Smart Structures and Systems, Vol. 11, No. 1 (2013) 19-34 19
DOI: http://dx.doi.org/10.12989/sss.2013.11.1.019

Application of recursive SSA as data pre-processing filter for stochastic subspace identification

Chin-Hsiung Loh* and Yi-Cheng Liu

Department of Civil Engineering, National Taiwan University, Taipei 10617, Taiwan

(Received March 22, 2012, Revised January 8, 2013, Accepted January 9, 2013)

Abstract. The objective of this paper is to develop on-line system parameter estimation and damage detection technique from the response measurements through using the Recursive Covariance-Driven Stochastic Subspace identification (RSSI-COV) approach. To reduce the effect of noise on the results of identification, discussion on the pre-processing of data using recursive singular spectrum analysis (rSSA) is presented to remove the noise contaminant measurements so as to enhance the stability of data analysis. Through the application of rSSA-SSI-COV to the vibration measurement of bridge during scouring experiment, the ability of the proposed algorithm was proved to be robust to the noise perturbations and offers a very good online tracking capability. The accuracy and robustness offered by rSSA-SSI-COV provides a key to obtain the evidence of imminent bridge settlement and a very stable modal frequency tracking which makes it possible for early warning. The peak values of the identified 1st mode shape slope ratio has shown to be a good indicator for damage location, meanwhile, the drastic movements of the peak of 2nd mode slope ratio could be used as another feature to indicate imminent pier settlement.

Keywords: stochastic subspace identification; singular spectrum analysis; recursive identification; bridge scouring

1. Introduction

The major reason for bridge collapse during typhoon and flood is the bridge scoring and this scoring may empty the foundation soil and cause the reduction of bridge bearing capacity, and is the primary cause of bridge failures. Because lack of bridge monitoring system as well as the monitoring techniques, it is impossible to send an early warning message before collapse. Based on all damage cases and harsh environmental conditions, it is necessary to upgrade the current bridge monitoring system on bridge structure, and develop reliable self-diagnosis monitoring and early warning system.

For output-only measurements the Stochastic Subspace Identification (SSI) technique is a well-known multivariate identification technique. It was proved by several researchers to be numerically stable, robust to noise perturbation and suitable for conducting non-stationarity of the ambient excitations although its assumption is violated (Michèle *et al.* 2001). The stochastic realization algorithm mainly focused on SSI-DATA was fully enhanced by (Van and De Moor

∗Corresponding author, Professor, E-mail: lohc0220@ccms.ntu.edu.tw

1996, Bart 2000, Bart and Guido 2000), its application in understanding the dynamic characteristics of a cable-stayed bridge had been studied (Weng *et al.* 2008). As opposed to SSI-DATA, the SSI-COV algorithm avoids the computation of orthogonal projection; instead, it is replaced by converting raw time histories in co-variances of the so-called Toeplitz matrix. The merits of SSI are: (a) The identified modes are given in frequency stability diagram, from which the operator can easily distinguish structural from the computational ones, (b) Since the maximum model order is changeable for the operator, a relative large model order can improve the quality of the identified modal parameters, (c) Mode shapes are simultaneously available with the poles, without requiring a second step to identified them, (d) These methods are robust against non-stationary excitation and thus are applicable to structural vibration study in the presence of ambient excitation.

Different from the off-line analysis, the on-line system identification and damage detection, based on vibration data measured from the ambient vibration of structure has received considerable attention recently. Therefore, the recursive stochastic realization by either the classical Covariance-driven SSI algorithm (RSSI-COV) or Data-driven SSI algorithm (RSSI-DAA) was proposed in (Goethals *et al.* 2004, Loh *et al.* 2011). In this paper, the recursive stochastic subspace identification (RSSI-COV) are discussed and used to estimate time-varying system natural frequencies directly from the response measurements. To avoid time-consumption of SVD in RSSI, the Extended Instrumental Variable version (EIV-PAST) is applied in SSI-COV. In addition, to consider the noise contaminated data, a recursive pre-processing technique called recursive singular spectrum analysis technique (rSSA) is proposed in this studt to enhance the accuracy and stability in the online tracking capability.

2. Stochastic subspace identification

Assuming a structure under consideration is being excited by un-measurable stochastic input forces, the discrete time stochastic state-space-model can be expressed as

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\, x_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\, x_k + \mathbf{v}_k \end{aligned} \tag{1}$$

where $\mathbf{x}_k \in \Re^{2n\times 1}$ is the state vector and $\mathbf{y}_k \in \Re^{l\times 1}$ is the measurement vector, $\mathbf{w_k}$ and $\mathbf{v_k}$ represent the system noise and measurement noise respectively. Basic Covariance-Driven Stochastic Subspace Identification method (SSI-COV) is to solve the problem through identifying a stochastic state-space model (matrices **A** and **C**) from output-only data. For the application of SSI-COV, instead of arranging the block covariances in the form of Toeplitz matrix, it must adopt the form of a Hankel Covariance matrix, which is the way as it is outlined in ERA (Eigensystem Realization Algorithm) (Caicedo *et al.* 2004). The Hankel Covariance matrix has the following factorization properties

$$\mathbf{H}^{\mathrm{cov}} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \dots & \mathbf{R}_{\mathrm{i}} \\ \mathbf{R}_2 & \mathbf{R}_3 & \dots & \mathbf{R}_{\mathrm{i+1}} \\ \dots & \dots & \dots & \dots \\ \mathbf{R}_i & \mathbf{R}_{i+1} & \dots & \mathbf{R}_{\mathrm{2i\text{-}1}} \end{bmatrix} = \mathbf{O}_i \mathbf{\Omega}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \dots \\ \mathbf{CA}^{i-1} \end{bmatrix} \begin{bmatrix} \mathbf{G} & \mathbf{AG} & \dots & \mathbf{A}^{i-1}\mathbf{G} \end{bmatrix} \tag{2}$$

where $\mathbf{O}_i \in \Re^{li\times 2n}$ is the same observability matrix and $\mathbf{\Omega}_i \in \Re^{2n\times li}$ is the stochastic controllability matrix. The observability matrix can be obtained by applying SVD to the Hankel covariance matrix, and then the system matrices and modal parameters can be extracted in the same manner than that presented in SSI-COV off-line analysis. The Hankel Covariance matrix can be constructed by arranging the output measurements data vectors as follows

$$\mathbf{H}_N^{\text{cov}} = E[\mathrm{y}_k^+ \mathrm{y}_k^{-T}] = \frac{1}{\tilde{p}} \sum_{k=2i}^{N} \mathrm{y}_k^+ \mathrm{y}_k^{-T}$$
$$= \begin{bmatrix} \mathrm{y}_{i+1} & \mathrm{y}_{i+2} & \cdots & \mathrm{y}_{N-i+1} \\ \mathrm{y}_{i+2} & \mathrm{y}_{i+3} & \cdots & \mathrm{y}_{N-i+2} \\ \vdots & \vdots & & \vdots \\ \mathrm{y}_{2i} & \mathrm{y}_{2i+1} & \cdots & \mathrm{y}_N \end{bmatrix} \begin{bmatrix} \mathrm{y}_i^T & \mathrm{y}_{i-1}^T & \cdots & \mathrm{y}_1^T \\ \mathrm{y}_{i+1}^T & \mathrm{y}_i^T & \cdots & \mathrm{y}_2^T \\ \vdots & \vdots & & \vdots \\ \mathrm{y}_{N-i}^T & \mathrm{y}_{N-i-1}^T & \cdots & \mathrm{y}_{N-2i+1}^T \end{bmatrix} = \mathbf{Y}_k^+ \mathbf{Y}_k^{-T} \tag{3}$$

where $\mathrm{y}_k^+ \in \Re^{il\times 1}$ and $\mathrm{y}_k^{-T} \in \Re^{1\times il}$. l is the number of sensors and i is number of block rows which forms the Hankel covariance matrix, and k is ranging over the entire set of available data and $\tilde{p}$ is an optional normalization parameter ($\tilde{p}$ =1, in this study). The software library LAPACK is used in MATLAB to compute SVD, which uses the classical algorithms like Householder reflections and QR algorithm (Golub *et al.* 2006), but it is not suitable for on-line application. The need of an online application of SVD for implementing to SSI-COV becomes an important issue for continuous monitoring.

