

柴油机振动信号快速稀疏分解与二维特征编码*

王 旭, 岳应娟, 蔡艳平

(火箭军工程大学理学院 西安, 710025)

摘要 针对柴油机故障诊断方法中的信号时频表征及特征提取问题,提出一种基于振动信号快速稀疏分解与二维时频特征编码识别的柴油机智能故障诊断方法。首先,为了获得时、频聚集性优良的时频图像,提出一种随分解残差信号自适应更新 Gabor 字典的改进匹配追踪(adaptive matching pursuit, 简称 AMP)算法,利用 AMP 算法将柴油机振动信号分解后叠加各原子分量的 Wigner-Ville 分布,获取原信号的稀疏分解时频图像;然后,为提取时频图像的特征参量,提出了双向二维非负矩阵分解(two-directional, 2-dimensional non-negative matrix factorization, 简称 TD2DNMF)算法,用于对时频图像的幅值矩阵进行特征编码,获取蕴含在时频图像内部的低维特征,并利用最近邻分类器实现了时频图像的自动分类识别。将提出的方法应用于 4 种不同状态柴油机气门故障的诊断试验中,结果表明,该方法能够获得无交叉项干扰、聚集性好的时频图像,使各时频分量的物理意义更加明确,并改进了传统图像模式识别中的特征参数提取方法,是一种有效的柴油机故障诊断方法。

关键词 柴油机; 匹配追踪; 二维非负矩阵分解; 特征编码; 时频分布

中图分类号 TH17; TK428

引 言

柴油机运行中既有旋转运动,又有往复运动,其振动响应信号复杂,耦合严重,具有较强的非线性、非平稳时变特征^[1]。采用时频分析是通过对柴油机振动时频图像的识别来实现对故障的诊断,其技术关键在于准确的时频表征和有效的特征提取,众多学者对此进行了大量的探索。文献[2]提出利用灰度共生矩阵提取柴油机振动信号三阶累积量图像的故障诊断方法。文献[3]将 Wigner 分布与分形维数相结合,对柴油机进行故障诊断。文献[4]将图像分割理论引入柴油机故障诊断之中。文献[5]提出基于经验模态分解和维格纳分布的柴油机振动信号时频表征方法。以上方法将柴油机的故障诊断问题转化为柴油机振动时频图像的模式识别问题,取得了较好的效果。

时频表征方面,利用小波分析(wavelet analysis, 简称 WT)、维格纳分布(Wigner-Ville distribution, 简称 WVD)及短时傅里叶变换(short time Fourier transformation, 简称 STFT)等方法对柴油机缸盖振动信号进行时频表征是应用较广泛的信号

处理手段^[6]。匹配追踪算法(matching pursuit, 简称 MP)是一种基于过完备冗余时频字典对信号进行稀疏分解的方法,相比于以上时频分析方法,MP 算法自适应性更好,对时频分布中各分量的刻画能力更强。但是 MP 算法计算量和存储量相当大,限制了其在柴油机振动信号处理中的应用^[7]。

特征提取方面,为了不依靠先验知识实现对不同工况时频图像的自动分类,需要对时频图像进行特征提取,其本质是对时频图像数据降维的同时最大限度保留不同样本间的差异化信息。非负矩阵分解算法(non-negative matrix factorization, 简称 NMF)是一种基于局部特征的特征提取方法,由于添加了非负的限制条件,能够保证分解结果的可解释性。文献[8-9]将 NMF 应用于柴油机和轴承时频图像矩阵的特征提取,取得了较好的效果。但是 NMF 算法在降维前需将图像矩阵向量化处理,破坏了图像矩阵间的空间位置信息,并且向量化处理后的图像矩阵往往维数较高,使得 NMF 的计算效率较低。对此,文献[10-12]从不同角度提出了二维非负矩阵分解(2-dimensional non-negative matrix factorization, 简称 2DNMF)方法,分解前不需将图像矩阵向量化,提高了特征提取效果,被成功应用于

* 国家自然科学基金资助项目(51405498);中国博士后基金资助项目(2015M582642)

收稿日期:2017-04-01;修回日期:2017-05-18

故障诊断领域^[13-14]。但是2DNMF在分解前都是将所有训练图像矩阵按行或列拼合,拼合后的初始分解矩阵维度依然较大,并且2DNMF算法没有考虑到不同类图像间的差异化信息,将所有图像样本一致对待,统一求解投影矩阵,这显然是不利于模式识别的。

基于以上分析,笔者着眼于柴油机故障诊断中的时频表征与特征提取过程,针对MP算法分解效率问题和2DNMF特征提取性能问题,提出了基于AMP算法和双向二维非负矩阵分解的柴油机振动信号时频分析方法。将该方法应用于气门间隙正常、气门间隙过小、气门间隙过大和气门漏气4种不同气门状态信号的诊断试验中,故障识别正确率最高可达100%,充分证明该方法用于柴油机自动故障诊断的有效性。

1 信号的稀疏分解及时频表征

1.1 改进的自适应MP算法实现

匹配追踪算法在计算和存储上的瓶颈主要在于过完备字典的制备,在分解时需要遍历一个庞大的原子字典中的所有原子后再找到最匹配的原子组合。柴油机振动信号通常会聚集于有限宽度的频带中,而Gabor原子具有最好的时频聚集性,所以利用Gabor原子字典分解信号时对其他频率范围内原子的搜索过程其实是无意义的。如果令字典随残余信号的功率谱分布自适应更新,缩小原子搜索范围,在保证算法稀疏性的基础上可以大幅提高算法的计算效率。基于此,笔者提出了自适应匹配追踪算法。

设 H 为Hilbert空间, $D = \{g_\gamma(t)\}_{\gamma \in \Gamma}$ 为 H 中过完备字典。原子 $g_\gamma(t)$ 由参数 γ 描述。Gabor原子^[15]可表示为

$$g_\gamma = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t} \quad (1)$$

其中: $g(t) = e^{-\pi t^2}$ 为高斯窗函数;原子参数 $\gamma = (s, u, \xi)$; s 为尺度因子; u 为位移因子; ξ 为频率因子。

