

短时傅里叶变换的时频聚集性度量准则研究*

赵学智, 叶邦彦, 陈统坚

(华南理工大学机械与汽车工程学院 广州, 510640)

摘要 为了使短时傅里叶变换(short time Fourier transform,简称 STFT)获得良好的时频聚集性,必须根据信号的具体特点来选择窗长。分析了现有窗长选择准则的选择机理及其优缺点,发现归一化 3 阶 Renyi 熵准则与 Stankovic 准则一般只会取到窗长选择范围的最大值,并分析了出现这种问题的原因,而其他准则得到的也不是最优的窗长。提出了一种新的基于对数窗能量的窗长选择准则。对数窗能量与窗长是一种非线性关系,这显著区别于普通窗能量随窗长线性增长的特性,且其增长速度与窗型无关,并在短窗和长窗具有不同的增长速度,因而能够在短窗和长窗之间取得良好的折衷。提供了仿真信号和实际信号的处理实例,其结果证明对数窗能量准则使 STFT 获得了良好的时频聚集性,效果优于现有的窗长选择准则。

关键词 短时傅里叶变换;时频分析;故障诊断;窗长选择;时频聚集性

中图分类号 TH113; TN911.7

引言

短时傅里叶变换是一种重要的时频分析方法,尽管现在小波分析^[1]、经验模态分解^[2]和局部均值分解^[3]等方法很流行,但并不能取代 STFT,因为无论是小波分析,还是经验模态分解或局部均值分解,其实质都是将原信号分解为一系列不同频段的分量信号,而确知这些分量信号的具体频率成分最终还是要通过傅里叶变换,对这些分量信号也都可以采用 STFT 做进一步分析。STFT 在信号处理、系统状态监测、故障诊断和瞬时频率估计等方面应用广泛。Astuti 等^[4]利用地电场信号的 STFT 分析结果来预测地震。Sivasankari 等^[5]利用 STFT 与独立分量分析相结合,提取癫痫病人脑电信号的特征,在心电信号的处理中,STFT 可用来计算各种心音分裂类型的分裂时间^[6]。此外,STFT 还被应用于对多普勒超声血栓信号的处理,以实现了对脑血栓的探测和定位^[7]。STFT 还可用于对语音的增强^[8]。在故障诊断方面,通过计算 STFT 结果的峭度并选择最大峭度所对应频段的信号进行解调分析,可以获取轴承的故障频率^[9]。另外,在双曲调频信号的解调和瞬时频率估计等方面 STFT 都有着重要的应用^[10]。

当应用 STFT 时,首先面临的一个重要问题就是窗长选择。在上面列举的所有 STFT 应用中,有的是根据经验来选择窗长,有的则是根据已有的选择准则来选取窗长。窗长是决定 STFT 时频分析效果的关键因素。对 STFT 来说,窗太短则 STFT 的频率分辨率会很差,窗太长则时间分辨率会很差。为了同时在时域和频域都获得良好的分辨率,即良好的时频聚集性,必须折衷地选择窗长。窗长的选择一直是 STFT 中具有挑战性的问题,尽管现在已经有一些选择准则,但是这些准则在实用中的效果并不太好,有些准则在很多情况下只能取到窗长选择范围的最大值,而其他准则确定出来的窗长也并不能使 STFT 取得很好的时频聚集性。

笔者总结了目前的 5 种窗长选择准则,提出了一种新的基于对数窗能量的窗长选择准则。与目前的窗长选择准则相比,对数窗能量的增长速度与窗型无关,而仅由窗长决定,因而对不同的窗型具有鲁棒性。它的另一显著特点是:与普通的窗能量随窗长线性增长的特性相比,对数窗能量在短窗时的增长速度相对较快,而在长窗时的增长速度则相对变缓,因而在短窗和长窗之间能够取得良好的折衷,有效地改变了以往准则的窗长选择曲线随窗长的单调下降性,在短窗和长窗之间实现了折衷的窗长选择。实例分析表明,基于对数窗能量的窗长选择准则能

