X}$ are the identified and the reconstructed modal features (natural frequencies, modal shape vectors, or COMAC) of the $j$th mode, respectively. The reconstructed modal features are obtained by ICA. Cspe$_{ej}$ is the contribution of component $j$ of the vibration characteristic to the SPE norm.

3.3. The Flowchart of ICA-Based Dam Health Monitoring. Now, a detailed step-by-step description of the dam health monitoring method based on AVT and two different BSS methods will be given in the following. The flow chart of this dam health monitoring method is shown in Figure 2.

Step 1. Through the vibration measurement system of a concrete dam, the AVT data can be obtained at different times, when the structure can be deemed as normal. The original AVT data may include the disturbance of measurement noise. Therefore, some preprocessing procedures are necessary to be adopted to improve the signal-to-noise ratio (SNR) of the AVT data.

Step 2. With these preprocessed AVT data, some basic modal parameters of structure are identified using the proposed modal identification procedure shown in Figure 1. Then, using (14), the COMAC index is calculated. For each type of modal feature, construct a data matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$, which includes the reference data for dam health monitoring.

Step 3. After normalizing these data of the modal features by their mean $\mu$ and variance $\sigma$, the ICA is adopted to analyse the normalized multidimensional time series $\mathbf{x}_1, \ldots, \mathbf{x}_n$ to extract some ICs. The determined component is $\mathbf{x}(t)$, and noise residuals are $\varepsilon(t)$. The SPE metrics are calculated using (22). Given a testing level $\alpha$, (26) determines the UCL of SPE for a dam without damage.

Step 4. For each newly obtained AVT data, the above-mentioned Steps 1–3 are repeated and the SPE index is calculated. These newly calculated SPE indices are compared with the UCL calculated using (26) to determine whether the structure is normal. If the new SPE metrics exceed the control limit, the dam is abnormal; otherwise, it is still in the normal state.

Step 5. When the calculated SPE index is greater than the UCL and the structure is deemed as abnormal, the SPE contribution plots and the local modal features are used to locate the structural damage. Local features are those modal features which can provide damage location information. If some obvious peaks appeared on the SPE contribution plots around some sensors number, damage may occur around these sensors.

4. Case Study

4.1. Numerical Verification. Before analysing practical AVT data, the simulated AVT data using FEM are used to validate the proposed dam health monitoring method. The 3D finite element model of the maximum dam block of a concrete gravity dam is constructed as shown in Figure 3. The locations of 10 acceleration sensors with three measurement directions and a simulated crack are shown in Figure 3(a). The width of the dam block is 10 m. The dynamic elastic modulus, Poisson’s ratio, and mass density of dam concrete are 31.0 GPa, 0.2, and 2643 kg/m$^3$, respectively; the dynamic elastic modulus and Poisson’s ratio of foundation rock are 20.0 GPa and 0.25, respectively. A massless foundation is assumed to model the dam-foundation interaction when calculating the vibration response of a structure. The ambient excitation used in each FEM calculation is a white noise time series, which is generated using the MATLAB program. The AVT data are obtained by adding the calculation results of the FEM software MSC.Marc [32] using the mode superposition method (10 modes are used) with some noise. The simulated AVT record of sensor #1 in 3 measurement channels is shown in Figure 4. The sampling frequency of the simulated acceleration measurement is 100 Hz and the total number of sample is 1000.
Referenced AVT data

Modal features

Calculate SPE

Determine upper control limit UCL

New AVT data

Modal features

Mean $\mu$ and variance $\sigma$

De-mixing matrix $W$

ICA

Calculate SPE

Normalize data

Global features

Local features

Yes SPE > UCL?

No SPE > UCL?

SPE > UCL?

Yes Abnormal

No Normal

Figure 2: The dam health monitoring method based on AVT and ICA.

Figure 3: (a) Model size, a simulated crack and the location of acceleration sensors; (b) the finite element model of the dam-reservoir-foundation system.
Table 1: A comparison of the calculated and identified no-damping natural frequencies (unit: Hz).