3. Recursive stochastic subspace identification

Instead of solving the SVD problem using classical approaches, a new approach called Projection Approximation Subspace Tracking (PAST) was initially developed by Yang (1995), who takes advantage of a mathematical lemma to find the required column subspace as an unconstrained optimization problem and try to update the column subspace. The PAST is originally a fast dominant-eigenvectors updating algorithm which is based on the following unconstrained cost function

$$V[\mathbf{W}(t)] = \sum_{k=1}^{t} \left\| \mathrm{z}(k) - \mathbf{W}(t)\mathbf{W}^H(t)\mathrm{z}(k) \right\|^2 \tag{4}$$

where $\mathrm{z}(t) \in \Re^{m\times 1}$ is a random vector. To adapt the solution to a Recursive Least Square (RLS) approach, an “approximation” is introduced

$$\mathrm{h}(k) = \mathbf{W}^H(k-1)\mathrm{z}(k) \tag{5}$$

which replace $\mathbf{W}^H(t)\mathrm{z}(k)$, and the assumption under this approximation is that the signal subspace is slow varying comparing to the sampling rate of data point. With this assumption, since the dominant subspace $\mathbf{W}^H(k-1)$ is already known from the previous step k-1, the original cost function is converted to a quadratic criterion

$$\bar{V}[\mathbf{W}(t)] = \sum_{k=1}^{t} \left\| z(k) - \mathbf{W}(t) h(k) \right\|^2 \tag{6}$$

which is a typical optimization function in Least Square problems and can be minimized by

$$\mathbf{W}(t) = \overline{\mathbf{U}}'(t) = \mathbf{C}_{zh}(t)\mathbf{C}_{h}^{-1}(t) \tag{7}$$

where $\mathbf{C}_{zh}$ is the covariance matrix formed by $\boldsymbol{z}$(k) and $\boldsymbol{h}$(k)

$$\mathbf{C}_{zh}(t) = \sum_{k=1}^{t} z(k) h^H(k) = \mathbf{C}_{zh}(t-1) + z(t) h^H(t) \tag{7a}$$

$$\mathbf{C}_{h}(t) = \sum_{k=1}^{t} h(k) h^H(k) = \mathbf{C}_{h}(t-1) + h(t) h^H(t) \tag{7b}$$

Thus, when there is a new incoming data at instant *t*, the matrix inversion lemma can be applied to Eq. (7) and the well-known RLS algorithm can be easily derived for updating $\mathbf{W}(t)$.

In output-only SSI, the input source $\mathbf{B}u_k$ is unknown, and assumed to be a stationary and spatially white noise. For this type of noise, it was proved in (Söderström and Stoica 1989) that the normal least square formulation will lead to a biased solution and it is not appropriate to handle this type of problem; instead, an **Extended Instrumental Variable** (EIV) approach must be used. The word "Extended" means that to solve the ill-conditioned for the inversion of cross-covariance matrix and to obtain a stable condition in the inversion process. The cost function to be minimized can be replaced by its corresponding EIV formulation

$$\bar{V}[\mathbf{W}(t)] = \left\| \sum_{k=1}^{t} z(k) \xi^H(k) - \mathbf{W}(t) \sum_{k=1}^{t} h(k) \xi^H(k) \right\|_F^2 = \left\| \mathbf{C}_{z\xi}(t) - \mathbf{W}(t) \mathbf{C}_{h\xi}(t) \right\|_F^2 \tag{8}$$

where the subscript $_F$ denotes the Frobenius norm defined as $\sqrt{tr(\boldsymbol{\sigma\sigma}^H)}$. $\mathbf{C}_{z\xi}(t)$ and $\mathbf{C}_{h\xi}(t)$ are defined as follows

$$\mathbf{C}_{z\xi}(t) = \sum_{k=1}^{t} z(k) \xi^H(k) = \mathbf{C}_{z\xi}(t-1) + z(t) \xi^H(t) \tag{8a}$$

$$\mathbf{C}_{h\xi}(t) = \sum_{k=1}^{t} h(k) \xi^H(k) = \mathbf{C}_{h\xi}(t-1) + h(t) \xi^H(t) \tag{8b}$$

The least square solution of (8) is readily found to be

$$\mathbf{W}(t) = \overline{\mathbf{U}}'_{EIV}(t) = \mathbf{C}_{z\xi}(t)\mathbf{C}_{h\xi}^T(t)\left[\mathbf{C}_{h\xi}(t)\mathbf{C}_{h\xi}^T(t)\right]^{-1} \tag{9}$$

The effect of the extra-added Frobenius norm, which is able to fulfill the need of Recursive Least Square (RLS) via matrix inversion lemma, is applied in EIV-PAST.

4. Application of EIV-PAST to RSSI-COV

Let the random vector $\mathbf{z}(t)$ in EIV-PAST formulation be replaced by the corresponding data $\mathrm{y}_k^+ \in \Re^{il\times 1}$; on the other side, the substitution of the instrument $\xi(t)$ is evidently $\mathrm{y}_k^{-T} \in \Re^{1\times il}$. For convenience in notation, a new variable $s = 2i - 1$ is introduced. EIV Solution to the cost function will become

$$\bar{V}[\mathbf{W}(t)] = \left\| \sum_{k=2i}^{t} \mathrm{y}_k^+ \mathrm{y}_k^{-T} - \mathbf{W}(t) \sum_{k=2i}^{t} \bar{\mathrm{h}}(k+s) \mathrm{y}_k^{-T} \right\|_F^2 = \left\| \mathbf{H}_t^{\mathrm{cov}} - \mathbf{W}(t) \overline{\mathbf{H}}_t^{\mathrm{cov}} \right\|_F^2 \tag{10}$$

where $\bar{\mathrm{h}}(k+s) = \mathbf{W}^H(k+s-1)\mathrm{y}_k^+$ is the above mentioned approximation, and $\mathbf{H}_t^{\mathrm{cov}} = \dfrac{1}{p}\sum_{k=2i}^{t} \mathbf{y}_k^+ \mathbf{y}_k^{-T}$ is the definition of Hankel covariance matrix, and $\overline{\mathbf{H}}_t^{\mathrm{cov}} = \sum_{k=2i}^{t} \bar{\mathrm{h}}(k+s) \mathbf{y}_k^{-T}$.
The dominant eigenvectors can be found by

$$\mathbf{W}(t) = \mathbf{U}_1(t) = \left(\mathbf{H}_t^{\mathrm{cov}} \overline{\mathbf{H}}_t^{\mathrm{cov}\,T} \right) \left(\overline{\mathbf{H}}_t^{\mathrm{cov}} \overline{\mathbf{H}}_t^{\mathrm{cov}\,T} \right)^{-1} \tag{11}$$