* 国家自然科学基金资助项目(51275174,51375178);广东省自然科学基金资助项目(S2012010008789)
收稿日期:2015-09-10;修回日期:2015-11-20

够在不同情况下稳定地取得最优窗长。

1 短时傅里叶变换窗长选择的必要性

STFT 的实质是利用一个移动的窗函数将信号在不同时刻的部分依次取入窗口中,然后对窗口内的信号进行傅里叶变换,随着窗口的不断移动,就可分析出信号在不同时间点的频率情况。STFT 的定义为

$$G(t, f) = \int_R x(\tau) \omega(\tau - t) e^{-j2\pi f\tau} d\tau \quad (1)$$

其中: $x(\tau)$ 为待分析的信号; $\omega(\tau)$ 为窗函数。 $G(t, f)$ 反映了信号 $x(\tau)$ 在时刻 t 的频谱。

对于离散的数字信号 $x(k)$, 设窗函数的窗长为 N , 两相邻窗的移动步长为 s 个采样时间间隔, 则离散的 STFT 为

$$G(m, n) = \sum_{k=ms}^{ms+N-1} x(k) \omega(k - ms) e^{-j\frac{2\pi kn}{N}} \quad (2)$$

$(m=0, 1, 2, \dots, M-1; n=0, 1, 2, \dots, N-1)$

其中: m 为窗口个数变量, 它对应 STFT 的时间参数; n 对应频率参数。

设原始信号的采样频率为 f_s , 则 $G(m, n)$ 为信号 $x(k)$ 在时刻 ms/f_s 的频谱, 其中每一个 n 对应的频率为 nf_s/N 。

对于 STFT 来说, 窗长 N 为决定其时频分析效果的一个最重要参数, 尤其是对于时变信号, 窗长的选择决定着能否正确分析出原始信号中的频率变化情况。STFT 的频率分辨率为 f_s/N , 显然窗长 N 越大, STFT 的频率分辨率就越高, 但同时时间分辨率却越差, 尤其是对于那些存在局部变化的信号, 只有采用短窗才能捕捉到信号中的局部变化及其发生时刻, 另外太长的窗长将使 STFT 失去“短时”的意义, 但是太短的窗长却又使得频率分辨率变得很粗糙。为了同时在时域和频域获得较好的分辨率, 即较好的时频聚集性, 必须折衷地选择窗长。

2 窗长选择准则及其选择机理分析

2.1 5 种窗长选择准则

1) 最大模值准则^[11]。Daubechies 是著名的小波分析专家, 以其名字命名的 Daubechies 小波应用广泛, 并对 STFT 也颇有研究, 提出了窗长选择的最大模值准则。STFT 结果的模值之和为

$$D(N) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)| \quad (3)$$

则使 $D(N)$ 取得最大值的 N 就是最优窗长, 这就是 Daubechies 提出的最大模值准则^[11]。

2) 最大峭度准则^[12]。Jones 等根据统计学的峭度概念, 通过计算 $G(m, n)$ 的峭度来选择最优窗长

$$K(N) = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^4}{\left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2 \right)^2} \quad (4)$$

在不同的窗长中, Jones 等^[12] 认为使峭度 $K(N)$ 取得最大值的窗长 N 是最优的。

3) Renyi 熵准则^[13]。Renyi 熵定义为

$$R_\alpha(N) = \frac{1}{1-\alpha} \log_2 \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^\alpha \right) \quad (\alpha \geq 2) \quad (5)$$

在应用中一般取 $\alpha=2$ 。当 Renyi 熵值较小时, 表明 STFT 的时频聚集性较好, 而聚集性差时的 Renyi 熵值较大。当 Renyi 熵取得最小值时, 此时的窗长是最优的^[13]。

