

基于 AR 模型和主成分分析的损伤识别方法*

吴 森^{1,2}, 韦灼彬², 王绍忠¹, 王 斌¹, 李 扬¹

(1. 海军驻沈阳地区发动机专业军事代表室 沈阳, 110043) (2. 海军工程大学后勤指挥与工程系 天津, 300450)

摘要 基于结构加速度时间序列提出了一种新的损伤识别方法。首先, 获取结构在无损伤状态下的加速度数据并进行分段, 以各段数据的 AR(auto-regressive)模型系数向量作为结构的参考状态样本, 将未知状态的加速度 AR 模型系数向量分别加入参考状态样本中, 构成多个原始数据矩阵; 其次, 对多个原始数据矩阵分别进行主成分分析得到前两阶主成分, 并建立相应的椭圆控制图, 以前两阶主成分在控制椭圆中的分布情况来判别结构是否存在损伤; 最后, 以一钢框架结构试验为例识别结构的两种损伤模式。结果显示, 该方法能够准确、直观地识别结构是否存在损伤, 相对于马氏距离判别法具有更强的稳定性。

关键词 AR 模型; 主成分分析; 控制椭圆; 马氏距离; 损伤识别

中图分类号 TH165.3

引 言

基于结构振动特性的损伤识别方法是结构健康监测领域研究的热点, 模态参数是反映结构振动特性的主要参数, 但基于模态参数的损伤识别方法在实际工程中并不能得到很好的应用, 主要原因为: 模态频率测量的准确性虽高, 但对结构损伤并不敏感; 模态振型实际测量不准, 且很难获取高阶模态振型。结构加速度的测量相对于结构模态参数十分方便快捷, 基于结构加速度测量的损伤识别方法更能满足结构健康监测的实时性要求。从传感器获得的加速度数据不能直接反映结构的损伤信息, 需要通过某种变换从加速度测量数据中提取反映结构损伤信息的损伤指标, 这类变换方法有很多^[1-6]。

笔者引入 AR 模型对结构加速度数据进行拟合, 采用拟合系数构建结构损伤特征向量。以结构无损状态损伤特征向量为参考样本, 结构损伤状态损伤特征向量为待检样本, 把待检样本逐个加入到参考样本中构建多个原始数据矩阵。用主成分分析法提取原始数据矩阵前两个主成分并构造相应的控制椭圆, 以前两阶主成分散点图在控制椭圆的分布状况来判断结构是否存在损伤, 并以一钢框架结构试验验证了本研究方法的有效性。

1 自回归模型的建立

AR 模型的参数估计是线性估计, 相对于 AR-

MA 模型参数估计具有计算简单、速度快的特点。一般情况下, 用 ARMA 模型参数对结构进行损伤识别也是采用 ARMA 模型的 AR 部分的参数^[7]; 所以, 笔者用 AR 模型对结构加速度数据进行拟合。

对时间序列 $\{X(t)\} (t=1, 2, \dots, n)$ 进行数据标准化处理

$$\tilde{X}(t) = (X(t) - \tilde{\mu}) / \tilde{\sigma} \quad (1)$$

其中: $\tilde{\mu}, \tilde{\sigma}$ 分别为 $\{X(t)\}$ 的期望和方差的估计。

对 $\{\tilde{X}(t)\}$ 建立 AR 模型如下

$$\tilde{X}(t) = \sum_{k=1}^p \varphi_k \tilde{X}(t-k) + \epsilon(t) \quad (2)$$

其中: φ_k 为第 k 阶 AR 模型的自回归系数, 为无量纲; p 为 AR 模型阶次; $\epsilon(t)$ 为残余项。

笔者采用 Burg 法对自回归系数 φ_k 进行计算, 由 $\{\tilde{X}(t)\}$ 的偏自相关函数结合 AIC(an information criterion)准则来确定 AR 模型阶次 p , 其中 AIC 准则函数如下

$$AIC(p) = N \ln \sigma_i^2 + 2p \quad (3)$$

其中: N 为时序长度; σ_i^2 为 AR 模型拟合残差的方差。

2 结构损伤统计模式识别

由于土木结构庞大复杂, 所处环境恶劣, 使得传感器在实际采集过程中不可避免地受到周围环境各种干扰的影响, 而这些干扰往往会对结构损伤造成误判, 阻碍了损伤识别方法在实际工程中的应用; 所