It is derived to update the dominant eigenvectors of the covariance matrix.
However, the SVD of a Hankel Covariance matrix is defined as

$$\mathbf{H}^{\mathrm{cov}} = \mathbf{USV}^T = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \mathbf{S}_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{pmatrix} \tag{12}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices, $\mathbf{S}$ is a diagonal matrix containing the singular values.
Since $\mathbf{H}^{\mathrm{cov}} \mathbf{H}^{\mathrm{cov}\,T}$ can also be expressed as

$$\mathbf{H}^{\mathrm{cov}} \mathbf{H}^{\mathrm{cov}\,T} = \mathbf{USV}^T \mathbf{VS}^T \mathbf{U}^T = \mathbf{U}\left(\mathbf{SS}^T\right)\mathbf{U}^T = \mathbf{U}\left(\mathbf{SS}^T\right)\mathbf{U}^{-1} \tag{13}$$

which indicated that $\mathbf{U}$ can also be obtained from the ED of the Hankel Covariance matrix multiplied by its transpose. From the relationships shown above, the desired observability matrix $\mathbf{O}_i$ is the same as the column subspace $\mathbf{U}_1$ extracted from Hankel Covariance matrix using SVD. Hence, the so-called Extended Instrumental Variable Recursive Least Square (EIV-RLS) algorithm can be applied to solve the EIV-PAST problem, which fulfills the SVD-updating requirement of RSSI-COV to track the time-varying subspace $\mathbf{U}_1$(t). The explicit formulas to be implemented in RSSI-COV are shown below. Complete derivation of these formulas of EIV-RLS algorithm can be found in (Söderström and Stoica 1989), and listed as follows:

(1) From an initial SVD one can initialize the recursive algorithm with $\mathbf{U}_1(t)$, $\mathbf{P}(t)$ and $\overline{\mathbf{H}}_t$

$$\overline{\mathbf{H}}_t = \mathbf{U}_1^T(t)\mathbf{H}_t \quad \& \quad \mathbf{P}(t) = \left[\overline{\mathbf{H}}_t \overline{\mathbf{H}}_t^T\right]^{-1} \tag{14a}$$

(2) Given a new incoming data vector $\mathbf{y}_{t+1}$, $\mathbf{U}_1(t+1)$, $\mathbf{P}(t+1)$, and using the following algorithm

$$\bar{\mathbf{h}}(t+1) = \mathbf{U}_1^T(t)\mathbf{y}_{t+1}^+ \quad , \quad \mathbf{w}(t+1) = \overline{\mathbf{H}}_t \mathbf{y}_{t+1}^- \tag{14b}$$

$$\mathbf{v}(t+1)=\begin{bmatrix}\mathbf{H}_t\mathbf{y}_{t+1}^{-} & \mathbf{y}_{t+1}^{+}\end{bmatrix} \quad \text{and} \quad \boldsymbol{\psi}(t+1)=\begin{bmatrix}\mathbf{w}(t+1) & \overline{\mathbf{h}}(t+1)\end{bmatrix} \tag{14c}$$

$$\boldsymbol{\Lambda}(t+1)=\frac{1}{\mu^2}\begin{bmatrix}-\left(\mathbf{y}_{t+1}^{-}\right)^T\mathbf{y}_{t+1}^{-} & \mu \\ \mu & 0\end{bmatrix} \tag{14d}$$

$$\mathbf{K}(t+1)=\left[\mu^2\boldsymbol{\Lambda}(t+1)+\boldsymbol{\psi}(t+1)^T\mathbf{P}(t)\boldsymbol{\psi}(t+1)\right]^{-1}\boldsymbol{\psi}(t+1)\mathbf{P}(t) \tag{14e}$$

$$\mathbf{U}_1(t+1)=\mathbf{U}_1(t)+\left[\mathbf{v}(t+1)-\mathbf{U}_1(t)\boldsymbol{\psi}(t+1)\right]\mathbf{K}(t+1) \tag{14f}$$

$$\mathbf{P}(t+1)=\frac{1}{\mu^2}\left[\mathbf{P}(t)-\mathbf{P}(t)\boldsymbol{\psi}(t+1)\mathbf{K}(t+1)\right] \tag{14g}$$

$$\overline{\mathbf{H}}_{t+1}=\mu\overline{\mathbf{H}}_t+\overline{\mathbf{h}}(t+1)\left(\mathbf{y}_{t+1}^{-}\right)^T \text{ and } \mathbf{H}_{t+1}=\mu\mathbf{H}_t+\mathbf{y}_{t+1}^{+}\left(\mathbf{y}_{t+1}^{-}\right)^T \tag{14h}$$

Then $\mathbf{H}_{t+1}$ and $\overline{\mathbf{H}}_{t+1}$ can be updated.

The use of moving window technique implies the same procedure shown above has to be done twice for each new incoming data: after adding the new incoming data (up-dating), the oldest data has to be subtracted from the moving window (down-dating). The same formulas can be applied for updating by setting forgetting factor equal to one, meanwhile several sign changes in the last four formulas.

5. Recursive singular spectrum analysis (RSSA)

SSA is a novel non-parametric technique used in the analysis of time series based on multivariate statistics. This method was firstly applied to extract tendencies and harmonic components in meteorological and geophysical time series (Alonso *et al.* 2005). Except the extraction of tendency, SSA can be applied to eliminate noise effect, or to detect the singularities, e.g., to extract structural residual deformation (Loh *et al.* 2010). Basically, this method is capable of decomposing the original series into a summation of principal components, so that each component in this sum can be identified as a tendency, periodic components (stationary), non-stationary signal or noise. The SSA procedure starts from: (1) Embedding, (2) Singular Value Decomposition, (3) The plot of the singular values in descending order is called the singular spectrum and is essential in deciding the index from where to truncate the summation, (4) Grouping. To be able to apply SSA in online filtering of vibration measurements, an on-line version of the algorithm that describe the current structure at each time instant is required. For subspace-based algorithms, the moving window approach is adopted and the number of block rows *i'* is kept constant, meanwhile a new data point is appended as a new column to the moving window Hankel matrix

$$\mathbf{X}(N+1)=\left[\begin{array}{cccc|c} \mathrm{y}_{N\text{-}L'+1} & \mathrm{y}_{N\text{-}L'+2} & \cdots & \mathrm{y}_{N-i'+1} & \mathrm{y}_{N-i'+2} \\ \mathrm{y}_{N\text{-}L'+2} & \mathrm{y}_{N\text{-}L'+3} & \cdots & \mathrm{y}_{N-i'+2} & \mathrm{y}_{N-i'+3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathrm{y}_{N\text{-}L'+i'} & \mathrm{y}_{N\text{-}L'+i'+1} & \cdots & \mathrm{y}_{N} & \mathrm{y}_{N+1} \end{array}\right]$$
$$=\left[\mathrm{X}_{N-L'+i'} \quad \mathrm{X}_{N-L'+i'+1} \quad \cdots \quad \mathrm{X}_N \,\middle|\, \mathrm{X}_{N+1}\right]=\left[\mathbf{X}(N)\middle|\mathrm{X}_{N+1}\right] \tag{15}$$

where $\mathbf{X}(N) \in \Re^{il \times K}$, l is the number of sensors, i' is the sliding window vector order, i.e., number of block rows of the Hankel data matrix; L' is the length of moving window, $K = L' - i' + 1$ is the number of columns, with $K > i'$.

