Journal of Vibration, Measurement & Diagnosis

doi:10.16450/j. cnki. issn. 1004-6801. 2015. 06. 025

局域波信息熵在高速自动机故障诊断中的应用

潘宏侠1, 都 衡2, 马春茂3

(1.中北大学机械与动力工程学院 太原,030051) (2.重庆青山工业有限责任公司 重庆,402761)(3.西北机电工程研究所 咸阳,712099)

摘要 针对自动机工作时的短时冲击信号特征,首先,对其运动形态进行分解,截取与故障构件动作相对应的振动 信号进行分析;其次,利用多层小波包分解截取信号,对其频率成分及能量分布进行研究;然后,将小波包分解后频 带能量大的信号重构进行局域波分解,同时提取局域波奇异谱熵、边际谱熵和特征空间谱熵定量描述信号状态的 时域、频域和能量的变化,并将其作为故障特征量;最后,利用遗传算法的全局寻优能力对支持向量机的参数进行 优化,建立了遗传支持向量机(genetic algorithms-support vector machine,简称 GA-SVM)模型,将提取的特征量输 入其中进行故障分类识别,并将识别结果与空间穷尽搜索支持向量机(support vector machine,简称 SVM)的识别 结果进行对比。

关键词 局域波;信息熵;自动机;遗传支持向量机;小波包分解 中图分类号 TH17

引 言

高速自动机是小口径高炮的核心部件,其工作 环境恶劣,运动规律复杂,在运动过程中经常伴随着 激烈的撞击、摩擦、振动和跳动等,因此,其工作可靠 性、各构件的裂纹及磨损逐渐成为火炮监测与诊断 领域关注的焦点。自动机作为一个复杂的机械系 统,当构件出现裂纹等故障时会影响其响应信号的 频率能量特性,因此,笔者采用现代测试与分析方 法,对自动机射击工作过程中的振动信号进行采集 并应用信号处理方法提取对故障信息敏感的特征 量,从而对故障加以识别。

为了使故障信息得以凸显,首先对自动机运动 形态进行分解,截取与故障构件动作相对应的振动 信号进行分析,并对截取信号进行多层小波包分解, 计算各频带归一化能量,了解信号的频率组成及能 量分布。自动机备件出现裂纹故障时,其频率、能量 成分必然会发生变化,而频带能量高的信号所包含 的故障信息量相对丰富,所以选择频带能量大的信 号作为分析对象。局域波是基于信号波形的瞬时和 局部特征,因此更加适合于非平稳信号的分析,而笔 者研究的自动机振动信号正是包含大量冲击的非平

1 小波包频带能量分析

本次试验故障设置在闭锁片和枪机上,而闭锁 片和枪机的主要动作是在闭锁行程和开锁行程完成 的,因此,截取该两段行程对应的信号进行分析。具 体故障设置如下:故障1在闭锁片闭锁斜面的圆角 处,沿其半径方向设置2.0 mm 深的裂纹槽,如图1 所示;故障2沿经过闭锁片回转圆心且垂直于闭锁 片内平面的方向设置1.5 mm 深的裂纹槽,如图2 所示;故障3 在枪机左右两侧的圆角矩形窗后端的 两对圆角上,沿圆角直径各成±45°切割1.5 mm 深 的裂纹槽,如图3 所示。

稳信号。局域波可以自适应地将信号分为多个反应 其模式组成的基本模式分量(intrinsic mode function,简称 IMF)分量,并且各 IMF 分量经过 Hilbert 变换后得到的时频谱,能同时反应时域、频域信息。 因此,用局域波对小波包分解后频带能量高的信号 进行分解可进一步细分频谱,出现故障后的信号频 率及能量变化会得到更加精确表现。反应复杂性测 度的信息熵是对信号不确定性的定量描述,因此可 应用信息熵理论对局域波分解后的信号提取特征量 来反映信号的状态变化。