* 收稿日期: 2011-01-12; 修改稿收到日期: 2011-04-03

以,有必要在损伤识别中引入统计理论对损伤识别结果进行统计分析,用对结构损伤的统计判别来代替传统的对结构损伤的确定性判别^[8]。

2.1 主成分数据缩减

结构损伤统计模式识别可分为5个过程:运行评价、数据采集、特征提取、数据压缩以及统计模型建立^[9]。求AR模型系数即特征提取过程,数据压缩通常采用主成分分析,其目的是用原始变量重新组合成少数几个互相无关的综合变量,且这少数几个新的变量几乎包含了原始变量的全部信息,从而可使用这几个新的变量代替原变量分析问题,达到了数据压缩的目的。其原理简述如下^[10]:

设有 n 个样本,每个样本有 p 个观测变量,分别记为 $\varphi_1, \varphi_2, \dots, \varphi_n$,则这 n 个样本可构成原始数据资料矩阵 $\varphi \in R^{n \times p}$

$$\varphi = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1p} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{np} \end{bmatrix} \quad (4)$$

其中: $\varphi_{ij} (i \in 1, 2, \dots, n; j \in 1, 2, \dots, p)$ 为第 i 个样本的第 j 个变量。

求 φ 的协方差矩阵 Σ 为

$$\Sigma = E(\varphi - E\varphi)(\varphi - E\varphi)^T \quad (5)$$

其中: Σ 的特征根为 $\lambda_1, \lambda_2, \dots, \lambda_p$,且 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ 。

相应的单位特征向量为 u_1, u_2, \dots, u_p ,令

$$U = (u_1, u_2, \dots, u_p) = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1p} \\ u_{21} & u_{22} & \cdots & u_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pp} \end{bmatrix} \quad (6)$$

φ 的第 i 主成分可表示为

$$Y_i = u_{1i} \varphi_1 + u_{2i} \varphi_2 + \cdots + u_{pi} \varphi_p \quad (7)$$

通常不能得到总体的协方差矩阵 Σ ,可用样本协方差矩阵估计值 S 来代替, S 的估计值计算如下

$$S = (s_{ij})_{p \times p} = \frac{1}{n-1} \sum_{k=1}^n (\varphi_k - \bar{\varphi})^T (\varphi_k - \bar{\varphi}) \quad (8)$$

其中: $\bar{\varphi} = (\bar{\varphi}_1, \bar{\varphi}_2, \dots, \bar{\varphi}_p)^T$; $\bar{\varphi}_j = \frac{1}{n} \sum_{i=1}^n \varphi_{ij}$; $s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (\varphi_{ik} - \bar{\varphi}_i)(\varphi_{jk} - \bar{\varphi}_j)$; $i, j = 1, 2, \dots, p$ 。

用累积贡献率来确定主成分个数 k ,如前 k 个主成分的方差和在全部分方差所占比重 $\sum_{i=1}^k \lambda_i / \sum_{i=1}^p \lambda_i \geq$

85%,说明前 k 个主成分 Y_1, Y_2, \dots, Y_k 包含了原始数据的绝大部分信息。

2.2 椭圆控制图

多元控制图来自于多元统计分析中的过程控制 and 产品质量控制理论,目的是识别引起系统变动的特殊原因。在结构损伤识别中,可以把结构损伤特征向量(如AR,ARMA模型系数)作为结构状态的观测变量,对这些变量进行多次观测获得结构一个过程的观察数据,用多元控制理论来分析数据,从而推断结构是否存在损伤。

椭圆控制图可直观地构造过程为两个变量的控制域,虽然大多数过程控制的观测变量都大于两个,但通过对原始数据矩阵进行主成分分析可得包含原始数据主要信息的前两阶主成分,因而可用椭圆控制图对前两阶主成分进行离群值分离。