By applying SVD to the Hankel data matrix $\mathbf{X}(N+1)$, the left singular vectors $\mathbf{U}$ can be computed via Eigen-Decomposition (ED) of the covariance of sliding window vectors

$$\mathbf{C}_{SSA}(N+1) = \sum_{j=1}^{K} \mathrm{X}_{N-L'+i'+j} \mathrm{X}_{N-L'+i'+j}^{T} = \mathbf{X}(N)\mathbf{X}^{T}(N) + \mathrm{X}_{N+1}\mathrm{X}_{N+1}^{T} - \mathrm{X}_{N-L'+i'}\mathrm{X}_{N-L'+i'}^{T} \quad (16)$$

where $\mathbf{C}_{SSA}(N+1)$ is the covariance matrix of the sliding window vector $\boldsymbol{X}_{N-L'+i'+j}$. Since the eigenvectors of the covariance matrix $\mathbf{C}_{SSA}(N+1)$ correspond to the desired column subspace, this is actually a typical rank-two modification of the symmetric eigen-problem. The above-mentioned projection approximation subspace tracking (PAST) algorithm is suitable to be implement to rSSA, because it is able to update in a recursive fashion the dominant eigenvectors of the signal covariance $\mathbf{C}_{SSA}(N+1)$, i.e., the subspace of $\mathbf{X}(N+1)$.

The PAST is a fast dominant eigenvector updating algorithm. To conduct the rSSA the cost function can be defined as

$$V[\mathbf{W}(N+1)] = \sum_{k=N-L'+i'}^{N+1} \left\| \mathrm{X}_k - \mathbf{W}(N+1)\mathbf{W}^{H}(N+1)\mathrm{X}_k \right\|^2 \quad (17)$$

By introducing the same approximation, $\overline{\mathrm{h}'}(k) = \mathbf{W}^{H}(k-1)\mathrm{X}_k$, the original cost function is converted into a quadratic criterion

$$\overline{V}[\mathbf{W}(N+1)] = \sum_{k=N-L'+i'}^{N+1} \left\| \mathrm{X}_k - \mathbf{W}(N+1)\overline{\mathrm{h}'}(k) \right\|^2 \quad (18)$$

This became a typical optimization function in Least Square problems which can be minimized by

$$\mathbf{W}(N+1) = \mathbf{U}_1'(N+1) = \mathbf{C}_{SSA,X\bar{h}'}(N+1)\mathbf{C}_{SSA,\bar{h}'}^{-1}(N+1) \quad (19)$$

where $\mathbf{C}_{SSA,X\bar{h}'}$ is the covariance matrix formed by the sliding window vector $\boldsymbol{X}_k$ and $\overline{\mathrm{h}'}(k)$, and $\mathbf{C}_{SSA,\bar{h}'}$ is the covariance matrix formed by $\overline{\mathrm{h}'}(k)$ and $\overline{\mathrm{h}'}(k)^{T}$

$$\mathbf{C}_{SSA,X\bar{h}'}(N+1) = \mathbf{C}_{SSA,X\bar{h}'}(N) + \mathrm{X}_{N+1}\overline{\mathrm{h}}'^{T}_{N+1} - \mathrm{X}_{N-L'+i'}\overline{\mathrm{h}}'^{T}_{N-L'+i'} \quad (20a)$$

$$\mathbf{C}_{SSA,\bar{h}'}(N+1) = \mathbf{C}_{SSA,\bar{h}'}(N) + \overline{\mathrm{h}}'_{N+1}\overline{\mathrm{h}}'^{T}_{N+1} - \overline{\mathrm{h}}'_{N-L'+i'}\overline{\mathrm{h}}'^{T}_{N-L'+i'} \quad (20b)$$

When the matrix inversion lemma is applied to (19), the well-known RLS algorithm can be easily derived. The procedures are listed as follows:

(1) From the initial SVD the recursive algorithm can be initialized with $\mathbf{U}_1'(N)$, later $\mathbf{C}_{SSA,\bar{h}'}(N)$ and $\mathbf{P}'(N)$ can be computed

$$\mathbf{C}_{SSA,\bar{h}'}(N) = \left[\mathbf{U}_1'^T(N)\mathbf{X}(N)\right]\left[\mathbf{U}_1'^T(N)\mathbf{X}(N)\right]^T \tag{21a}$$

$$\mathbf{P}'(N) = [\mathbf{C}_{SSA,\bar{h}'}(N)]^{-1} \tag{21b}$$

(2) Given a new incoming sliding window vector $\boldsymbol{X}_{N+1}$, $\mathbf{U}_1'^T(N+1)$, $\mathbf{P}'(N+1)$ and $\overline{\mathrm{h}}'(N+1)$ can be updated using the following algorithm

$$\overline{\mathrm{h}}'(N+1)^* = \mathbf{U}_1'^T(N)\mathrm{X}_{N+1} \tag{22a}$$

$$\mathbf{K}'(N+1)^* = \frac{\overline{\mathrm{h}}'(N+1)^{*T}\mathbf{P}'(N)}{[\mu + \overline{\mathrm{h}}'(N+1)^{*T}\mathbf{P}'(N)\overline{\mathrm{h}}'(N+1)^*]} \tag{22b}$$

$$\mathbf{U}_1'(N+1)^* = \mathbf{U}_1'(N) + \left[\mathrm{X}_{N+1} - \mathbf{U}_1'(N)\overline{\mathrm{h}}'(N+1)^*\right]\mathbf{K}'(N+1)^* \tag{22c}$$

$$\mathbf{P}'(N+1)^* = \left(\frac{1}{\mu}\right)\left[\mathbf{P}'(N) - \mathbf{P}'(N)\overline{\mathrm{h}}'(N+1)^*\mathbf{K}'(N+1)^*\right] \tag{22d}$$

Therefore, after the column subspace $\mathbf{U}_1'(N+1)$ is updated, the filtered sliding window vector $\tilde{\mathrm{X}}_{N+1}$ can be computed by the projection