原子经归一化处理后 $\|g_\gamma(t)\| = 1$ 。设 f 为待分解信号,且 $f \in H$,有

$$f = \sum_{n=0}^{m-1} \langle R_f^n, g_{\gamma_n} \rangle g_{\gamma_n} + R_f^m \quad (2)$$

$$\|f\|^2 = \sum_{n=0}^{m-1} \left| \langle R_f^n, g_{\gamma_n} \rangle \right|^2 + \|R_f^m\|^2 \quad (3)$$

其中: R_f^m 为第 m 次迭代残留信号; g_{γ_n} 为第 n 次迭代中的匹配原子。

为使 R_f^m 最小,应满足

$$\left| \langle R_f^n, g_{\gamma_n} \rangle \right| \geq \alpha \sup_{\gamma \in \Gamma} \left| \langle R_f^n, g_\gamma \rangle \right| \quad (0 \leq \alpha \leq 1) \quad (4)$$

为减小字典存储量,利用残留信号自适应地更新原子字典。令 F_R 为残余信号的Fourier变换,定义信号频率域峰值 $\max F_R$ 对应频率为 ξ_0 。定义新字典 $D_n (D = \{g_\gamma(t)\}_{\gamma \in \Gamma_n})$,其中 $\gamma = (s, u, \xi_0)$ 。将残留信号 R_f^n 在 D_n 中寻找最匹配的原子,得到新的残留信号 R_f^{n+1} ,满足

$$R_f^{n+1} = f - \sum_{i=1}^n \langle R_f^i, g_{\gamma_i} \rangle g_{\gamma_i} \quad (5)$$

Gabor原子在频域的能量主要集中在以调制频率为中心的频率区域内,自适应字典 D_n 中所有的原子主要能量均聚集于中心频率 ξ_0 附近,使得该字典中的原子和残留信号能有较好的匹配性。新字典为 D_n 的参数集合 $\Gamma_n \in R^+ \times R$ 中包含 $\gamma = (a^j, pa^j \Delta u, ka^j \Delta \xi)$, $a = 2, \Delta u = 1/2, \Delta \xi = \pi, 0 < j < \log_2 N, 0 \leq p < N2^{-j+1}, 0 \leq k < 2^{j+1}$ 。将字典中每个原子簇的 N 个原子与残留信号的 N 次内积转化为一次互相关操作,通过相关系数搜索字典中的原子,进一步减少计算量。

AMP算法流程如下。

1) 定义 $s = 2^j$,其中: $j \in (0, \log_2 N)$; $u = s$ 。计算原子 g_γ 的包络 $g((t-u)/s)$,并保存为包络库。由于Gabor原子的能量主要集中在区域 $[0, 2s]$ 内,为了减少计算量,每个包络可以只计算区域 $[0, 2s]$ 内的数值点。

2) 对残余信号作 N 点的Fourier变换,找出幅值最大处对应的频率 ξ_0 ,并保存Fourier变换的结果 $\text{FFT}(R_f^m)$ 。

3) 从包络库中取出对应的包络,用 ω_0 进行频率调制并归一化处理得到Gabor原子 $g(2^j, 2^j, \xi^m)$,和残余信号互相关得到 $\text{FFT}(R_f^m) \cdot \overline{\text{FFT}(g(2^j, 2^j, \xi^m))}$ 。

4) 找到上步运算中的互相关系数最大值,确定相应的匹配到的原子参数。

5) 判断停机条件,如果 $R_f^m/f \leq \zeta$ 或限定迭代次数,满足时停止迭代,否则转到步骤2。

1.2 仿真分析

为分析AMP算法的性能,建立一个多分量仿真信号 $x(t)$,信号长度为256,由5个具有高斯包络的原子信号分量叠加而成,采样频率归一化为1 Hz,5个信号分量的时、频域分布中心分别位于 $(t_1,$

$f_1) = (60, 0.1), (t_2, f_2) = (60, 0.4), (t_3, f_3) = (130, 0.25), (t_4, f_4) = (200, 0.1)$ 和 $(t_5, f_5) = (200, 0.4)$, 仿真信号表达式为式(6)。信号的时域波形与时频分布对应如图1所示。

$$x(t) = \left\{ \begin{array}{l} \cos(2\pi 0.01t) + \\ \cos(2\pi 0.04t) \end{array} \right\} \exp\left\{-\left(\frac{t-60}{0.07 \times 256}\right)^2\right\} + \cos(2\pi 0.025t) \exp\left\{-\left(\frac{t-130}{0.07 \times 256}\right)^2\right\} + \left\{ \begin{array}{l} \cos(2\pi 0.01t) + \\ \cos(2\pi 0.04t) \end{array} \right\} \exp\left\{-\left(\frac{t-200}{0.07 \times 256}\right)^2\right\} \quad (t=1, 2, \dots, 256) \quad (6)$$

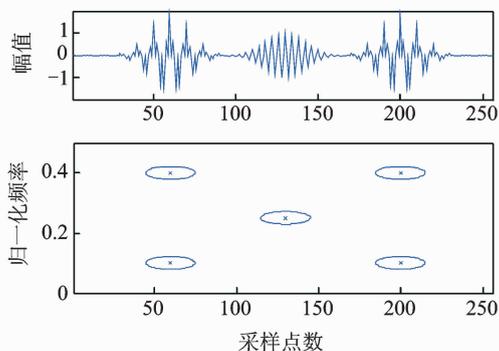


图1 仿真信号时域波形及时频分布对应图

Fig. 1 The time and frequency domain waveform

表1 6次迭代残余信号能量及耗时

Tab. 1 Residual signal energy and time cost in 6 iterations

迭代次数	MP耗时/ms	AMP耗时/ms	MP残余能量/%	AMP残余能量/%
1	504.043	4.649	82.64	79.07
2	394.473	2.436	64.62	58.17
3	382.044	2.132	44.04	37.27
4	397.459	2.002	23.31	19.17
5	418.764	2.501	2.95	1.07
6	481.936	2.109	2.15	0.82