设 Y_1, Y_2 为原始数据矩阵中心化处理后的前两个主成分,其相对应的样本方差为 λ_1 和 λ_2 ,则对这两个主成分的 $1-\alpha$ 置信椭圆满足式(9)

$$Y_1^2/\lambda_1 + Y_2^2/\lambda_2 \leq \chi_2^2(\alpha) \quad (9)$$

其中: α 为置信水平,通常取为0.05。

2.3 损伤识别方法

AR(p)模型有 $\varphi_1, \varphi_2, \dots, \varphi_p, \sigma_i^2$ 共 $p+1$ 个参数,由于自回归参数 $\varphi_1, \varphi_2, \dots, \varphi_p$ 反映了系统的固有特性,因此笔者用 $\varphi_1, \varphi_2, \dots, \varphi_p$ 这 p 个参数构建结构状态的特征向量 $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_p]$ 。损伤识别方法的具体步骤如下:

1) 在结构无损伤状态下测量结构加速度数据并把数据分成 n 段;

2) 确定AR模型阶次 p ,并分别计算各段数据AR模型系数,得到结构无损状态 n 个损伤特征向量样本 $\{\varphi_1^u, \varphi_2^u, \dots, \varphi_n^u\}$;

3) 测量结构在未知状态下的加速度数据,根据步骤1、步骤2得到结构在未知状态下的 m 个损伤特征向量样本 $\{\varphi_1^d, \varphi_2^d, \dots, \varphi_m^d\}$;

4) 把结构未知状态 m 个损伤特征向量分别添加到结构无损状态 n 个特征向量样本中,构造 m 个原始数据矩阵 $\varphi_i = [\varphi_1^u, \varphi_2^u, \dots, \varphi_n^u, \varphi_i^d]^T (i = 1, 2, \dots, m)$,假设这 m 个原始数据矩阵中的 $n+1$ 个样本独立分布,且分别服从 $N_p(\mu_i, \Sigma_i) (i = 1, 2, \dots, m)$;

5) 对 φ_i 进行中心化处理得到新的 m 个矩阵 $\tilde{\varphi}_i (i = 1, 2, \dots, m)$,如式(10)所示

$$\tilde{\varphi}_i = \varphi_i - (\varphi_1^u + \varphi_2^u + \cdots + \varphi_n^u + \varphi_i^d)/(n+1) \quad (10)$$

分别求取 $\tilde{\varphi}_i$ 的前两阶主成分

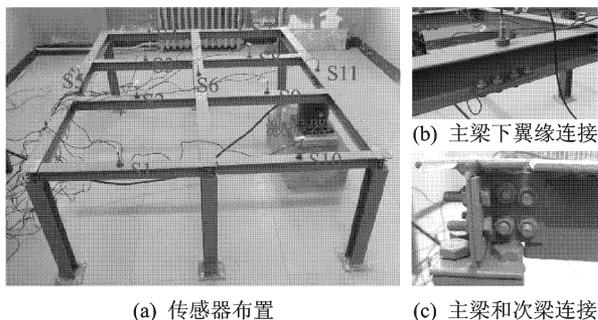
$$Y_1^i = [y_{11}^i, y_{21}^i, \dots, y_{(n+1)1}^i]^T, Y_2^i = [y_{11}^i, y_{22}^i, \dots, y_{(n+1)2}^i]^T$$

6) 分别构造 m 个控制椭圆, 观测并记录数据对 $(y_{(n+1)1}^i, y_{(n+1)2}^i)$ 离群的个数, 如果绝大多数 $(y_{(n+1)1}^i, y_{(n+1)2}^i)$ 都在相对应的椭圆控制域外, 则判断结构存在损伤。

3 试验研究

本研究对象为一钢框架结构, 其模型整体尺寸为 $1\ 500\ \text{mm} \times 1\ 150\ \text{mm} \times 564\ \text{mm}$, 由 3 根主梁、8 根次梁和 6 根柱组成, 主梁、次梁及柱之间由螺栓连接, 整个结构由螺栓固定在地面上, 如图 1 所示。