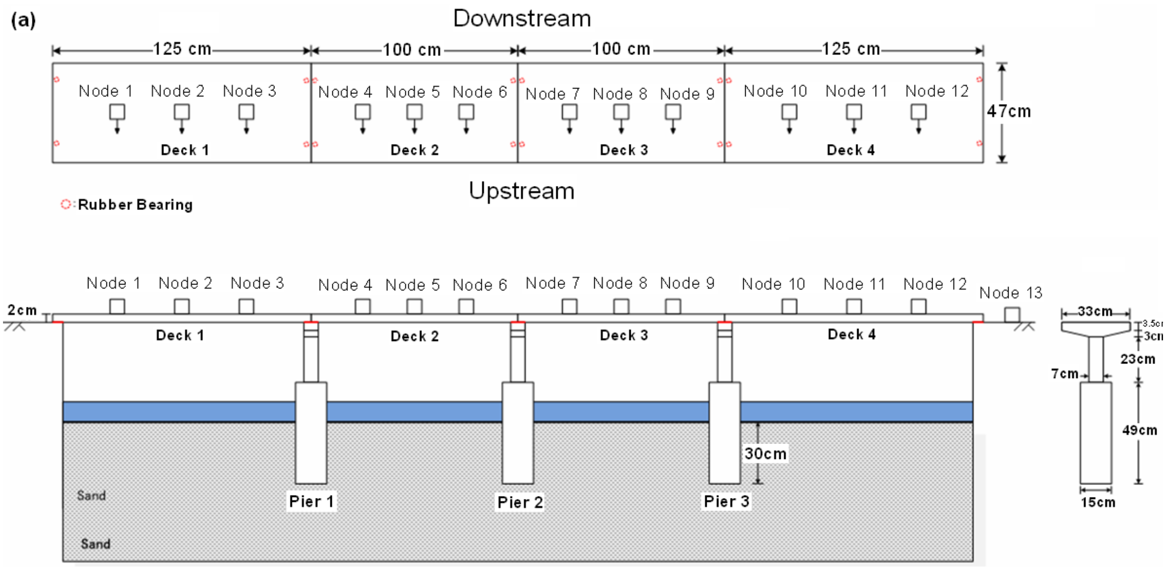
$$\tilde{\mathrm{X}}_{N+1} = \mathbf{U}_1'(N+1)\mathbf{U}_1'(N+1)^T\,\mathrm{X}_{N+1} \tag{23}$$

where $\tilde{\mathbf{x}}_{N+1}$ is the reconstructed data vector. Hence, for each new incoming data, a new vector column $\mathbf{X}_{N+1}$ is appended, the column subspace is updated to $\mathbf{U}_1'(N+1)$, and the reconstructed data vector $\tilde{\mathbf{x}}_{N+1}$ can be obtained by procedure shown above. Finally, elements of the same time instant in Hankel data matrix (in the anti-diagonal direction) are averaged to reconstruct the signal

$$\tilde{\mathbf{X}}(N+1) = \left[\begin{array}{cccc|c} \tilde{y}_{N-L'+1} & \tilde{y}_{N-L'+2} & \cdots & \tilde{y}_{N-i'+1} & \tilde{y}_{N-i'+2} \\ \tilde{y}_{N-L'+2} & \tilde{y}_{N-L'+3} & \cdots & \tilde{y}_{N-i'+2} & \tilde{y}_{N-i'+3} \\ \vdots & \vdots & & \vdots & \vdots \\ \tilde{y}_{N-L'+i'} & \tilde{y}_{N-L'+i'+1} & \cdots & \tilde{y}_{N} & \tilde{y}_{N+1} \end{array}\right] \tag{24}$$

6. Application of rSSA-RSSI-COV method to bridge scouring monitoring

From experimental study, consider a four span bridge with its steel decks simply supported on three cylindrical piers as shown in **Fig. 1**. The piers are buried with 30 cm of depth and confined by coarse sand. The goal of the experiment is to monitor the state of the bridge under continuous scouring, and to extract its vibrating features directly using output-only vibration measurements which allow for early warning of the pier settlement or failure of the bridge foundation as well as to locate damage. In order to create a local damage caused by scouring at a single bridge pier,



(a) A schematic diagram of the test flume which shows the location of sensors and the dimension of the test flume

(b) Model structure during the scouring test in the large flume

(c) Model structure after the scouring test

Fig. 1

brick wall was used in flume bed so as to enhance the bridge scouring at a single bridge pier. **Fig. 1(c)** shows the final state of the bridge pier after scouring test. A total of twelve VSE-15D velocity sensors (Tokyo Sokushin Corporation) as well as twelve acceleration sensors (AS-2000) were installed uniformly in the longitudinal direction and along the center line of the deck, as shown in **Fig. 1**. To analyze the vibration data from the measurement of the bridge under scouring test, the time-varying modal frequencies will be extracted along time axis to identify the abnormal change of the bridge system natural frequencies. Not only the rSSA-SSI-COV algorithm will be the main tool to carry out this analysis, but also the RSSI-COV algorithm is applied for comparison purpose.

In this study all the acceleration data from twelve sensors was used simultaneously. The acceleration data was filtered in the field by an analog band-pass filter having its plateau zone in

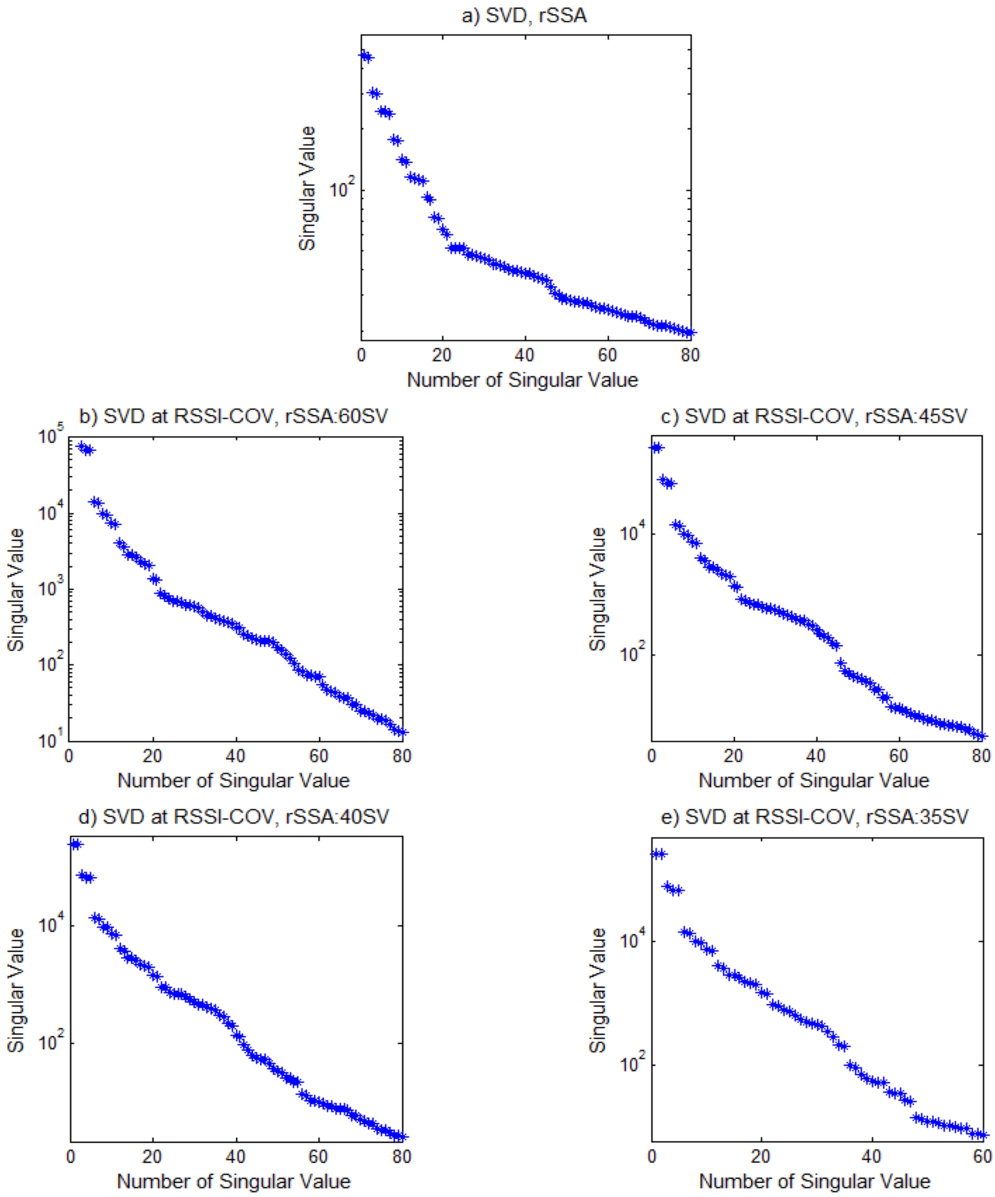


Fig. 2 (a) Singular spectrum for different choices of singular values in rSSA, (b)~(e): Distribution of singular spectrum in SSI-COV analysis by using different number of singular spectrum from rSSA