^{*} 国家自然科学基金资助项目(51175480) 收稿日期:2014-12-17;修回日期:2015-06-03





图 1 故障 1 实物图 Fig. 1 A physical failure of fault 1

图 2 故障 2 实物图 Fig. 2 A physical failure of fault 2



图 3 故障 3 实物图 Fig. 3 A physical failure of fault 3

对通过试验采集的五连发射击振动信号截取测 点一 *x* 向各故障状况后四发振动信号进行分析。 考虑到采样频率较高,为 204.8 kHz,为了能更清楚 地了解信号频率分布情况,采用 db4 小波函数对截 取信号进行 6 层小波包分解,将其分解到 64 个频带 上,重构后计算归一化能量值,如图 4~图 7 所示。



图 4 五连发正常状况的 6 层小波包分解能量图

Fig. 4 Energy distribution of 6 layer wavelet packet decomposition of normal five series





从小波包分解归一化能量图中可以看出:4种 状况下的能量都主要集中在低频部分(频带1~8, 频率0~10kHz);在频带35~40,虽也出现能量较 大值,但相对低频部分要小的多,而4种状况的频带 3和频带7的能量值均为峰值。对比4种状况的频 带-能量图可发现,相互之间虽存在差别(如故障1 频带7能量值相对要低一些,频带35上能量值也不



图 6 五连发故障 2 的 6 层小波包分解能量图 Fig. 6 Energy distribution of 6 layer wavelet packet decomposition of fault 2 five series



图 7 五连发故障 3 的 6 层小波包分解能量图 Fig. 7 Energy distribution of 6 layer wavelet packet decomposition of fault 3 five series

同),但并不明显,也无规律可言。为进一步对故障 进行识别,提取频带3和频带7信号进行局域波分 解。图8,9为五连发射击故障1小波包重构前后截 取的信号时域图。

2 局域波信息熵特征提取

2.1 局域波奇异谱熵

奇异值具有好的稳定性是矩阵的固有特征,因 此可被提取作为特征量。

假设信号 X 采样所得的 M 点一维时间序列表



Fig. 9 Wavelet packet reconstruction signal of fault 1 five series

示为 { $X(i), i = 1, 2, \dots, M$ },采用延时嵌陷技术,嵌入维数为 m,时间延迟为 τ ,根据 Tekens 嵌入定 理^[1],将其嵌入到 $m \times n$ 维的相空间为

$$\mathbf{x}_{k} = \begin{bmatrix} X(k), X(k+\tau), X(k+2\tau), \cdots, \\ X(k+(m-1)\tau) \end{bmatrix}^{\mathrm{T}}$$
(1)

其中: $k=1,2,\dots,n;n=M-(m-1)\tau_{\circ}$

轨道矩阵 $x = [X_1, X_2, \dots, X_n]$ 代表相空间中 n个坐标点,形成一条运动的轨迹^[2]。对该矩阵进行 奇异值分解,得奇异值 $\delta_k, 1 \leq k \leq M - (m-1)\tau$ 。

奇异值分解需要延时重构相空间,而相空间重 构质量很大程度上要受重构参数中嵌入维数和延迟 时间的影响^[3],而这些参数的确定极为不便。局域 波可将信号分解为一系列 IMF 分量,这些分量分别 包含了从高到低不同频率成分、不同频带的信息,因 此,将这些 IMF 组成初始特征向量矩阵

$$\mathbf{A} = \begin{bmatrix} h_1, h_2, \cdots, h_n \end{bmatrix}$$
(2)

对其进行奇异值分解,得到奇异值 δ_i (1 $\leq i \leq$ n)。奇异值 δ_i 的大小是相应的模式在总的信号模 式中所占比重的反映,而非零奇异值的个数是对信 号频率成分组成多少的反映。{ δ_i }是对信号在基本 模式空间中的一种划分,称之为局域波奇异谱。由 此可定义奇异谱熵为

$$\begin{cases} \boldsymbol{H}_{qy} = -\sum_{i=1}^{n} p_{i} \log p_{i} \\ p_{i} = \delta_{i} / \sum_{i=1}^{n} \delta_{i} \end{cases}$$
(3)