表1为6次迭代过程中MP与AMP两种算法对应的原子匹配时间和残余信号能量百分比。理论上,仿真信号由5个原子分量复合而成,所以无论MP算法还是AMP算法都应在5次迭代后匹配出所有原子分量。MP与AMP算法在第5次迭代后,残余信号的能量分别达到了2.15%和0.82%,均较好地匹配出了信号的原子分量,实现了信号的稀疏分解。由于本研究中AMP算法根据残余信号功率谱分布将搜索字典限定在固定频带中,其搜索到的原子对于分解信号的匹配性更好,所以每次迭代中AMP算法的残余能量都低于MP算法。同时,由于AMP算法中自适应匹配的原子字典较MP算法

中的过完备冗余字典维度有较大幅度的压缩,使得AMP算法较MP算法的计算效率提高了150~200倍,从而能够利用AMP算法对高维度、强耦合的信号进行分析。

为了获得信号的时频分布,利用AMP算法将仿真信号分解为5个原子分量,通过二次叠加每个原子的Wigner-Ville分布得到原信号的时频分布。仿真信号的WVD分布与AMP稀疏分解时频分布情况如图2所示。

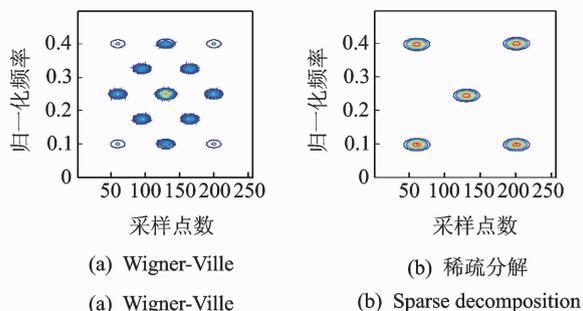


图2 仿真信号时频分布

Fig. 2 Time-frequency distribution of the simulation signal

由图2看出,直接对仿真信号进行WVD分析,每两个原子之间都会产生一个交叉项,严重影响对原信号分量的分析。利用AMP算法得到的信号稀疏分解时频分布中交叉项的问题得到了很好的解决,所有原子分量能量分布均衡,且保持了优良的时频聚集性,所以利用AMP算法对柴油机振动信号进行稀疏分解,能够使各时频分量物理意义更加明确,有利于对柴油机故障的判别。

2 基于双向2DNMF的时频特征编码方法

为了获得更好的特征提取效果,笔者提出一种双向二维非负矩阵分解算法,其与NMF,2DNMF有着相同的目标函数及迭代规则,但是架构不同。TD2DNMF算法分别对行基、列基矩阵进行求解,进一步得到TD2DNMF的二维基,用于对时频图像进行特征编码。

首先,对行基进行求解,求解时考虑到不同类别样本的差异性,对各类别样本矩阵并行运算得到相应的基矩阵和系数矩阵;然后,组合成整个训练样本的行基投影矩阵和系数矩阵,由此得出的系数矩阵具有更好的稀疏度。假设有 k 类模式 $\omega_1, \omega_2, \dots, \omega_k$,每类模式有 m 个训练样本图像 $\mathbf{A}_{a,b}, a=1, 2,$

$\dots, k; b=1, 2, \dots, m$, 每个图像对应大小为 $p \times q$ 的矩阵。将所有图像用矩阵 \mathbf{X} 表示为

$$\mathbf{X}_{p \times qm} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k] \quad (7)$$

其中: $\mathbf{X}_a = [\mathbf{A}_{a,1}, \mathbf{A}_{a,2}, \dots, \mathbf{A}_{a,m}]$; $a=1, 2, \dots, k$ 。

对于非负矩阵 \mathbf{X}_a , 求解非负矩阵 \mathbf{L}_a 和 \mathbf{H}_a , 满足

$$\mathbf{X}_a \stackrel{\text{NMF}}{\approx} \mathbf{L}_a \mathbf{H}_a \quad (8)$$

其中: $\mathbf{X}_a \in \mathbb{R}^{p \times qm}$, $\mathbf{L}_a \in \mathbb{R}^{p \times r}$, $\mathbf{H}_a \in \mathbb{R}^{r \times qm}$; r 为特征维数, 满足 $(p + qm)r < pqm$ 。

为了描述 $\mathbf{X} \approx \mathbf{L} \cdot \mathbf{H}$ 的近似效果, 利用矩阵 \mathbf{X} 与 $\mathbf{L} \cdot \mathbf{H}$ 间的 K-L 散度作为近似误差, 对应的目标函数^[16]为

$$D(\mathbf{V} \parallel \mathbf{WH}) \approx \sum_{ij} (\mathbf{V}_{ij} \log \frac{\mathbf{V}_{ij}}{(\mathbf{WH})_{ij}} - \mathbf{V}_{ij} + (\mathbf{WH})_{ij}) \quad (9)$$

其对应的优化问题为

$$\min D(\mathbf{V} \parallel \mathbf{WH}) \quad \text{s. t. } \mathbf{W}, \mathbf{H} \geq 0 \quad (10)$$

相应的迭代规则为

$$\mathbf{H}_{rj} \leftarrow \mathbf{H}_{rj} \sum_i \mathbf{L}_{ir} \mathbf{X}_{ij} / (\mathbf{LH})_{ij} \quad (11)$$

$$\mathbf{L}_{ir} \leftarrow \mathbf{L}_{ir} \sum_j \mathbf{H}_{rj} \mathbf{X}_{ij} / (\mathbf{LH})_{ij} \quad (12)$$

$$\mathbf{L}_{ir} \leftarrow \frac{\mathbf{L}_{ir}}{\sum_i \mathbf{L}_{ir}} \quad (13)$$