<table>
<thead>
<tr>
<th>Method</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEM</td>
<td>1.05</td>
<td>2.51</td>
<td>3.77</td>
<td>5.56</td>
<td>6.25</td>
<td>7.91</td>
<td>8.34</td>
<td>10.91</td>
<td>11.93</td>
<td>12.53</td>
</tr>
<tr>
<td>SSI</td>
<td>1.06</td>
<td>2.50</td>
<td>3.78</td>
<td>5.57</td>
<td>6.24</td>
<td>7.85</td>
<td>8.35</td>
<td>10.91</td>
<td>11.92</td>
<td>12.53</td>
</tr>
<tr>
<td>SOBI</td>
<td>1.06</td>
<td>2.51</td>
<td>3.77</td>
<td>5.57</td>
<td>6.24</td>
<td>7.93</td>
<td>8.35</td>
<td>10.91</td>
<td>11.93</td>
<td>12.54</td>
</tr>
</tbody>
</table>

Table 2: The false alarm rate and missing alarm rate of the dam health monitoring.

<table>
<thead>
<tr>
<th>Case</th>
<th>Alarm rate</th>
<th>PCA</th>
<th>KPCA</th>
<th>ICA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>COMAC</td>
<td>Frequency</td>
<td>COMAC</td>
</tr>
<tr>
<td>1</td>
<td>$P_F$</td>
<td>2.5%</td>
<td>0.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>$P_C$</td>
<td>97.5%</td>
<td>99.5%</td>
<td>100.0%</td>
</tr>
<tr>
<td>2</td>
<td>$P_F$</td>
<td>2.0%</td>
<td>0.0%</td>
<td>1.5%</td>
</tr>
<tr>
<td></td>
<td>$P_C$</td>
<td>98.0%</td>
<td>100.0%</td>
<td>98.5%</td>
</tr>
<tr>
<td>3</td>
<td>$P_F$</td>
<td>57.5%</td>
<td>7.5%</td>
<td>7.5%</td>
</tr>
<tr>
<td></td>
<td>$P_C$</td>
<td>42.5%</td>
<td>92.5%</td>
<td>92.5%</td>
</tr>
<tr>
<td>4</td>
<td>$P_F$</td>
<td>34.0%</td>
<td>0.0%</td>
<td>15.5%</td>
</tr>
<tr>
<td></td>
<td>$P_C$</td>
<td>66.0%</td>
<td>100.0%</td>
<td>84.5%</td>
</tr>
<tr>
<td>5</td>
<td>$P_F$</td>
<td>11.0%</td>
<td>0.0%</td>
<td>5.0%</td>
</tr>
<tr>
<td></td>
<td>$P_C$</td>
<td>89.0%</td>
<td>100.0%</td>
<td>9.5%</td>
</tr>
</tbody>
</table>

Figure 4: Simulated AVT data of sensor #1 in 3 measurement channels.

To validate the modal identification procedure based on SOBI and modal contribution, the simulated AVT data are used to extract structural modal parameters using SOBI and SSI. The practical modal parameters of the dam are calculated using FEM. The identified and calculated no-damping natural frequencies are shown in Table 1. As shown in Table 1, the maximum relative errors of the modal identification results for the SOBI and SSI methods are 0.95% and 1.03%, respectively.

Setting $n_{\text{max}} = 50$, the modal contribution index and singular value spectrum are shown in Figure 5. As shown in Figure 5(a), when the system order is higher than 20, the modal contribution index tends to become a stable constant, whereas in the singular value spectrum, no obvious drop is observed. The increase in noise level will increase the modal contribution index, but the position of the turning point will not be changed. The modal contribution of 10 activated modes is shown in Figure 5(c).

The relationship between the identified first- and second-mode natural frequencies and the water level is shown in Figure 6. Because the reservoir water will provide additional mass to the dam-foundation-reservoir system, Figure 6 shows that the natural frequencies decrease with increasing water level.

In this numerical example, to evaluate the impact of environmental variables (here, only the water level is simulated) and damage extent on the identification results of vibration features, five simulation cases are designed. The crack lengths corresponding to the five cases are 0 m, 0 m, 2 m, 4 m, and 6 m. For each of the five cases, 200 different water levels are selected randomly within the water level range [75 90] m, and then, the acceleration responses measured by sensors numbered #1~#10 are calculated using the FEM. Case number 1 is designed to generate referenced data, and case number 2 is used to verify the false alarm rate of the proposed method. For the simulated AVT data of case number 1 and case number 5, the identified modal natural frequencies of the 1st mode and the 10th mode are shown in Figure 7. As shown in this figure, the effect of damage cannot be distinguished directly as a result of the variation in water level.

Using the modal features, that is, the natural frequencies and COMAC, PCA, KPCA, and ICA are adopted to monitor the dam health. A comparison of the false alarm rate and missing alarm rate of the three analysis methods is shown in Table 2. For the PCA and ICA method, the SPE index and UCL are shown in Figure 8. The false alarm rate $P_F$ and correct alarm rate $P_C$ can be defined as

$$P_F = \frac{N_F}{N} \times 100\%;$$
$$P_C = 1 - P_F,$$

(28)
where, $N$ is the total number of modal features and $N_F$ is the number of false alarming.