4) 归一化 3 阶 Renyi 熵准则^[14]。归一化 3 阶 Renyi 熵的定义为

$$RV_3(N) = -\frac{1}{2} \log_2 \left(\frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^3}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|} \right) \quad (6)$$

式(6)与式(5)的主要区别是在对数当中除除了 $G(m, n)$ 的模值之和, 并且取式(5)中的 $\alpha=3$ 。此准则认为当 $RV_3(N)$ 取得最小值时的窗长是最优的^[14]。

5) Stankovic 准则^[15]。以上 4 种方法都只针对 STFT 的结果 $G(m, n)$ 进行各种运算, 以便由此确定一个最佳的窗长。与这些方法有重要区别的是, Stankovic 在选择窗长时引入了窗函数的能量, 提出利用式(7)来度量 STFT 的时频聚集性

$$H(N) = \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \left(\frac{1}{E(N)} |G(m, n)|^2 \right)^{\frac{1}{2}} \right)^2 \quad (7)$$

其中: $E(N)$ 为窗函数的能量, $E(N) = \sum_{k=0}^{N-1} |\omega(k)|^2$; $\omega(k)$ 为窗函数。

Stankovic 准则认为 $H(N)$ 越小, 则 STFT 的时频聚集性越好^[15]。

2.2 5 种窗长选择准则的选择机理分析

以上 5 种准则都涉及到对 $G(m, n)$ 的 1~4 次

方进行求和运算,为了从本质上探讨这些准则的选择机理,首先来分析 $G(m, n)$ 求和的实质。对于第 m 个窗,窗口里面的局部信号是 $x(k)\omega(k - ms)$, 它的傅里叶变换是 $G(m, n)$,则由离散傅里叶逆变换知,它可由式(8)重构

$$x(k)\omega(k - ms) = \frac{1}{N} \sum_{n=0}^{N-1} G(m, n) e^{\frac{j2\pi nk}{N}} \quad (8)$$

$(k = ms, ms + 1, ms + 2, \dots, ms + N - 1)$

对式(8)两边乘以 $x(k)\omega(k - ms)$, 然后对 k 求和

$$\begin{aligned} & \sum_{k=ms}^{ms+N-1} [x(k)\omega(k - ms)]^2 = \\ & \sum_{k=ms}^{ms+N-1} \left[x(k)\omega(k - ms) \frac{1}{N} \sum_{n=0}^{N-1} G(m, n) e^{\frac{j2\pi nk}{N}} \right] = \\ & \frac{1}{N} \sum_{n=0}^{N-1} G(m, n) \sum_{k=ms}^{ms+N-1} x(k)\omega(k - ms) e^{\frac{j2\pi nk}{N}} = \\ & \frac{1}{N} \sum_{n=0}^{N-1} G(m, n) \sum_{k=ms}^{ms+N-1} \overline{x(k)\omega(k - ms) e^{-\frac{j2\pi nk}{N}}} = \\ & \frac{1}{N} \sum_{n=0}^{N-1} G(m, n) \overline{G(m, n)} = \frac{1}{N} \sum_{n=0}^{N-1} |G(m, n)|^2 \end{aligned}$$

在两边同时对所有的 m 求和,可得

$$\sum_{m=0}^{M-1} \sum_{k=ms}^{ms+N-1} [x(k)\omega(k - ms)]^2 = \frac{1}{N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2$$

即

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2 = N \sum_{m=0}^{M-1} \sum_{k=ms}^{ms+N-1} [x(k)\omega(k - ms)]^2 \quad (9)$$

其中: M 为STFT中窗口的总个数。

从式(9)可见, $G(m, n)$ 模值平方和的实质是时域中所有窗口内信号的能量之和的 N 倍。因为总是有 $|G(m, n)| \geq 0$, 因此 $|G(m, n)|$, $|G(m, n)|^3$, $|G(m, n)|^4$ 的单调性与 $|G(m, n)|^2$ 是一致的, 因此若