由此得到每一类别图像数据矩阵对应的分解矩阵因子, 对于矩阵 $\mathbf{X}_{p \times qm}$ 对应的 TD2DNMF 分解因子可由每一类别的分解因子组合而成

$$\mathbf{X}_{p \times qm} \stackrel{\text{TD2DNMF}}{\approx} \mathbf{L}_{p \times R} \mathbf{H}_{R \times qm} \quad (14a)$$

$$\mathbf{R} = kr \quad (14b)$$

$$\mathbf{L}_{p \times R} = [\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_k] \quad (15)$$

$$\mathbf{H}_{R \times qm} = \begin{bmatrix} \mathbf{H}_1 & & & \\ & \mathbf{H}_2 & & \\ & & \ddots & \\ & & & \mathbf{H}_k \end{bmatrix} \quad (16)$$

然后, 对列基进行求解, 在此直接利用求解出的系数矩阵 $\mathbf{H}_{R \times qm}$ 来求解列基。每一类图像样本的系数矩阵 $\mathbf{H}_a \in \mathbb{R}^{r \times qm}$ 由 m 个图像系数 $\mathbf{h}_c \in \mathbb{R}^{r \times q}$, $c=1, 2, \dots, m$ 组合而成, 将 \mathbf{h}_c 转置后得到 $\mathbf{h}_c^T \in \mathbb{R}^{q \times r}$, 将所有转置后的系数按行排列为如下形式

$$\mathbf{H}^T = [\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_k^T] \quad (\mathbf{H}^T \in \mathbb{R}^{q \times Rm}) \quad (17)$$

$$\mathbf{H}_a^T = [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_m^T] \quad (\mathbf{H}_a^T \in \mathbb{R}^{q \times rm}; a=1, 2, \dots, k) \quad (18)$$

与行基的求解过程类似, 根据 K-L 散度构造目标函数进行迭代运算, 确定列基分解的特征维数 d , 得到矩阵 \mathbf{H}^T 对应的 TD2DNMF 分解因子

$$\mathbf{H}_{q \times Rm}^T \stackrel{\text{TD2DNMF}}{\approx} \mathbf{R}_{q \times d} \mathbf{W}_{d \times Rm} \quad (19)$$

任意图像矩阵 $\mathbf{A}_{p \times q}$ 在行基和列基上的投影系数矩阵为

$$\mathbf{F}_{R \times d} = \mathbf{L}_{R \times p}^T \mathbf{A}_{p \times q} \mathbf{R}_{q \times d} \quad (20)$$

TD2DNMF 的二维基定义为

$$\mathbf{B} = \mathbf{LR}^T \quad (21)$$

通过将柴油机振动时频图像向 TD2DNMF 的二维基进行投影, 可得到相对应的特征编码。利用分类器对特征编码进行识别, 可实现时频图像的自动分类, 进一步可实现对故障的判别。

3 柴油机故障诊断实例分析

基于 AMP 振动信号稀疏分解与 TD2DNMF 二维时频特征编码的柴油机故障诊断方法流程如图 3 所示。



图 3 故障诊断流程框图

Fig. 3 The flowchart of fault diagnosis

3.1 试验设备与工况设置

以 6135G 型柴油机为研究对象, 用加速度传感器测量缸盖振动信号, 试验平台如图 4 所示。分别将气门间隙设置成不同状态, 用以模拟气门的不同间隙异常故障。在转速稳定于 1 500 r/min 时测量第 2 缸的缸盖振动信号, 单通道采样频率为 25 kHz, 试验具体工况设置如表 2 所示。试验中运行状态为空载, 气门正常间隙值为 0.30 mm, 使用 0.06 mm 模拟气阀间隙过小, 使用 0.50 mm 模拟气阀间隙过大, 在气阀上开 4 mm × 1 mm 孔来模拟严

重漏气故障。共采集柴油机气门4种状态下各60组振动信号,总计240个样本。

表2 4种试验工况设置

Tab. 2 Four states of IC engine's valve train mm

工况编号	进气门	排气门
1	0.30	0.30
2	0.30	0.06
3	0.30	0.50
4	0.30	开口 4×1



(a) 试验装置 (b) 振动加速度传感器位置
(a) Experimental device (b) Vibration acceleration sensor

图4 试验平台及传感装置设置

Fig. 4 Experimental Platform and sensing device

3.2 信号稀疏分解生成时频图像

图5为4种工况下振动信号的时域波形。进气门开启的角度在 310° 附近,关闭的角度在 -120° 附近;排气门开启的角度在 120° 附近,关闭的角度在 -310° 附近,柴油机在 0° 产生燃烧激振。

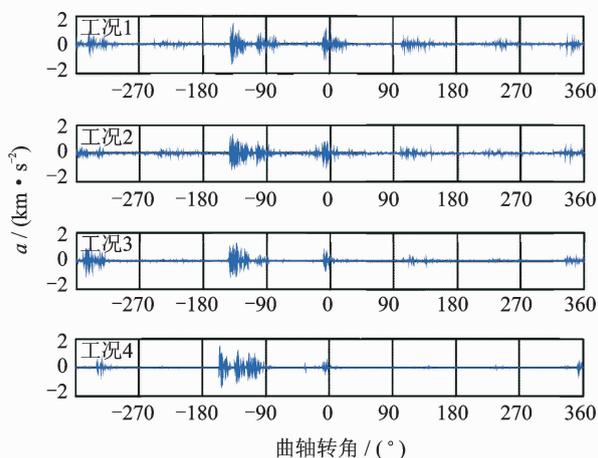
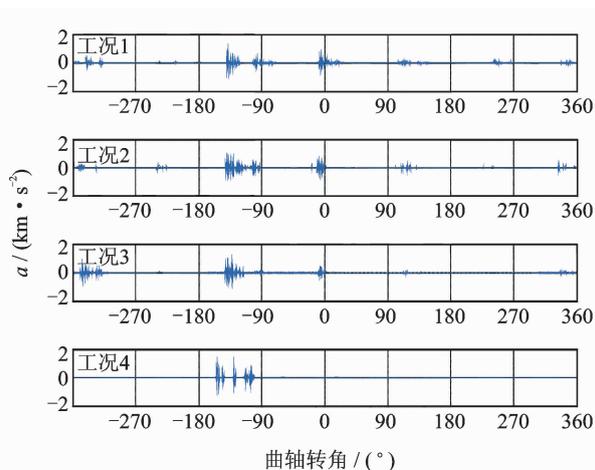