其中: *p_i* 为在整个谱中第*i* 个奇异值所占的比例分额,即第*i* 个模式占的份额。

奇异谱熵反映了奇异谱划分下的系统振动能量 的不确定性。信号越简单,熵值越小,能量越集中于 少数几个模式;相反信号越复杂,熵值就越大,能量 越分散。

2.2 局域波特征空间谱熵与边际谱熵

Hilbert 时频谱 H(ω,t) 为基本模式分量在时 频空间上的能量分布矩阵。由于矩阵的奇异值具有 好的稳定性是矩阵的固有特征,因此对 H 进行奇异 值分解,定义局域波特征空间谱熵为

$$\begin{cases} \boldsymbol{H}_{kj} = -\sum_{i=1}^{n} p_i \log p_i \\ p_i = \delta_i / \sum_{i=1}^{n} \delta_i \end{cases}$$

$$(4)$$

其中: δ_i 为H 奇异值分解的第个奇异值; H_{kj} 反映 了局域波特征空间中局部能量分布的不确定性。

Hilbert 边际谱 *h*(ω) 反映了信号的频率-能量 结构,其含义与傅里叶频谱相类似。边际谱是对信 号频域的一种划分,局域波边际谱熵定义为

$$\begin{cases} \mathbf{H}_{bj} = -\sum_{i=1}^{n} p_i \log p_i \\ p_i = h(\boldsymbol{\omega}_i) / \sum_{i=1}^{n} h(\boldsymbol{\omega}_i) \end{cases}$$
(5)

H_{bj} 刻画了信号的瞬时谱型结构情况。*H_{bj}* 越 大,则能量在整个频率成分上分布的越均匀,不确定 性程度也就越大。由于局域波边际谱和时频谱中的 频率具有瞬时意义,因此 *H_{bj}*,*H_{bj}* 与瞬时频率成分 和振动能量在各模态的分布有关。

4种工况提取的部分特征量如表1所示。

表 1 自动机五连发射击部分特征量 Tab. 1 Characteristics of automaton shooting part five series

信旦业太		特征向量	
佰乡扒恋	$oldsymbol{H}_{kj}$	$oldsymbol{H}_{bj}$	$oldsymbol{H}_{qy}$
工业	8.126 7	7.693 4	1.594 4
止币	7.853 8	7.534 9	1.568 8
+++ 17字 1	7.268 1	6.981 4	1.190 7
₩□	7.518 9	7.202 1	1.203 0
抽座の	7.974 6	7.332 6	0.842 4
₩Ψ Δ	7.850 6	7.169 6	0.740 1
抽座っ	7.950 1	7.705 6	1.846 6
以悼 3	8.102 0	7.683 3	1.743 7

3 GA-SVM 模型

支持向量机(support vector machine,简称 SVM)的核心思想是通过核函数将输入空间的样本 从原空间映射到高维特征空间。给定样本集{ $(x_1, y_1), (x_2, y_2), \dots, \{(x_n, y_n)\},$ 其中输入样本 $x_i \in R^n$,输出样本 $y_i \in \{-1, +1\},$ SVM 通过核函数 $K(x_i, x_j)$ 将输入样本映射到 N 维特征空间,并在 N 维特征空间构造最优分类超平面 $\sum_{i=1}^{N} w_i k(x, x_i) + b = 0$ 。SVM 要求最优超平面中分类间隔 $2/\|w\|$ 达最大,由此,引入松弛变量,则最优超平 面转化为如下约束优化问题

$$\begin{cases} \min_{\boldsymbol{w},b,e} J(\boldsymbol{w},e) = \frac{1}{2} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{w} + C \sum_{k=1}^{n} e_{k}^{2} \\ \text{s. t.} \quad y_{k}(\boldsymbol{w}^{\mathrm{T}} \varphi(x_{k}) + b) \geqslant 1 - e_{k} \end{cases} (k = 1, 2, \cdots, n)$$

$$(6)$$

其中: e_k 为松弛变量; C 为惩罚系数。

利用拉格朗日乘子,将其转化为对偶最优化问题,则最终的分类决策函数为

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{N} y_{i}\alpha_{i}k(x, x_{i}) + b\right)$$
(7)