试验测量用的是 DH5920N 动态测试系统, 传感器为 DH131E 加速度传感器, 用 JZK-10 激振器对结构进行激励, 传感器布置和激励部位如图 1(a) 所示。激励信号为随机激励, 采样频率为 $2\ \text{kHz}$ 。在激振器开始激振 5 s 后进行采样, 采样时间为 16 s, 以每段 3 000 个数据把采样样本分成 30 段, 数据重合率为 80%。



(a) 传感器布置

(c) 主梁和次梁连接

图 1 钢框架结构及其损伤部位

3.1 结构损伤的模拟

如图 1 所示, 钢框架结构主梁下翼缘中间部位人为断开, 由两块面板通过螺栓进行等强度连接, 主梁和次梁也是由螺栓和角铁进行连接。采用两种模式对损伤进行模拟: a. 去除框架结构左边主梁(传感器 S5 位置)下翼缘一边连接面板, 模拟结构断裂损伤, 如图 1(b) 所示; b. 松开主梁和次梁(传感器 S4 位置)连接部位的部分螺栓, 模拟结构连接松动损伤, 如图 1(c) 所示。

3.2 马氏距离判别的不稳定性

基于 AR 模型系数的损伤识别方法中, 在特征提取后通常采用马氏距离来判别结构的损伤状态。马氏距离具有许多优点, 如排除了各变量之间相关

性的干扰、不受变量量纲的影响等。它的缺点是夸大了变化微小的变量的作用, 计算具有不稳定性, 这种不稳定性在变量多的情况下表现更加明显。下面以一实例说明马氏距离的这一特性。

在结构无损状态时, 根据本研究方法以传感器 S5 采集的数据为例, 第 1 次采样计算得到 30 个结构损伤特征向量 $\{\varphi_1^{n1}, \varphi_2^{n1}, \dots, \varphi_{30}^{n1}\}$, 上标 $U1$ 表示结构无损状态第 1 次采样。以这 30 个特征向量为训练样本, 则参考总体 $G_R = [\varphi_1^{n1}, \varphi_2^{n1}, \dots, \varphi_{30}^{n1}]^T$ 。以第 2 次采样计算得到的 30 个特征向量 $\{\varphi_1^{n2}, \varphi_2^{n2}, \dots, \varphi_{30}^{n2}\}$ 作为检验样本, 分别计算这 60 个结构损伤特征向量到参考总体的马氏距离, 计算公式如下

$$DSF^n = (\varphi_i - \mu_R)^T C_R^{-1} (\varphi_i - \mu_R) \quad (i = 1, 2, \dots, 60) \quad (11)$$

均值向量 μ_R 和总体协方差矩阵 C_R 由参考总体 G_R 估计得到, φ_i 为损伤特征向量, 其中包括结构第 1 次和第 2 次测量的所有特征向量。结构损伤敏感特征值 DSF^n 计算结构如图 2 所示。

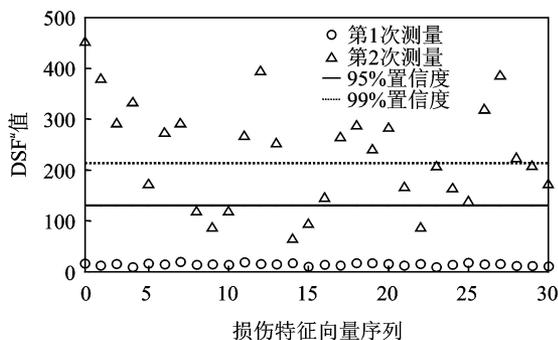


图 2 马氏距离计算结果

由于 $(\varphi_i - \mu_R)^T C_R^{-1} (\varphi_i - \mu_R)$ 服从 $\frac{(n^2-1)p}{n(n-p)} F_{p, n-p}$ 分布, 则计算 DSF^n 的 95% 置信度的控制上限为 $\frac{(n^2-1)p}{n(n-p)} F_{p, n-p}(0.05)$ 。由图 2 可知, 训练样本到总体的马氏距离都在控制域以内, 而同一状态获得的检验样本与参考总体的马氏距离大多在控制域以外, 即使扩大控制上限到 99% 置信度, 第 2 次测量大多数样本仍然被拒绝, 即按照图 2 判断, 检验样本应来自结构损伤状态, 这与事实并不相符。