图5 信号时域波形

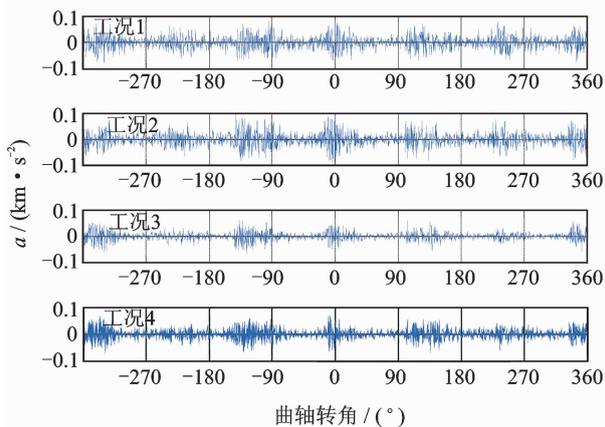
Fig. 5 Time domain waveform of signal

用AMP算法对振动响应信号进行分解和重构,迭代200次后残余信号能量达到原信号能量的5%以内。4种工况分解的重构信号和残余信号如图6所示。可以看到,原信号中的所有冲击分量均得到了很好的匹配。在时频域上对振动信号进行分析,将AMP分解原子的Wigner-Ville分布进行叠



(a) 重构信号

(a) Reconstructed signal



(b) 残余信号

(b) Residual signal

图6 4类工况信号分解重构信号及残余信号

Fig. 6 Signal construction and residues of four states

加得到各工况的时频分布图像,如图7所示。

对柴油机振动信号进行AMP时频表征,4类工况信号分别耗时9.022,8.678,8.869和8.955s,平均耗时8.881s,而相同条件下对振动信号直接进行WVD时频表征,4类工况信号平均耗时814.65s,说明AMP时频分析方法更能够满足柴油机故障诊断的时效性。从图7可以看出,正常工况对应的频率能量较大,曲轴转角 0° 处对应能量更高,说明混合气体燃烧效率更高。状态2,3,4对应的频率能量小,说明气体燃烧做功不充分。对于状态2,4,由于气门漏气会导致燃烧不充分,状态4的漏气最严重,所以状态4的燃烧功率最低。气门间隙故障不但影响混合气体的燃烧效率,而且影响气门落座对缸盖的冲击。对比状态2,3和4曲轴转角 -310° 处频率能量,状态3的频率能量明显大于状态2,可以知道状态3的排气门间隙大,而状态2的排气门间隙小。

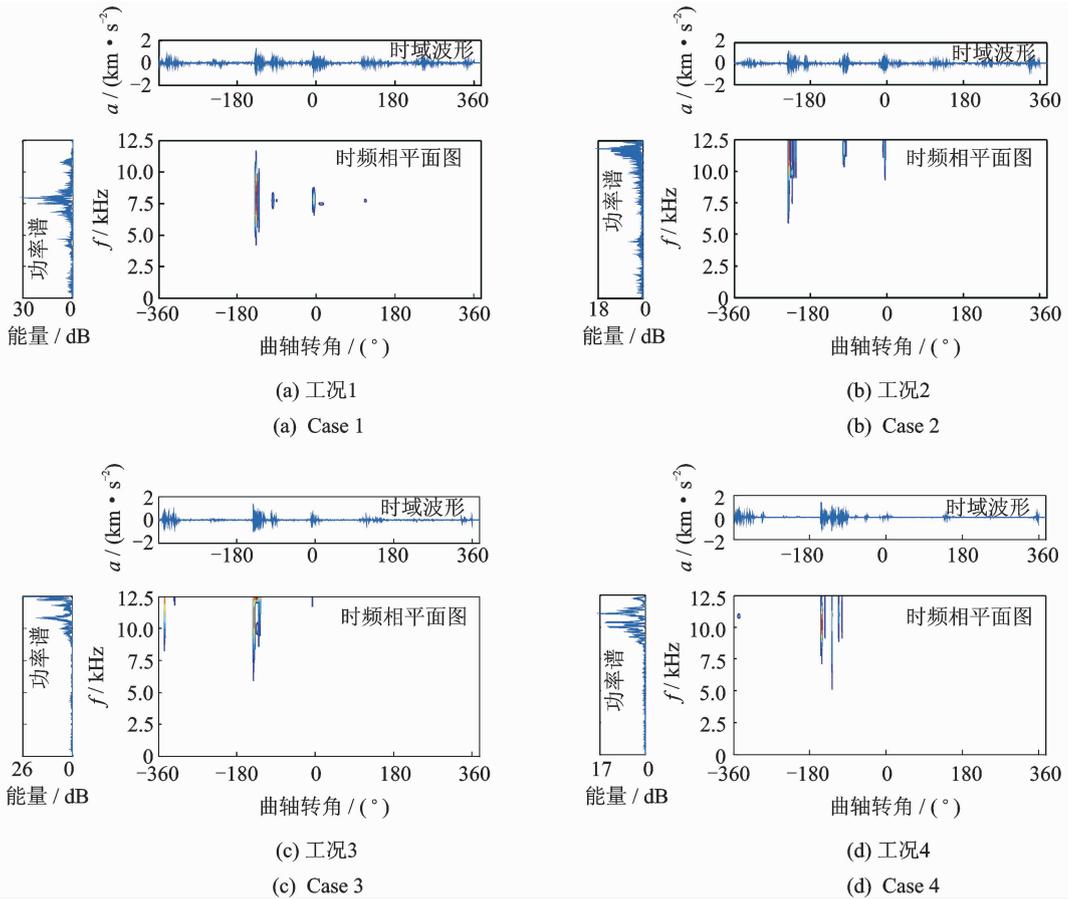


图 7 4 类工况的 AMP 时频分布情况

Fig. 7 AMP time-frequency distribution of four states

通过以上分析,时频图像清楚地表明不同间隙的气门落座冲击以及混合气体的燃烧效率所占的频率分量是不同的。因此,柴油机振动信号的稀疏分解时频图像能够较好地反映出故障特征。但是,以上分析需要一定的先验知识做前提,要依靠计算机实现故障的自动诊断,需进一步提取时频图像中的特征参量。