SVM 通过结构风险最小化原理来提高泛化能 力,具有直观的集合解释和良好的分类能力,已广泛 应用于故障诊断、控制理论等领域,但是 SVM 模型 参数的选择仍未解决。SVM 的优越性能需要选取 合适的参数值才能体现,参数取值的好坏很大程度 上影响了故障诊断的精度和效率[4]。支持向量机的 参数选取通常采用参数空间穷尽搜索法,在实际问 题中,因为很难事先知道最优参数所在的范围,所以 这种方法不但很难精确地找到最优参数,而且采用 遍选的方式效率也较低^[5]。影响 SVM 性能主要有 3个因素:分类间隔、对偶和核函数,而核函数更是 其中的关键^[6]。核函数参数σ主要影响样本数据在 高维特征空间中分布的复杂程度,而误差惩罚因子 的作用是在确定的特征空间中调节学习机的置信范 围和经验风险的比例^[7]。因此要提高 SVM 的分类 识别能力,就需要选择合适的参数 σ 和c。

遗传算法是一种借鉴生物界自然选择和优化机制的高度并行、随机及自适应搜索的全局寻优算法, 特别适合于传统搜索算法解决不好的复杂的和非线 性问题^[8]。遗传算法的初始群体是随机产生形成 的,然后不断进化更新群体。适应度高的个体被直 接复制到下一代群体中,其产生后代的概率较高。 交叉和变异分别是以一定概率在染色体上进行基因 交换或基因变异操作^[9]。

支持向量机可较好地解决小样本局部极小值问题,这一特点正好解决了自动机故障特征量样本少的问题。遗传算法的优点是能够处理各种类型的目标函数和约束方便^[10],对所解的优化问题数学要求少,并且具有不依赖于特定求解模型的全局搜索能力,其优化过程中所需的较高计算量也可通过利用支持向量机的快速映射能力得到满足^[11]。因此,可用遗传算法来优化支持向量机的参数,避免人为选择参数的盲目性。因此,笔者将二者结合,建立了GA-SVM模型以提高对故障的识别率。

4 自动机故障诊断试验

试验中各工况分别进行了一次五连发射击,考 虑到第1发受初始因素影响,截取后4发的振动信 号进行分析并提取特征值,这样每种状况包含4组 样本。采用支持向量机进行分类,以每种状况的两 组样本做训练样本,其余两组做测试样本输入支持 向量机中进行识别。

为了获取最佳参数以提高故障识别准确率,分 别建立了参数空间穷尽搜索 SVM 和 GA-SVM 分 类识别模型。两种模型分别通过空间穷尽搜索和遗 传算法来优化误差惩罚因子和核函数参数。空间穷 尽搜索 SVM 模型中和的范围均设在[-2¹⁰,2¹⁰]范 围内,以一定的间隔取值,然后选择使得期望风险上 界最小的一组参数组合作为最优参数值。对于 GA-SVM 模型,设定群体规模 m = 50,最大遗传代 数 t=70, 变量数目 n=3, 交叉概率 $p_{\epsilon}=0.35$, 变异 概率 $p_m = 0.02$, 产生初始种群, 利用遗传算法自动 进行全局搜索寻找最优参数。寻优过程中适应度函 数随进化代数变化曲线如图 10 所示。将表1特征 量作为训练样本分别输入利用空间穷尽搜索法和遗 传算法得到最优参数的两种模型中,对其进行训练。 相应的输出设为:正常→1;故障 1→2;故障 2→3; 故障3→4。

最后将测试样本输入两种模型中,得出诊断结 果。自动机五连发射击测试样本特征量及诊断结果 如表2所示。两种模型得到的最优参数及诊断正确 率如表3所示。结果表明,两种模型对训练样本的 识别率都能达到100%,但测试样本 GA-SVM 模型





Fig. 10 Fitness function curve of genetic algorithm optimizing of five series

对故障识别率更高,准确率为87.5%,而空间穷尽 搜索SVM识别率只有75%,说明了遗传算法相对 传统的空间穷尽搜索方法可以更好地优化参数。

表 2 自动机五连发射击测试样本特征量及诊断结果

Tab. 2 Test sample characteristics and diagnosis results of automaton five series firing