Table 1 Model parameters for rSSA-SSI-COV and RSSI-COV analysis from the 2011/03/29 test

Parameter	rSSA-SSI-COV		RSSI-COV used alone
	rSSA	RSSI-COV	
Window length	L' = 3000 points	L = 5000 points	L = 5000 points
Block rows	i' = 100	i = 80	i = 100
Order	First 45 singular values	First 44 singular values	First 46 singular values

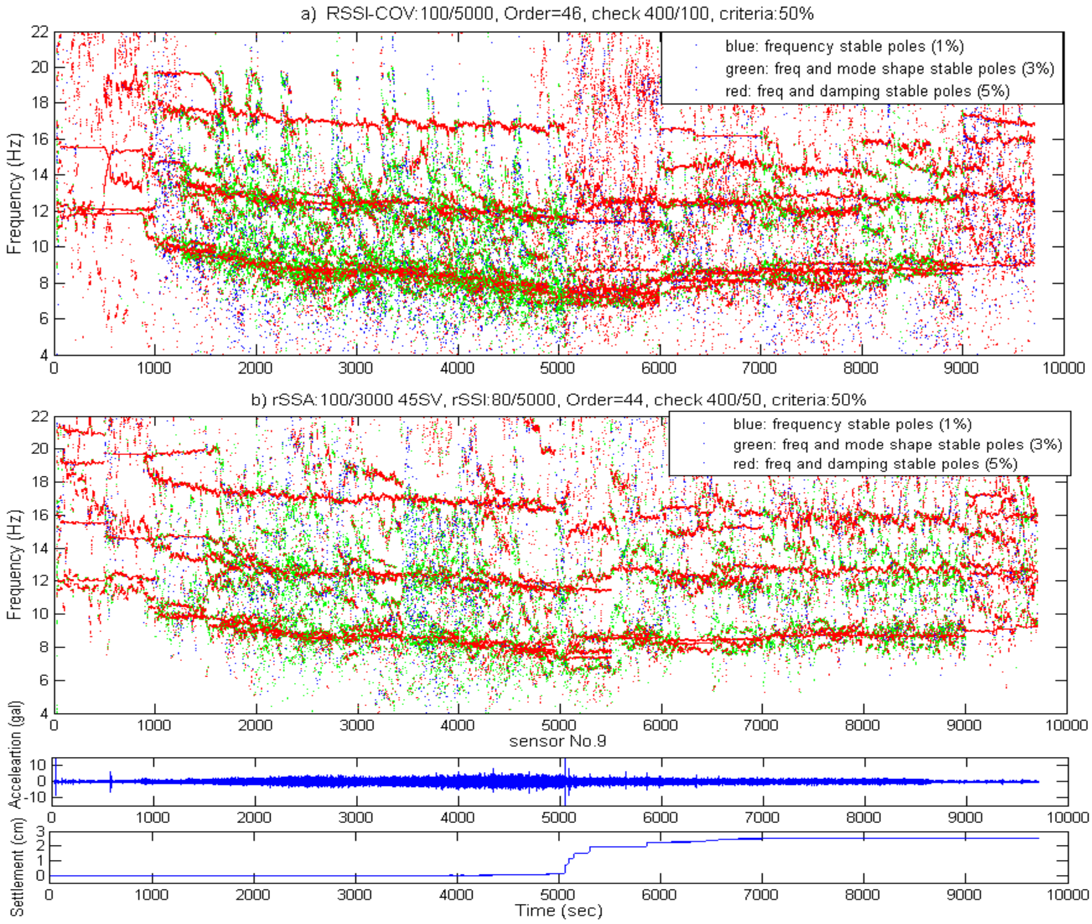


Fig. 3 Evolution of bridge modal frequencies traced by both (a) RSSI-COV and (b) rSSA-SSI-COV, applying stability criterion, test conducted in 2011/03/29 measured by accelerometers

the frequency response function between 0.02 Hz and 50 Hz. To apply the SSA the order must be determined in advance. The singular spectrum was first constructed using the initial data set. **Fig. 2(a)** shows the distribution of the singular spectrum from the test data of 2011/03/29. Using different order chosen from SSA, the singular spectrum constructed in SSI-COV was generated. It is proved that the subspace order selected from SSA which shows the appearance of a jump in the distribution of singular spectrum of SSI-COV can lead to a better choice of for the system order. **Fig. 2(b)** to **Fig. 2(e)** show the constructed singular spectrum from SSI-COV by using the reconstructed data using different singular value of SSA (from Fig. 2(a)). From these figures it is observed that a clear jump appears at 44 SV in the singular value distribution of SSI-COV analysis by using 45 SV chosen from SSA (as shown in Fig. 2(c)). However, the jump becomes more clear with less SV be chosen from SSA. To be conservative and try to identify all the excited modes, the order shown in Fig. 2(c) is chosen. The model parameters for rSSA-SSI-COV are shown in **Table 1**.

Fig. 3 shows a comparison between the outcome of using RSSI-COV and rSSA-SSI-COV with the same stability criteria. Evidently the addition of rSSA to determine the system order before adopting RSSI-COV can enhance the tracking capability and stability. The pattern shown in the time-frequency plot indicated that the water head arrivals at about 860 seconds, after that the modal frequencies slowly decrease until the occurrence of the first settlement of the bridge pier which occurred at 5057 seconds.

A close picture was taken to the identified traces between 4500 and 5500 seconds, as shown in **Fig. 4**. The 4th mode appears at about 4650 seconds and decreases rapidly at about 4950 seconds. At the same moment in which the traces of the 1st and 2nd mode modal frequency are almost completely lost. This indicates that there is a very unstable dynamic behaviour before the 1st settlement which occurs at 5057 seconds. This phenomenon together with the fast decrement of the 4th mode constitutes a good indicator to identify the imminent bridge settlement. During the time period with successive settlements between 5057 and about 6000 seconds, the traces of modal frequencies are also very diversified and the reduction in 3rd modal frequency is evident. However, after 6000 seconds the modal frequencies slowly increase because the decks are getting stuck each other.

A close look on the identified time-varying system frequencies, as also shown in Fig. 4, there are about two identified frequency traces for 3rd mode, three traces for 2nd mode and about six traces for 1st mode. All are very close one to another. These closely-spaced frequencies also indicate the time varying characteristic of the bridge natural frequencies under the scouring test.