3.3 TD2DNMF 时频特征编码提取与故障识别

取采集到的 240 个信号作为研究对象并分别绘制时频图像,由于彩色图像对应的是三维矩阵,为了便于计算机处理,对图像进行预处理,采用阈值平均法将其转化为灰度图像,相应得到 240 个 420×560 像素点的矩阵样本。

从 4 类工况时频分布图中每一类随机选取 30 幅组成训练样本集 $\mathbf{X}_{420 \times 67200}$,其余 120 幅图像样本组成测试集;对样本集 \mathbf{X} 进行 TD2DNMF 特征提取,得到最优行基投影矩阵 $\mathbf{L}_{420 \times R}$ 和系数矩阵 $\mathbf{H}_{R \times 67200}$ 。样本类别数 $k=4$, R 为 TD2DNMF 提取的特征维数,其取值对特征提取效果有较大影响。

将系数矩阵 \mathbf{H} 按式(17),(18)进行重排,得到新的矩阵 $\mathbf{H}^T_{560 \times 480R}$ 。对 \mathbf{H}^T 进行 TD2DNMF 特征提取,得到最优列基投影矩阵 $\mathbf{R}_{560 \times d}$ 和系数矩阵 $\mathbf{W}_{d \times 480R}$ 。 d 同 R 一样,为 TD2DNMF 提取的特征维数,所不同的是 R 需要为 $k=4$ 的整数倍, d 无此限制。将 240 幅时频图像分别向行基矩阵 \mathbf{L} 和列基矩阵 \mathbf{R} 投影,得到其对应的编码矩阵 $\mathbf{F}_{R \times d}$,每个编码矩阵 \mathbf{F} 代表了所对应的时频图像。

图 8 为特征维数 $R=d=8$ 时,柴油机 4 类工况振动信号时频分布图像测试集对应的特征编码。图中每个像素的幅值严格与样本系数值对应,由于文章篇幅有限,每种工况下选取 5 个样本显示,每个样本编码图的横、纵坐标仅代表矩阵维度。图中每一行代表一种柴油机工况,从上到下依次为气门间隙正常、过小、过大和严重漏气工况。可以看出,TD2DNMF 对数据进行了有效降维,将 420×560 维数据压缩到 8×8 维,大大降低了模式识别的计算复杂度。同种工况的编码矩阵类内差异小,不同工况间编码矩阵的类间差异较大。对图像进行 TD2DNMF 特征提取,在对数据进行有效降维的同

时,不同工况对应时频图像间的差异化信息得到了较大程度的保留,有利于提高故障模式识别的正确率。

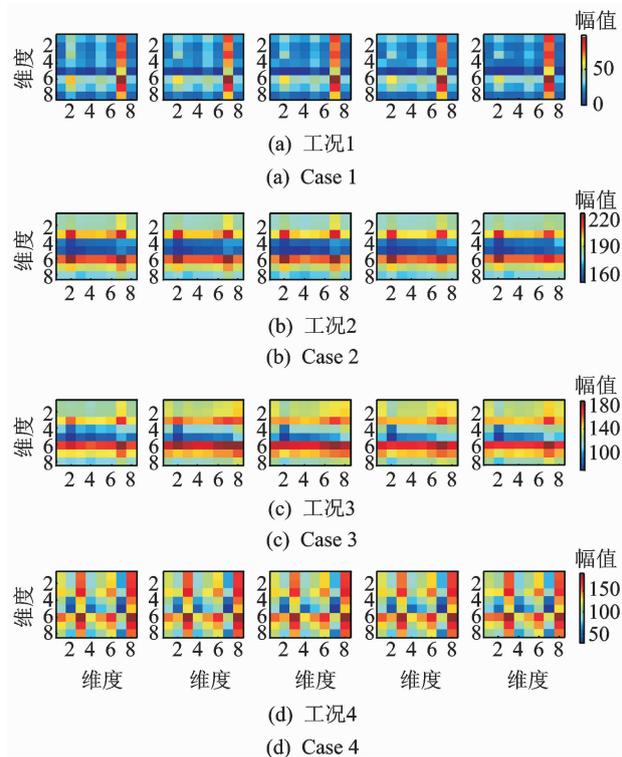


图 8 4 类工况时频图像特征编码

Fig. 8 Feature code of time-frequency distribution of 4 states

为比较不同算法的计算效率,分别利用传统非负矩阵分解算法、文献[10]提出的二维非负矩阵分解算法(2-dimensional NMF-Zhang,简称 2DNMF-Z)、文献[11]提出的二维非负矩阵分解算法(2-dimensional NMF-Gu,简称 2DNMF-G)、文献[12]提出的并行二维非负矩阵分解算法((2D)²NMF)以及笔者提出的 TD2DNMF 算法对 4 种工况所有样本进行特征提取。NMF 特征维数设定为二维方法编码矩阵行、列维数的乘积。4 种算法统一最大迭代步长为 100,目标函数容忍误差为 10^{-5} 。表 3 为 5 种算法特征提取的计算时间,均不包含图像载入的时间。试验环境为 Matlab R2012b,AMD A8 处理器,1.90 GHz 主频 CPU,4GB 内存,Win7 操作系统。

由表 3 可见,随着特征维数的增加,5 种算法特征提取时间整体呈增长的趋势。二维 NMF 的计算效率明显要高于一维 NMF,一维 NMF 对图像矩阵进行向量化后数据矩阵维度过大。例如,笔者所用图像维度为 420×560 ,120 幅训练图像组成的数据