信号		特征向量		期望	穷尽搜索	GA-SVM
状态	$oldsymbol{H}_{kj}$	$oldsymbol{H}_{bj}$	H_{qy}	输出	SVM 输出	模型输出
正常	8.150 5	7.7227	1.812 3	1	1	1
	7.815 0	7.524 9	1.377 8	1	3	1
故障1	7.246 9	7.0287	0.909 1	2	2	2
	7.206 7	7.012 9	0.833 8	2	2	2
故障2	7.896 8	7.423 9	1.133 3	3	3	3
	7.818 1	7.136 4	0.752 1	3	3	3
故障3	8.077 3	7.7977	2.278 3	4	4	4
	8.025 0	7.682 5	1.584 2	4	1	1

表 3 五连发射击两种诊断模型的最优参数及识别正确率 Tab. 3 Optimal parameters and identification accuracy on the

two diagnosis model of five series shooting

壮 陈 汨 則 措 刊	进而会	*/	识别准确率/%			
取障切刑侠堂	此收参数 (, ,)		训练样本	测试样本		
穷尽搜索 SVM	1 024,	1 024	100	75.0		
GA-SVM 模型	18.81,	228.53	100	87.5		

5 结 论

 提出了应用小波包分解提取频带能量大的 信号重构,并对其进行局域波分解进一步细化频谱, 同时分别利用局域波奇异谱熵、边际谱熵和特征空 间谱熵定量描述信号状态的时域、频域和能量变化 的特征量提取方法。 2)针对SVM的优越性能需要选取合适的参数 值才能体现,而参数选择尚未得到有效解决,运用具 有全局自动寻优能力的遗传算法对SVM进行参数 寻优,建立了GA-SVM故障识别模型对故障进行 分类识别。

3) 与空间穷尽搜索法寻优的 SVM 相比,GA-SVM 具有更高的识别正确率。两种故障模型都具 有较高的分类识别能力,体现了信息熵提取特征向 量方法的可行性。为自动机从故障特征量提取到分 类识别故障提供了一种有效的方法。

参考文献

- [1] Takens F. Detecting strange attractors in fluid turbulence[J]. Lecture Notes in Mathematics, 1981, 898: 363-381.
- [2] 卢德林,郭兴明. 基于奇异谱分析的心音信号小波包 去噪算法研究[J].振动与冲击,2013,32(18):63-69.
 Lu Delin, Guo Xingming. Wavelet packet denoising algorithm for heart sound signal based on singular spectrum analysis [J]. Journal of Vibration and Shock, 2013, 32(18): 63-69. (in Chinese)
- [3] 于大鹏,赵德有.时延法相空间重构双参数联合估计 方法研究[J].大连理工大学学报,2009,49(4):540-544.

Yu Dapeng, Zhao Deyou. Study of dual parameters of delay-time coordinates phase space reconstruction[J]. Journal of Dalian University of Technology, 2009,49 (4): 540-544. (in Chinese)

- [4] Yuan Shengfa, Chu Fulei. Fault diagnosis based on support vector machines with parameter optimization by artificial immunization algorithm [J]. Mechanical Systems and Signal Processing, 2007,21:1318-1330.
- [5] 陈法法,汤宝平,马婧华,等.基于遗传退火优化
 MSVM的齿轮箱故障诊断[J].振动、测试与诊断, 2014,34(4):699-704.
 Chen Fafa, Tang Baoping, Ma Jinghua, et al. Gear-

box fault diagnosis based on multi-kernel MSVM optimized by genetic simulated annealing algorithm [J]. Journal of Vibration, Measurement & Diagnosis, 2014, 34(4):699-704. (in Chinese)

[6] 徐玉秀,杨文平,吕轩,等.基于支持向量机的汽车发动 机故障诊断研究[J]. 振动与冲击,2013,32(8):143-146.

Xu Yuxiu, Yang Wenping, Lü Xuan, et al. Fault diagnosis for a car engine based on support vector machine[J]. Journal of Vibration and Shock, 2013, 32
(8):143-146. (in Chinese)

[7] 霍天龙,赵荣珍,胡宝权.基于熵带法与 PSO 优化的
 SVM 转子故障诊断[J].振动、测试与诊断,2011,31
 (3):279-284.