7. Damage location indicator: mode shape slope ratio

The curvature of mode shape has been widely used to figure out the damage location by many researches. Larger curvature can be identified at node where the loss of stiffness loss occurred. Since the mode shape is a relative quantity which can be scaled arbitrary, however, the mode shape curvature is not independent of the scaling criteria and consequently, therefore the

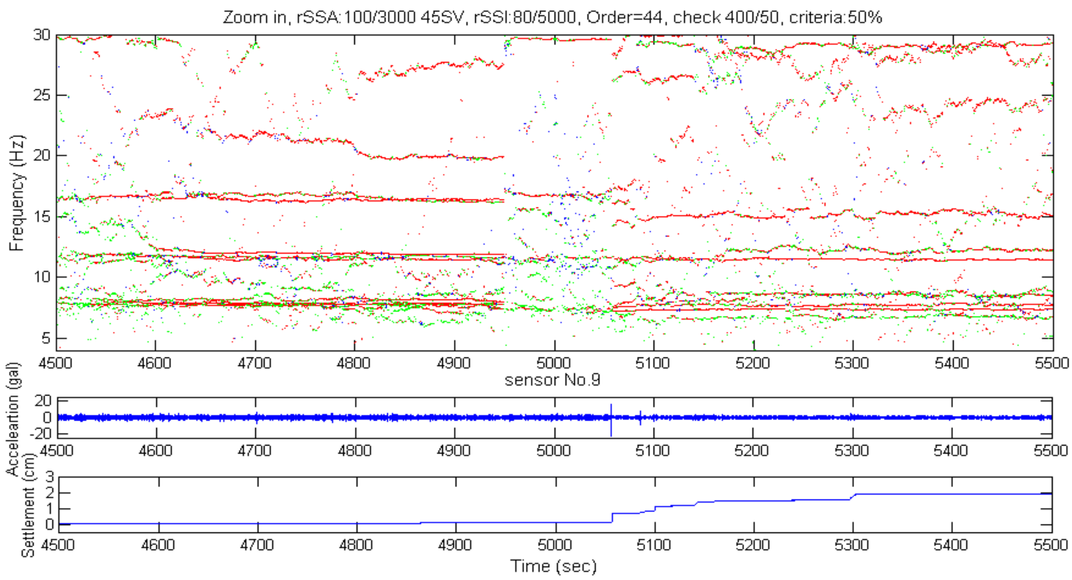


Fig. 4 Zoom in from Fig. 3(b) from time scale of 4500 sec to 5500 sec

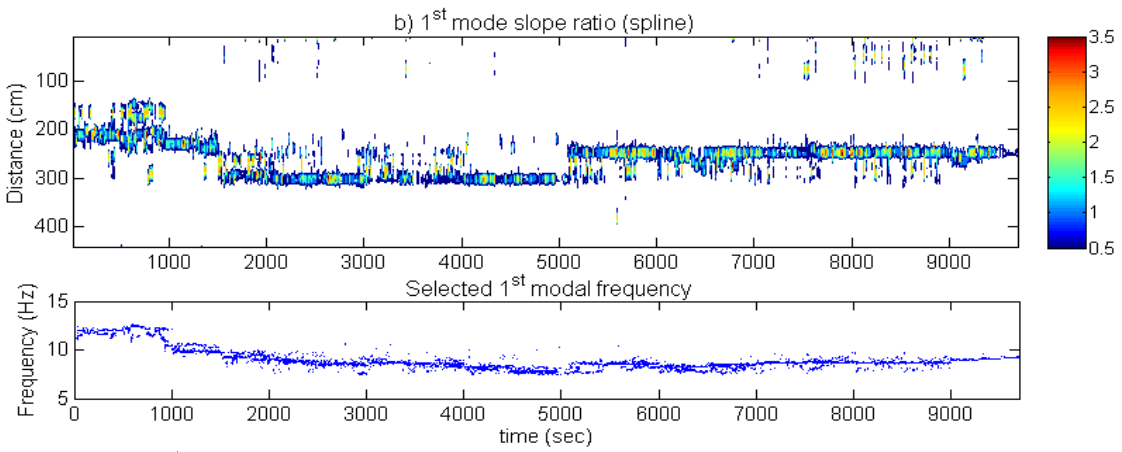


Fig. 5(a) Plot of 1st mode shape slope ratio and the identified time-varying system natural frequency from the 2011/03/29 test

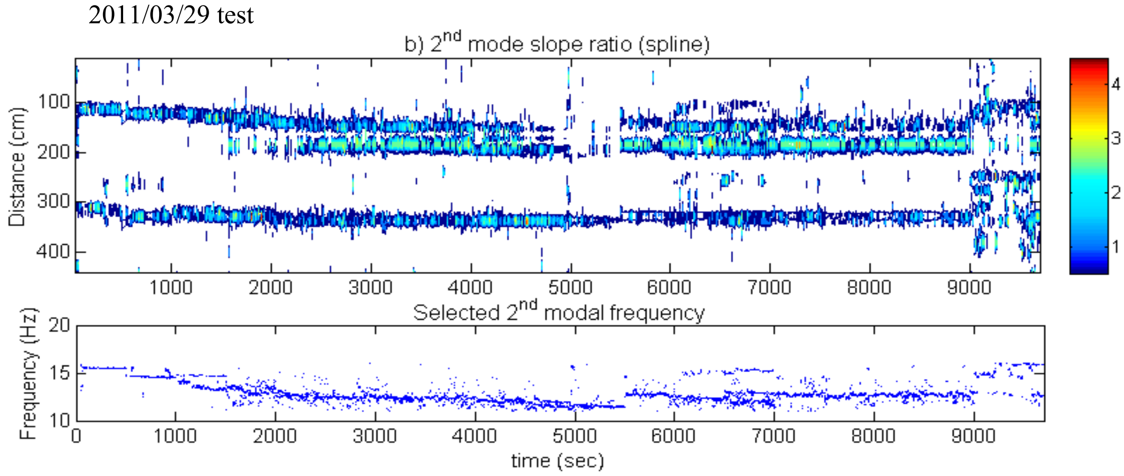


Fig. 5(b) Plot of the 2st mode shape slope ratio and the identified time-varying system natural frequency from the 2011/03/29 test

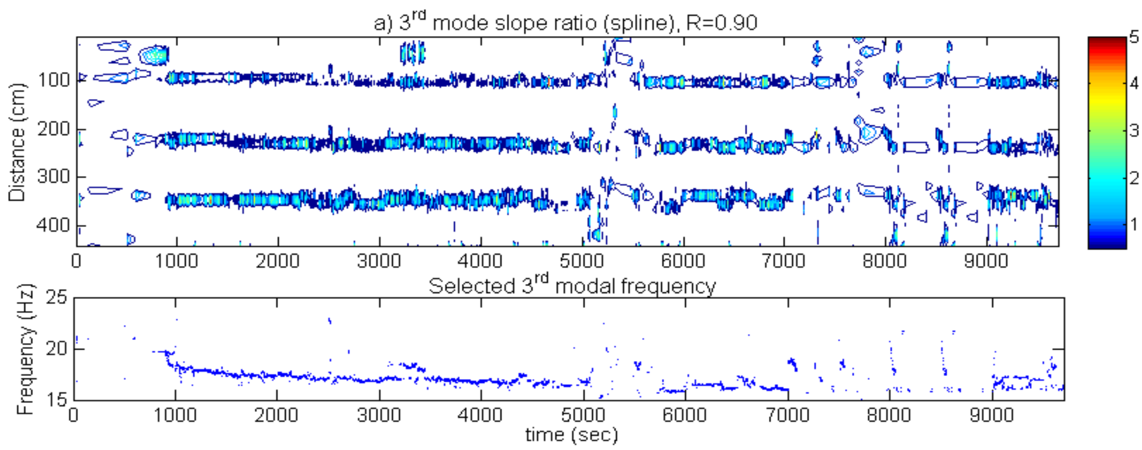


Fig. 5 (c) Plot of the 3rd mode shape slope ratio and the identified system natural frequency from the 2011/03/29 test identification of damage location depends on how one scale the mode shape.