矩阵维度为 $120 \times 235\ 200$,计算任务量十分繁重。二维 NMF 算法中,2DNMF-G 的计算效率略低于其他几种,这是由于 2DNMF-G 在计算时将原始图片分别按行、列拼合,初始分解矩阵维数分别为 $420 \times 67\ 200$ 和 $560 \times 50\ 400$,数据维度依然很大。(2D)²NMF 算法在初始分解矩阵的组成上与 2DNMF-G 相同,但是采用并行运算的方法,计算效率有所提高。2DNMF-Z 算法的初始分解矩阵仅将图像按行拼接,再将一次 NMF 分解后得到的系数矩阵分块转置再拼接后作为新的初始分解矩阵进行求解,两个初始分解矩阵的维度分别为 $420 \times 67\ 200$ 和 560×960 ,维度小于 2DNMF-G 与 (2D)²NMF,计算效率与后者相当。本研究方法 TD2DNMF 将图像矩阵按行拼接后按照各自的类别进行分解,将得到的 4 个系数矩阵各自分块转置再拼接,组成新的初始分解矩阵,两次分解的数据矩阵维度分别为 $420 \times 67\ 200$ 和 560×120 。与其他几类算法相比,本研究方法初始分解矩阵维度更低,有效提高了算法的计算效率。

为比较 5 种方法特征提取效果,采用最近邻分类器(nearest neighbor classifier,简称 NNC)对提取到的特征参数进行模式识别。诊断时分别采用柴油机振动信号的 Wigner-Ville 时频分布图和 AMP 稀疏分解时频分布图的 240 个样本,每种工况 60 个样本,从中随机抽取 30 个作为训练样本,其余作为测试样本。以 NNC 分类正确率评价 5 种方法的特征提取效果,为减小随机误差,试验重复 10 次取平均准确率作为最终结果,如图 9 所示。

表 3 不同方法特征提取效率对比

特征维数	特征提取方法				
	NMF	2DNMF-Z	2DNMF-G	(2D) ² NMF	TD2DNMF
8(4×2)	304.8	84.8	105.9	80.06	46.2
16(4×4)	367.2	86.9	106.3	81.43	47.3
24(8×3)	408.6	94.1	133.7	87.26	57.9
32(8×4)	463.8	98.1	135.2	93.5	59.8
40(8×5)	622.2	98.4	134.7	95.4	62.4
48(8×6)	882.8	98.7	136.9	96.3	63.9
56(8×7)	921.4	99.4	137.2	98.6	62.5
64(8×8)	960.9	99.6	136.5	99.7	64.8
72(8×9)	1002.5	99.3	140.9	100.3	66.5
80(16×5)	1021.8	118.9	164.5	114.8	72.6

图 9(a)可以看出,使用 3 种分类方法对 WVD 时频分布图进行特征提取,其中 NMF 的识别率相对较低,2DNMF-G 和 (2D)²NMF 的识别准确率相

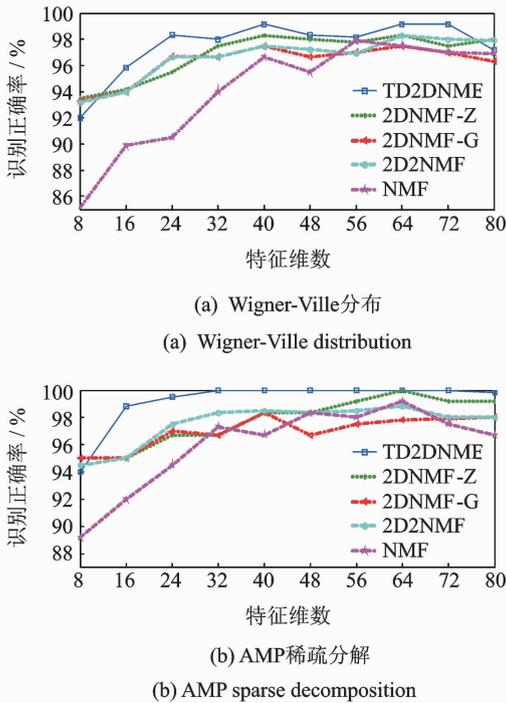


图9 不同特征提取方法分类性能对比

Fig.9 Performance of classifier based on different feature extraction methods

差不多,各个特征维度的识别准确率都在90%~98%之间,特征维数大于32时2DNMF-Z的识别率稍高,整体来看TD2DNMF特征提取效果最好。但是由于WVD分布中交叉项的干扰,会对故障类别带来较大影响,所以其识别正确率最高为99.17%,稳定于98%上下。图9(b)中,3种分类方法对AMP稀疏分解时频分布图进行特征提取,识别率与图9(a)中相比,5种方法均有所提高。这是由于采用AMP对柴油机振动信号进行时频分析时,生成的时频分布时频聚集性更好,各个工况间的差异更明显,也更利于分类器的分类。TD-2DPCA在特征矩阵维度为 8×4 和更高维度时,识别准确率稳定在100%,直到特征维数为80时识别率才降为99.83%。对比图9(a)和(b)可以发现,采用基于AMP稀疏分解与TD2DNMF特征编码的故障诊断方法更适用于柴油机气门间隙故障的诊断,并具有较高的诊断精度,且对于一个测试信号样本而言,整个诊断过程运行时间平均为9.483s,运行效率较高。

为了与图像方法形成对比,笔者在振动信号的频域进行故障特征的提取。利用经验模态分解方法(empirical mode decomposition,简称EMD)将信号分解为有限阶的内禀模态函数(intrinsic mode

function,简称IMF)之和,剔除伪分量后保留了前三阶IMF。利用AR参数模型求解前三阶IMF的功率谱,采用最终预测误差准则判断最优阶次,使用Yule-Walker方法估计AR模型参数,将每阶IMF的前8个AR参量作为特征向量,仍使用NNC作为分类器进行模式识别,训练集、测试集样本个数与前文保持一致,平均识别正确率为86.67%。相比而言,振动信号的时频图像中既包含了信号时域中各时频分量的产生及消亡时刻的信息,又包含了频域中信号的能量分布信息,对柴油机振动信号非平稳时变特征具有较强的刻画能力。

4 结论

1) AMP算法利用随分解残差信号自适应更新的Gabor字典,通过原子与残留信号的互相关运算选择最优原子,与原始MP算法相比,运算速度得到了大幅提高,同时也保持了良好的稀疏性能。