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2 \text{ 越大, 则相应地, } \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|, \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^3, \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^4 \text{ 也会越大。}$$

式(9)表明, $\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2$ 的实质就是所有窗口内信号总能量的 N 倍, 显然这总能量取决于两个参数: 一是窗长 N , 窗长 N 越大, 则取入此窗口内的信号数据越多, 其能量就越大; 二是窗的个数 M , 窗个数越多, 则总能量越大, 但是窗个数 M 却受窗长 N 制约。设信号 $x(k)$ 的长度为 L (注意它与窗长 N 的区别), 两相邻窗的距离为 s 个采样间隔, 如图1所示, 则窗口的个数 M 为

$$M = \text{int}\left(\frac{L - N}{s}\right) + 1 \quad (10)$$

其中: $\text{int}(\quad)$ 表示取整。

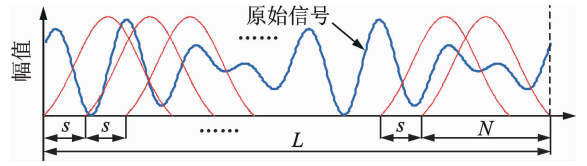


图1 窗长 N 、窗位移 s 与信号长度 L 的关系

Fig.1 Relationship among window length N , window shift s and signal length L

对于参数 s 来说, 虽然 s 会影响窗口的个数, 但是实际处理结果表明此参数对 STFT 时频聚集性的影响并不显著, 在计算 STFT 时, 一般取 $s = 1$ 。则从式(10)可见, 对于固定的 s , 窗个数 M 与窗长 N 成反比, 因此大的窗长 N 虽然使窗口内信号的能量增加, 但同时却会使窗口的个数 M 减少, 则 $\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2$ 又会因 M 减少而下降, 即所有窗口内信号的总能量又会有所降低。这两个参数综合的结果, 就是可能会存在一个折衷的 N 使得 $\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)|^2$ 最大, 因此对于 Daubechies 的最大模值准则和 Renyi 熵准则来说, 可能会取到一个折衷的窗长 N , 虽然这个窗长在实际当中未必是最优的。但是对于其他三个准则(即最大峭度准则、归一化 3 阶 Renyi 熵准则以及 Stankovic 准则)来说, 则未必能取到一个折衷的 N , 有可能总是只取到可选范围的最大值或最小值。以 Stankovic 准则为例来进行说明。

2.3 窗能量增长速度对选择曲线单调性的影响

Stankovic 准则的显著优点是引入了窗能量的倒数, 在目标函数中除以窗能量是一个很好的思想, 因为显然短窗能量小, 长窗能量大, 因此除以这种窗能量之后可以使短窗的 $H(N)$ 相对增大, 而长窗的 $H(N)$ 相对减小, 从而使最终的选择在短窗和长窗之间可以有一个更好的折衷。虽然表面上看起来是这样, 但是这一方法在很多情况下往往会失效, 得不到最优的窗长。可以对这一问题进行分析, 首先将式(7)简化为

$$H(N) = \frac{1}{E(N)} \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)| \right)^2 \quad (11)$$

上式中涉及到窗能量 $E(N)$, 对于常见的几种窗函数, 利用它们的表达式可以通过理论计算得到它们的能量 $E(N)$ 和窗长 N 的关系, 这里不给出计

算过程,只把结果列在表 1 中。从表 1 可见,所有窗函数的能量都与窗长成正比,即窗函数的能量 E 随窗长是线性增长的,但问题是对于很多实际信号来说,Stankovic 准则中的另一个因素即分子 $(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m,n)|)^2$ 随窗长 N 增加的速度在很多情况下赶不上 $E(N)$ 的那种线性增长速度,甚至当 N 增加时它可能还会减小(因为窗个数 M 会减少)。在这些情况下除以窗能量的后果就是 $H(N)$ 将变为一条单调下降的曲线,即 N 越大, $H(N)$ 越小,这样总是最大的窗长 N 被选为最优窗长,但这往往是错误的结果。