Taking into account of this fact, the curvature of the mode shape can be modified to a quantity which is independent of scaling. To cancel out arbitrary scaling, ratio between two consecutive slopes can be used to replace the rate of change of the slope of mode shape. Moreover, sign of the curvature indicates the concavity, and from the identified mode shapes the sign changes introduced in the concavity due to imperfection in the shape could make it difficult to identify the damage location. To avoid all these inconveniences, a new damage indicator defined as the mode shape slope ratio is presented:

$$\begin{cases} \text{If } m_i/m_{i+1} > 0 \text{ and } m_i > m_{i+1} \longrightarrow \text{slope ratio}(i) = m_i/m_{i+1} \\ \text{If } m_i/m_{i+1} > 0 \text{ and } m_i < m_{i+1} \longrightarrow \text{slope ratio}(i) = m_{i+1}/m_i \\ \text{If } m_i/m_{i+1} < 0 \longrightarrow \text{slope ratio}(i) = \{\text{slope ratio}(i) + \text{slope ratio}(i+1)\}/2 \end{cases}$$

where m_i is the slope of the i^{th} discrete segment of the smoothed interpolated mode shape.

By defining the slope ratio in this way, the sign problem can be avoided, and the resultant slope ratio will only reflect how big the slope change at a given point. However, at the peak point of the mode shape where the slope sign changes, the 3rd criterion can be applied by taking an average of the adjacent slope ratios. Finally, to avoid the disproportionate increment in slope ratio comparing to the others, when slope in the divisor is near zero, a base 10 logarithm can be applied to the slope ratio. For the implementation, the computed mode shapes can be smoothed by curve fitting and interpolated with a spline function, and then sampled at 52 points to obtain 50 slope ratios along the bridge.

Based on the identified mode shapes (smoothed) from online identification, **Figs. 5(a)** and **(b)** show the mode shape slope change of the 1st and 2nd modes. The y-axis of Fig. 5 indicates the spatial distance along the bridge deck, and the x-axis is the time.

From this figure it is observed that the zone with higher slope ratio become wider after 1000 seconds, indicating that the system begins to be changed. Initially the peak is located at the center of the bridge which is expected for a 1st mode shape; while the scouring depth become deeper and deeper, the peak moves from the location at 300 cm to 250 cm, specially between 5000 and 5500 seconds as that shown in Fig.5(a). The moment which corresponds to the imminent pier settlemen, precisely the pier 3 is located at 325 cm.

The 2nd mode shape slope ratio is shown in **Fig. 5(b)**. Two peaks exist in the 2nd mode shape as expected. Although the 2nd peak is located at 325 cm (pier 3 location), but it does not change at all along the time history, otherwise, the second peak located at 200 cm which is almost disappeared 100 cm at about 5500 seconds. Therefore, the 2nd mode shape is not appropriate to identify the damage location in the bridge, but the drastic change of the 2st peak position is also an indicator of imminent pier settlement. Same situation was also observed for 3rd mode, as shown in **Fig. 5(c)**.

8. Conclusions

The on-line system parameter estimation technique from output-only measurements is developed through using Covariance Driven Recursive Stochastic Subspace identification (RSSI-COV). To update the SVD through recursive analysis, Extended Instrumental Variable version of Projection Approximation Subspace Tracking algorithm (EIV-PAST) was adapted for this purpose. In order to have a stable result on identification, the recursive Singular Spectrum

Analysis (rSSA) algorithm which used the PAST concept is proposed as a data pre-processing filter before conducting the RSSI-COV for identification. To verify the effectiveness on using the rSSA as a data pre-processing tool, both RSSI-COV and rSSA-SSI-COV are applied to the experimental study of a bridge during scouring monitoring. Several conclusions were obtained from this study:

(1) In applying the RSSI-COV for online system identification, the tracking stability increases with the longer window length. But this does not affect the computation time, however, the longer the window length is used the more delay time will be encountered to detect system change is observed.
(2) The pre-subspace filtering using recursive SSA can enhance the tracking stability and allows tracking the time-varying modal information with less system order. From the experimental study on bridge scouring test, to observe the change of system natural frequencies, the rSSA-SSI-COV algorithm proves to be a very effective way on monitoring the structural system with time-varying modal characteristics.
(3) The proposed mode shapes slope change which was extracted recursively can be used not only identify the occurrence time of damage but also can identify the damage location.

Acknowledgements

This research was supported by National Science Council, Taiwan, under grant number NSC 98- 2625-M-002-018-MY3.

References

Alonso, F.J., Del Castillo, J.M. and Pintado, P. (2005), "Application of singular spectrum analysis to the smoothing of raw kinematic signals", *J. Biomech*., **38**(5), 1085-1092.
Bart, P. and Guido, D.R. (1999), "Reference-based stochastic subspace identification for output-only modal analysis", *Mech. Syst. Signal Pr*., **13**(6) 855-878.
Bart, P. (2000), *System Identification and Damage Detection in Civil Engineering*, Ph.D. Dissertation, Katholieke Universiteit, Leuven, December.
Caicedo, J.M., Dyke, S.J. and Johnson, E.A. (2004), "Natural excitation technique and eigensystem realization algorithm for phase I of the IASC-ASCE benchmark problem: simulated data", *J. Eng. Mech.-ASCE*, **130**(1), 49-60.
Goethals, I., Mevel, L., Benveniste, A. and De Moor, B. (2004),"Recursive output only subspace identification for in-flight flutter monitoring", *Proceedings of the 22nd International Modal Analysis Conference*, Dearborn, Michigan.
Golub, G.H. and Van Loan, C.F. (1996), *Matrix computations*, Johns Hopkins University Press.
Loh, C.H., Mao, C.H., Chao, S.H. and Weng, J.H. (2011), "System identification and damage evaluation of degrading hysteresis of reinforced concrete frames", *Earthq. Eng. Struct. D*., **40**(6), 623-640.
Loh, C.H., Weng, J.H., Liu, Y.C., Lin, P.Y. and Huang, S.K. (2011), "Structural damage diagnosis based on on-line recursive stochastic subspace identification", *Smart Mater. Struct*., **20**(5), 055004 (10pp).
Michèle B., Albert, B., Maurice, G., Luc, H., Laurent, M. and Herman, V.D.A. (2001), "Output-only subspace-based structural identification: From theory to industrial testing practice", *J. Dyn. Syst. Measure.Control*, **123**(4), 668-676.
Söderström, T. and Stoica, P. (1989), *System identification*, Prentice-Hall International.
Van Overschee, P. and De Moor, B. (1996), *Subspace Identification for Linear Systems: theory -*

Implementation – Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands.
Weng, J.H., Loh, C.H., Lynch, J.P., Lu, K.C., Linn, P.Y. and Wang, Y. (2008), "Output-only modal identification of a cable-stayed bridge using wireless monitoring systems", *J. Eng. Struct*., **30**(2), 1802-1830.
Yang, B. (1995), "Projection approximation subspace tracking", *IEEE T. Signal Process*., **43**(1), 95-107.