按照笔者介绍的椭圆控制图法, 对检验样本进行重新判断, 其中部分椭圆控制图如图 3 所示。PC1, PC2 分别表示第 1、第 2 主成分, 图 3 中圆点数据来自于结构参考样本, 三角点数据来自结构待检样本。由图 3 可知, 检验样本中的绝大多数样本和参考样本来自于结构同一状态, 30 个控制椭圆图中有 6 个产生了离群点, 分别为第 1, 11, 13, 21, 25 和 26 个控制椭圆。这些点包含在 99% 置信度的控制椭圆

之内,属于控制椭圆的可控范围,其部分图形如图 4

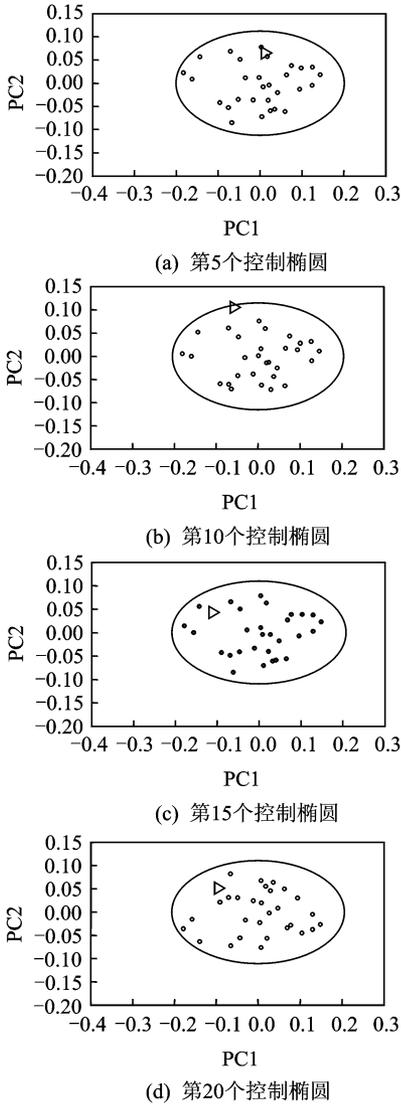


图 3 检验结果部分 95%置信控制椭圆

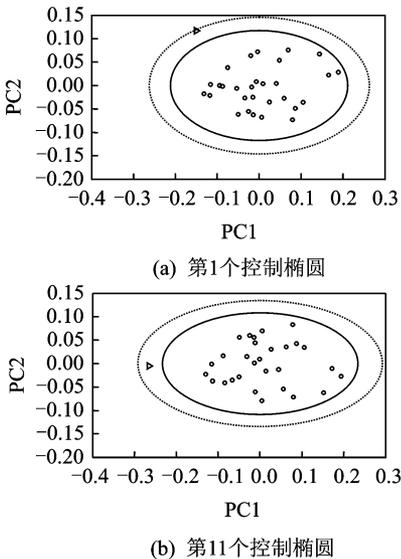


图 4 部分失控椭圆

所示。其中:实椭圆为 95%置信控制椭圆;虚椭圆为 99%置信控制椭圆。

从实例分析可知,马氏距离虽然是距离判别的常用方法,但其计算的不稳定性会影响判断结果。对资料数据进行主成分分析提取其前两个主成分,不但能够显示结构状态的大部分信息,而且避免了变量多维性对计算结果的影响,具有较强的稳定性。

3.3 损伤识别结果

用本研究方法对结构两种损伤进行识别。以传感器 S5 采集数据进行分析,分别对每段数据进行 AR 模型拟合,计算 AR 模型系数。其中 AR 模型阶数为 17,由偏自相关函数的截尾特性和 AIC 准则得到。