表 1 窗长为 N 的各种窗函数能量

Tab. 1 Energy of different window functions with length N

窗函数类型	窗函数能量 $E(N)$ /无量纲
Bartlett 窗(三角窗)	$(N-1)/3$
Hanning 窗	$0.375\ 0(N-1)$
Hamming 窗	$0.397\ 4(N-1)$
Blackman 窗	$0.304\ 6(N-1)$
Gauss 窗	$\sqrt{\pi}(N-1)/6$

很容易给出这样一个例子。对于函数 $x(t) = \cos(6\cos(3\pi t) + 60\pi t) + \cos(18\pi t^2 + 120\pi t)$ 以采样间隔 $T_s = 1/256$ 采集 512 点,即信号长度 $L = 512$,窗函数选用 Hanning 窗,窗长选择范围为 $8 \sim 256$,

窗位移 $s = 1$,令 $P(N) = (\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m,n)|)^2$,计算不同窗长下的 $P(N)$ 曲线、窗能量 $E(N)$ 曲线以及 $H(N)$ 曲线,所得结果如图 2 所示。可见 $P(N)$ 本来是一个先持续增长、后略为下降的曲线,但是除以窗能量 $E(N)$ 后,尽管 $P(N)$ 比 $E(N)$ 大了好几个数量级,但是它们相除得到的 $H(N)$ 曲线却并没有保持 $P(N)$ 的那种曲线特性,而是变成了一个纯粹的单调下降曲线。按照式(7)的准则,此时的最优窗长为 $N = 256$,但这不是正确的结果。

上面的结果不是偶然的,尽管 $P(N)$ 的数量级比 $E(N)$ 大很多,但是窗能量却是线性增长的,而 $P(N)$ 则未必,只要 $P(N)$ 的增长率追赶不上窗能量 $E(N)$ 的增长率,则不管 $P(N)$ 的数量级比 $E(N)$ 大多少,两者相除得到的 $H(N)$ 必将是成一个单调下降的曲线,证明如下。

$E(N)$ 的增长率为 $E(N)/E(N-1)$, $P(N)$ 的增长率为 $P(N)/P(N-1)$,如果 $P(N)$ 追赶不上 $E(N)$ 的增长率,则有

$$\frac{P(N)}{P(N-1)} < \frac{E(N)}{E(N-1)} \quad (12)$$

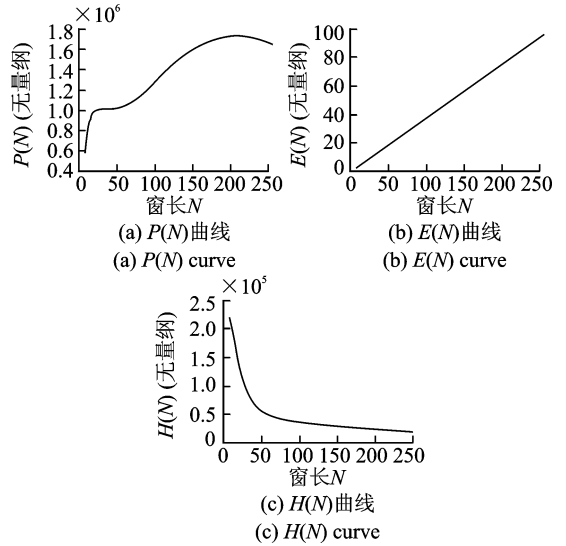


图 2 Stankovic 准则中的 3 条曲线

Fig. 2 Three curves in Stankovic rule

STFT 的窗长至少必须大于 1,从表 1 可见,各个窗的窗能量都是大于零的,即 $E(N) > 0$ 。另从 $P(N)$ 的表达式也可看出,所有的 $P(N)$ 也都是大于零的,即有 $P(N-1) > 0$ 。因此不等式(12)可移项,得