Huo Tianlong, Zhao Rongzhen, Hu Baoquan. Fault diagnosis for rotor systems based on entropy band method and support vector machine optimized by PSO [J]. Journal of Vibration, Measurement & Diagnosis, 2011, 31(3):279-284. (in Chinese)

[8] 李锋,汤宝平,刘文艺.遗传算法优化最小二乘支持向 量机的故障诊断[J].重庆大学学报,2010,33(12): 14-20.

Li Feng, Tang Baoping, Liu Wenyi. Fault diagnosis based on least square support vector machine optimized by genetic algorithm[J]. Journal of Chongqing University, 2010, 33(12): 14-20. (in Chinese)

[9] Acevedo-Rodríguez J, Maldonado-Bascón S, Lafuente-Arroyo S, et al. Computational load reduction in decision functions using support vector machines[J]. Signal Processing, 2009,89(10):2066-2071.

- [10] Hajnayeb A, Ghasemloonia A, Khadem S E, et al. Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis [J]. Expert Systems with Applications, 2011, 38(8): 10205-10209.
- [11] Jun Ye. Fault diagnosis of turbine based on fuzzy cross entropy of vague sets[J]. Expert Systems with Applications, 2009,36:8103-8106.



第一作者简介:潘宏侠,男,1950年10 月生,教授、博士生导师。主要研究方向 为机电系统检测诊断与控制。曾发表 《基于局域波降噪和双谱分析的自动机 故障诊断研究》(《兵工学报》2014年第 35卷第3期)等论文。

E-mail:panhx1015@163.com

向审稿人致谢

(排名不分先后)

姚志远	霍学晋	姚文席	周思达	张春宜	牛军川	赵荣珍	王义亮	金家楣	向志海	许金余
黄国勇	王广林	黄庆华	张梅军	贾宏宇	孙志峻	熊晓燕	危银涛	霍林生	石启印	曹茂森
李允公	时运来	曾周末	阎绍泽	高 亮	朱劲松	袁惠群	汤宝平	阎绍泽	陈怀海	向 玲
刘增华	杨运强	李兆霞	刘彦辉	李志农	左言言	袁慎芳	于秀敏	贾宏宇	边义祥	鹿存跃
杨文伟	张清东	刘芳华	金栋平	张云清	王 雪	纪跃波	沈振中	孙晓燕	王 见	郝如江
陈 超	苏怀智	蒋伟康	张小栋	伊廷华	徐伟炜	赵玉成	危银涛	张 方	束学道	郑 辉
卜长根	齐朝辉	李 辉	张周锁	吕延军	徐志伟	常 佶	张西宁	卢碧红	臧 勇	李瑰贤
洪 俊	谢官模	罗贵火	田小华	吴子燕	袁万城	谌 勇	刘应华	邱志成	徐存东	陈国平
张建辉	姚小虎	施 洲	陈仲生	姚国凤	张 昭	杨润林	姜金辉	奚 鹰	孙文磊	鲁亮
张孝远	李建华	荣伟彬	石启印	周瑾	孙增寿	叶雪荣	李仁年	张 洁	陈雪峰	许明恒
陈花铃	王 彤	夏毅敏	赵宏林	马 高	缪长青	杨庆山	陈焕新	孔德文	李华峰	谷立臣
王红军	陈小安	王文中	石可重	周星德	高 斌	朱 华	齐朝晖	李秀领	李 兵	鲍跃全
马 辉	温广瑞	郭 瑜	王 轲	林明星	周广东	冯志鹏	郝际平	王良模	刘宏伟	张 晞
张 明	蔡敢为	张 超	赵又群	岳 林	史铁林	王三民	傅 攀	李允公	王保华	吴洪涛
臧朝平	孔德文	骆 英	钟佑明	王从庆	吕延军	王书翰	董广明	温卫东	董明明	潘宏侠
王 立	李 亮	赵学智	夏齐霄	廖广兰	朱忠奎	余成波	安 琦	招启军	张新华	杜 宇
贺西平	陈定方	赵英凯	王国强	邢宗义	胥永刚	程 珩	左晓宝	洪 军	杨玉虎	上官文斌