对于结构两种损伤状态分别可以得到 30 个待检数据样本 $\{\varphi_1^{d1}, \varphi_2^{d1}, \dots, \varphi_{30}^{d1}\}$ 和 $\{\varphi_1^{d2}, \varphi_2^{d2}, \dots, \varphi_{30}^{d2}\}$ 。根据损伤识别步骤 4,分别构建两组原始数据矩阵 $G_i^{T1} = [\varphi_1^{u1}, \varphi_2^{u1}, \dots, \varphi_{30}^{u1}, \varphi_i^{d1}]^T \in R^{31 \times 17}$ 和 $G_i^{T2} = [\varphi_1^{u1}, \varphi_2^{u1}, \dots, \varphi_{30}^{u1}, \varphi_i^{d2}]^T \in R^{31 \times 17}$,其中 $i \in 1, 2, \dots, 30$ 。由 G_i^{T1} 和 G_i^{T2} 可以建立 30 个椭圆控制图,所有来自损伤状态的数据都在控制椭圆之外,其中部分结果如图 5 所示。其中:三角形为离群点;实椭圆为 95%置信控制椭圆;虚椭圆为 99%置信控制椭圆。由图 5 可知,结构参考样本数据的前两个主成分的散点图都在 95%置信度的椭圆控制域以内,而结构损伤状态样本数据前两个主成分的散点图都在椭圆控制域以外,证明两种状态下的结构加速度 AR 模型系数存在显著差异,判别结构存在损伤。

引用假设检验的概念对控制椭圆识别结果进行整体描述:原假设 H_0 为两种状态下结构加速度 AR 模型系数不存在显著差异;备择假设 H_1 为两种状态下结构加速度 AR 模型系数存在显著性差异。对无损损伤第 2 次测量以及两种损伤的假设检验结果如表 1 所示。

表 1 假设检验结果(显著性水平 $\alpha=0.05$)