$$\frac{P(N)}{E(N)} < \frac{P(N-1)}{E(N-1)}$$

即

$$H(N) < H(N-1) \quad (13)$$

因此当窗长 N 增大时,不管 $P(N)$ 的数量级比 $E(N)$ 大多少,得到的 $H(N)$ 必定是一单调下降曲线。

对于归一化 3 阶 Renyi 熵准则(即式(6)),也存在类似的问题。从式(6)可以看出,其分子是 $|G(m,n)|$ 的立方和,而其分母仅仅是 $|G(m,n)|$ 的和,则其分子的增长率必然会大于分母的增长率,因此两者相除的结果必然是单调上升的。但由于式(6)前面有一个负号,因此最终得到的 $RV_3(N)$ 是随窗长单调下降的。由于该准则认为使 $RV_3(N)$ 取得最小值时的窗长是最优的,故该准则总是只能取到窗长选择范围的最大值。

对于最大峭度准则,由于其分子分母的幂次相同,难以确定它们的增长率孰快孰慢,但显然它们会随 $|G(m,n)|$ 的变化而变化,而不同信号的 $|G(m,n)|$ 是千变万化的,因此两者相除的结果比较复杂。

仍以前面的函数 $x(t) = \cos(6\cos(3\pi t) + 60\pi t) + \cos(18\pi t^2 + 120\pi t)$ 为例,给出其余 4 种准则的窗长选择曲线如图 3 所示。由图可见:Daubechies 最大模值准则的窗长选择曲线先上升然后有所下降;

最大峭度准则的选择曲线先是上升,后略微下降,再略微上升;Renyi 熵和 3 阶归一化 Renyi 熵都是单调下降的。除 Daubechies 最大模值准则得到的最优窗长是 $N=209$ 外,其余所有准则得到的最优窗长都是 $N=256$,但从后面的结果可以看到,它们都不是正确的最优窗长。

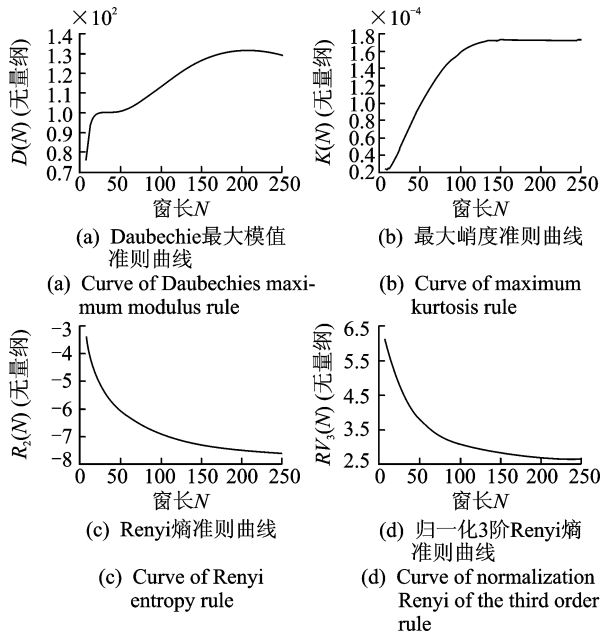


图 3 不同准则的窗长选择曲线

Fig. 3 Window length selection curves of different rules

3 基于对数窗能量的时频聚集性度量准则

3.1 对数窗能量增长速度与窗型的无关性

在上面已有的 5 种时频聚集性度量准则中,Stankovic 准则是比较独特的,其他的准则都只是对 $|G(m, n)|$ 本身进行各种运算,而 Stankovic 准则却引入了窗函数能量 $E(N)$ 因子。表面上看,短窗的 $E(N)$ 小,长窗的 $E(N)$ 大,因此除以窗能量后,短窗的 $H(N)$ 相对增大,长窗的 $H(N)$ 相对减小,从而似乎在短窗和长窗之间有一个更好折衷。但由于窗能量 $E(N)$ 随窗长 N 呈线性增长(见表 1),它在很多情况下会快于 $P(N)$ 的增长率,则从前面的证明过程可知,此时简单地将 $P(N)$ 除以 $E(N)$ 就会使 $H(N)$ 随 N 而单调下降。为了防止因 $E(N)$ 的线性增长而可能出现的这种问题,在除以窗能量之前,必须想办法降低 $E(N)$ 的那种线性增长速度,笔者提出采用对数函数来达到这一目的。如果对 $E(N)$ 取常用对数,则此时的对数窗能量随 N 的增长速度为