2) TD2DNMF算法将数据矩阵行、列维信息融合到一个判别分析框架中,将不同类别的数据信息并行运算,对柴油机振动时频图像样本进行特征编码。与现有几种二维NMF分解算法相比,TD2DNMF算法能更有效地提取柴油机振动时频图像的差异特征,计算效率也得到了提高。

3) 利用AMP算法分解振动信号,叠加匹配原子的Wigner-Ville分布所得到的时频分布具有很高的时、频分辨率且无交叉项的干扰。用该方法对柴油机不同气门间隙工况的振动信号进行分析,各工况的时频分布特征明显,时频分量物理意义明确。AMP稀疏分解与TD2DNMF特征编码相结合的故障诊断方法用于柴油机故障诊断中可获得较高的诊断精度,是一种有效的机械设备故障诊断方法。

参 考 文 献

- [1] Komorska I. Adaptive model of engine vibration signal for diagnostics of mechanical defects[J]. *Mechanics*, 2013, 19(3): 301-305.
- [2] 沈虹,赵红东,梅检民,等.基于高阶累积量图像特征的柴油机故障诊断研究[J]. *振动与冲击*, 2015, 34(11): 133-138.
Shen Hong, Zhao Hongdong, Mei Jianmin, et al. Diesel engine fault diagnosis based on high-order cumulant image features [J]. *Journal of Vibration and Shock*, 2015, 34(11):133-138. (in Chinese)
- [3] 刘昱,张俊红,毕凤荣,等.基于Wigner分布和分形

- 维数的柴油机故障诊断[J]. 振动、测试与诊断, 2016, 36(2): 240-245.
- Liu Yu, Zhang Junhong, Bi Fengrong, et al. Study on fault diagnosis of diesel valve trains based on Wigner distribution and fractal dimension[J]. Journal of Vibration, Measurement & Diagnosis, 2016, 36(2): 240-245. (in Chinese)
- [4] 蔡艳平, 李艾华, 王涛, 等. 基于时频谱图与图像分割的柴油机故障诊断[J]. 内燃机学报, 2011, 29(2): 181-186.
- Cai Yanping, Li Aihua, Wang Tao, et al. Diesel engine fault diagnosis based on time-frequency spectrum image and image segmentation [J]. Transactions of CSICE, 2011, 29(2): 181-186. (in Chinese)
- [5] 蔡艳平, 李艾华, 王涛, 等. 基于 EMD-Wigner-Ville 的内燃机振动时频分析[J]. 振动工程学报, 2010, 23(4): 430-437.
- Cai Yanping, Li Aihua, Wang Tao, et al. Engine vibration time-frequency analysis based on EMD-Wigner-Ville [J]. Journal of Vibration Engineering, 2010, 23(4): 430-437. (in Chinese)
- [6] Feng Zhipeng, Liang Min, Chu Fulei. Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples [J]. Mechanical Systems and Signal Processing, 2013, 38(1): 165-205.
- [7] Subhash C K S, Ashutosh M, Vinay S, et al. Comparison of matching pursuit algorithm with other signal processing techniques for computation of the time-frequency power spectrum of Brain Signals[J]. Neurosci, 2016, 36(12): 3399-3408.
- [8] Yang Yongsheng, Ming Anbo, Zhang Youyun, et al. Discriminative non-negative matrix factorization (DNMF) and its application to the fault diagnosis of diesel engine[J]. Mechanical Systems and Signal Processing, 2017, 10(95): 158-171.
- [9] 梁霖, 栗茂林, 李利邦, 等. 基于非负矩阵分解的单通道故障特征分离方法[J]. 振动、测试与诊断, 2016, 36(5): 835-840.
- Liang lin, Li Maolin, Li Libang, et al. Separation of fault features from a single channel mechanical signal using non-negative matrix factorization [J]. Journal of Vibration, Measurement & Diagnosis, 2016, 36(5): 835-840. (in Chinese)
- [10] Zhang Daoqiang, Chen Songcan, Zhou Zhihua. Two-dimensional non-negative matrix factorization for face representation and recognition [J]. Analysis and Modelling of Faces and Gestures, 2005, 3723: 350-363.
- [11] Gu Quanquan, Zhou Jie. Two dimensional nonnegative matrix factorization [C] // 16th IEEE International Conference on Image Processing. Cairo, Egypt: IEEE, 2009: 2069-2072.
- [12] 高宏娟. 基于 NMF 改进算法的人脸识别仿真研究 [J]. 计算机仿真, 2011, 28(12): 231-235, 303.
- Gao Hongjuan. Research and simulation of face recognition based on nmf and its improvements [J]. Computer Simulation, 2011, 28(12): 231-235, 303. (in Chinese)
- [13] 张焱, 汤宝平, 刘自然, 等. 时频表示特征约简的旋转机械故障特征提取方法[J]. 振动工程学报, 2015, 28(1): 156-163.
- Zhang Yan, Tang Baoping, Liu Ziran, et al. Rotating machine fault feature extraction based on reduced time frequency representation [J]. Journal of Vibration Engineering, 2015, 28(1): 156-163. (in Chinese)
- [14] 李兵, 米双山, 刘鹏远, 等. 二维非负矩阵分解在齿轮故障诊断中的应用[J]. 振动、测试与诊断, 2012, 32(5): 836-840.
- Li Bing, Mi Shuangshan, Liu Pengyuan, et al. Application of two-dimensional non-negative matrix factorization for gear fault diagnosis [J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(5): 836-840. (in Chinese)
- [15] Mallat S, Zhang Zhifeng. Matching pursuits with time-frequency dictionaries[J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3397-3415.
- [16] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(21): 788-791.



第一作者简介:王旭,男,1992年10月生,硕士。主要研究方向为机械设备状态监测与故障诊断。曾发表《内燃机变分模态 Rihaczek 谱纹理特征识别诊断》(《仪器仪表学报》2017年第38卷第10期)等论文。

E-mail: 15829220894@163.com

通信作者简介:岳应娟,女,1972年10月生,博士、教授。主要研究方向为压力容器检测与机电设备故障诊断。

E-mail: yingjuanyue@163.com