It can be seen in Table 2 and Figure 9 that the ICA-based dam health monitoring method has higher alarm accuracy than the other two methods, especially for a dam with a small damage extent. For the simulated case 4, the contribution plots of SPE are shown in Figure 10. As shown in this figure, an obvious peak of SPE can be found near sensor #5, which is just the location where the crack is simulated.

In practical engineering application, acceleration sensors may be not just installed near a crack zone, since the location of crack is commonly unknown in advance. In order to verify the performance of the proposed ICA-based SHM method when sensors are far away from structural damage zone, the acceleration sensors are rearranged, as is shown in Figure 11. For the new arrangement of sensors, the dam health monitoring results using COMAC index and ICA-based method are shown in Figure 12. The SPE control chart shows the damage can be detected effectively. The contribution plots show two peaks at the sensor numbered #5 and #6, but the two peaks are not very obvious since the distance between the two sensors and crack zone is little far (4.5 m indeed).
4.2. Analysing the Practically Measured AVT Data. A hydropower station is located in the middle stream of the Minjiang River in Fujian Province of China. The maximum height of the roller-compacted concrete (RCC) gravity dam is 101.0 m, and its normal flood level elevation is 65.0 m, with a corresponding storage capacity of $2.6 \times 10^9$ m$^3$. This project
Figure 12: Dam health monitoring results using COMAC index for the ICA-based method, (a) SPE control chart and (b) contribution plots at the sample number $n_1$ and (c) at the sample number $n_2$ of case 5.

is located near the Taiwan Strait seismic zone, so seismographs are installed on the 19th and 25th dam blocks. The arrangement and location of these seismographs are shown in Figure 13. Because each dam monolith in the concrete gravity dam works independently, in this study, only the vibration response record of dam monolith number 19 is studied.

The ambient vibration response record of 56 different measurement times is used. The sample size of each AVT record is 16,000, and the sampling frequency is 100 Hz. The AVT data of all nine channels measured on August 19, 2003 are shown in Figure 14.

Using the AVT data of 56 different measurement times, the modal parameters are identified based on the proposed automated modal identification procedure. Seven modes with larger modal contributions than the other modes are used to monitor the dam health, and their identification results are shown in Figure 15(a). The modal contribution of these seven modes is shown in Figure 15(b).

The identified natural frequencies (7 orders) and calculated COMAC index using the AVT data of the dam for the first 40 times are used as reference data. For the other 16 measurement times, the structural health is diagnosed using the PCA, KPCA, and ICA methods. The dam health diagnosis result based on these data is shown in Figures 16 and 17. UCL is calculated by setting the test level $\alpha = 0.01$ using (26). From the two figures, we can see that the health state of the
Figure 14: The AVT data of the dam monolith number 19 measured on August 19, 2003.

dam is normal at the 16 different times, which is consistent with the diagnosis result using the static monitoring data of displacement and flow seepage.

5. Conclusions

The dam health monitoring method presented in this work includes a modal identification procedure based on SOBI and modal contribution index and a SPC method based on ICA. The modal contribution index calculation method for SOBI-based modal identification is studied to determine the system orders. ICA and SPE contribution plots are used for the calculated modal features to detect structural damage and determine the location of it. The numerical example and the engineering example have verified the good performance of these proposed methods. The advantages of the vibration-based structural health monitoring method have been widely applied in many fields of civil engineering. For dam and some other concrete hydraulic structures, since much more attentions have been paid to the construction of vibration...
Figure 15: (a) Identification results of natural frequencies using the AVT data of 56 different times and (b) modal contribution index for each identified mode.

Figure 16: Dam health monitoring using the PCA, KPCA, and ICA methods for the referenced data and compared data of natural frequencies.
monitoring systems, the dam health monitoring method proposed in this work is expected to be widely applied in realistic engineering.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants nos. 51409205, 51279052, 51409018, and 51309189), Project Funded by China Postdoctoral Science Foundation (Grant no. 2015M572656XB) and Postdoctoral Science Foundation of Shanxi Province, Open Foundation of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering (Grant no. 2014491011), the Program 2013KCT-15 for Shanxi Provincial Innovative Research Team, and the Innovative Research Team of Institute of Water Resources and Hydroelectric Engineering, Xi’an University of Technology (Grant no. 2016ZZKT-14).

References