无损状态		损伤工况 1		损伤工况 2	
H_0	H_1	H_0	H_1	H_0	H_1
24/30	6/30	0/30	30/30	0/30	30/30

0/30 表示在 30 次假设检验中结果为 H_0 的次数为 0;
30/30 表示 30 次假设检验中结果为 H_1 的次数为 30

4 结束语

笔者研究了基于结构加速度时间序列的损伤识别方法,以结构加速度 AR 模型系数为结构损伤特征向量,分别获取结构在不同状态下的损伤特征向

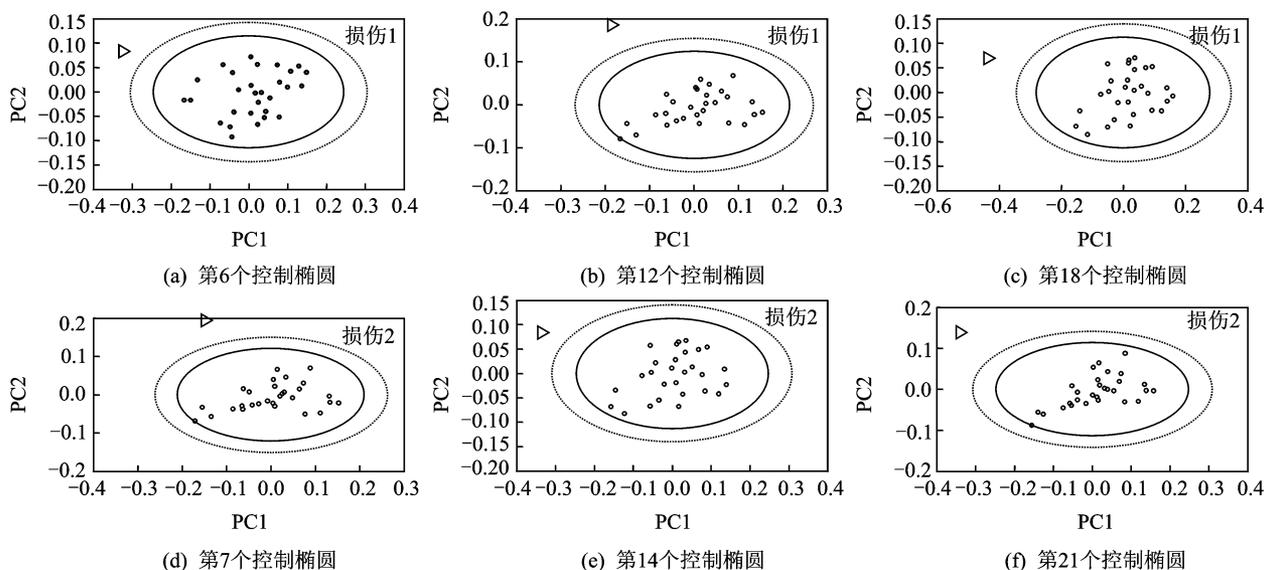


图5 损伤状态部分控制椭圆

量样本。以结构无损伤状态损伤特征向量为参考样本,以结构损伤状态特征向量为待检样本,把待检样本中各损伤特征向量逐个加入参考样本中,构建多个原始数据矩阵。对多个原始数据矩阵进行主成分分析,提取各原始数据矩阵的前两个主成分,并构造其相应的控制椭圆,由前两个主成分的散点图在控制椭圆中的分布判别结构是否损伤。用一钢结构框架试验验证了本研究方法的有效性和准确性。和马氏距离判别法相比,本研究方法具有更好的识别稳定性。

参 考 文 献

- [1] 郭健,陈勇,孙炳焯. 桥梁健康监测中损伤特征提取的小波包方法[J]. 浙江大学学报:工学版,2006,40(10):1767-1772.
Guo Jian, Chen Yong, Sun Bingnan. Wavelet packet method of damage feature extraction in bridge health monitoring[J]. Journal of Zhejiang University: Edition of Engineer Science, 2006, 40(10): 1767-1772. (in Chinese)
- [2] 孙增寿,任伟新. 基于小波熵指标的结构损伤检测[J]. 振动、测试与诊断,2008,28(3):233-237.
Sun Zengshou, Ren Weixin. Structural detection based on the index of wavelet entropy[J]. Journal of Vibration, Measurement & Diagnosis, 2008, 28(3): 233-237. (in Chinese)
- [3] Krishnan K N, Kiremidjian A S, Kincho H L. Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure[J]. Journal of Sound and Vibration, 2006, 291: 349-368.
- [4] Mustafa G, Necati F C. Statistical pattern recognition

for structure health monitoring using time series modeling: theory and experimental verifications[J]. Mechanical Systems and Signal Processing, 2009, 23: 2192-2204.

- [5] Hoon S, Charles R F. Damage diagnosis using time series analysis of vibration signals[J]. Smart Materials and Structures, 2001, 10: 446-451.
- [6] Michael L F, Hoon S, Charles R F. Vibration-based damage detection using statistical process control[J]. Mechanical Systems and Signal Processing, 2001, 15(4): 707-721.
- [7] 杨叔子,吴雅. 时间序列分析的工程应用(上)[M]. 武汉:华中科技大学出版社,2007:51-52.
- [8] 李卫东. 应用多元统计分析[M]. 北京:北京大学出版社,2008:196-198.
- [9] Housner G W, Bergman L A, et al. Structural control: past, present and future[J]. Journal Engineering Mechanics, 1997, 123(9): 897-971.
- [10] Farrar C R, Duffery T A, Doebling S W, et al. A statistical pattern recognition paradigm for vibration-based structural health monitoring[C]//The 2nd International Workshop on Structural Health Monitoring. Stanford, CA, 1999.



第一作者简介:吴森,男,1982年1月生,博士研究生。主要研究方向为结构健康监测、评价与维护。曾发表《振动分析在钢桁架结构损伤预警中的应用》(《海军工程大学学报》2010年第5期)等论文。

E-mail: Woserwusen201@sina.com