$$\frac{d\{\lg[E(N)]\}}{dN} = \frac{1}{E(N)\ln 10} \frac{dE(N)}{dN} \quad (14)$$

可见与自然窗相比,对数窗能量的增长速度下降了 $E(N)\ln 10$,并且这是一种非线性下降。例如对于 Hanning 窗,由表 1 知, $E(N)=0.3750(N-1)$,则从式(14)可得其对数窗能量的增长速度为 $\frac{1}{[(N-1)\ln 10]}$ 。实际上对表 1 的其他各窗,可以计算得到它们的对数窗能量的增长速度都是 $\frac{1}{[(N-1)\ln 10]}$ 。因此可以得到一个非常重要的结论:对数窗能量的增长速度与窗型无关,但却与窗长 N 成反比, N 越大,对数窗能量的增长速度越慢,这对长窗时的 $H(N)$ 的提升有显著效果;而 N 变小时,对数窗能量的增长速度会变快,这对短窗时的 $H(N)$ 的下降有显著效果,从而有效地改善了采用自然窗时 $H(N)$ 随 N 的那种单调下降性,使窗长在短窗和长窗之间有了一个更好的折衷。对数窗能量的增长速度与窗型无关的另一优势是:它对不同的窗型具有鲁棒性,可以有效可靠地适用于各种窗型。

相比之下,自然窗能量的增长速度与窗型有关,但却与窗长 N 无关,这使得它在长窗和短窗时的增长速度都是一样的。例如从表 1 可以得到: Hanning 窗能量的增长速度是 0.375 0, Hamming 窗能量的增长速度是 0.397 4, Blackman 窗能量的增长速度是 0.304 6, 高斯窗的增长速度是 0.295 4, 它们都与窗长没有关系。如果简单地除以自然窗能量,则当窗长变化时,这种恒定不变的增长速度不利于在长窗和短窗之间做折衷选择。图 4 给出了 Hanning 窗的自然窗和对数窗的能量及其增长速度与窗长 N 的关系对比,其他窗函数也存在类似关系。

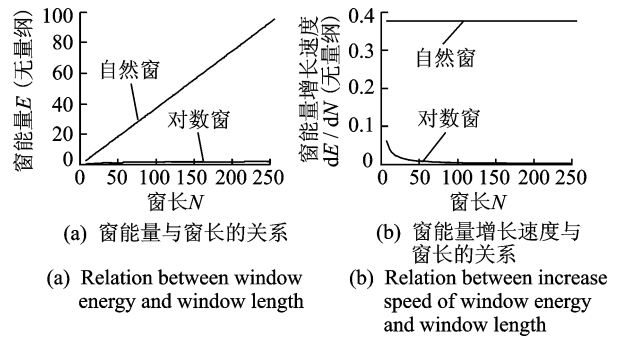


图 4 Hanning 窗的自然窗和对数窗能量及其增长速度与窗长的关系对比

Fig. 4 Relation between the energy, energy increase speed and length of natural and logarithm window of Hanning window

3.2 对数窗能量准则

根据以上分析,笔者提出一种新的时频聚集性度量准则:对长度为 L 的原始信号进行短时傅里叶变换时,采用下式进行时频聚集性度量

$$Q(N) = \frac{1}{\lg[E(N)]} \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |G(m, n)| \right)^2 \quad (15)$$

其中: M 为窗个数; N 为窗长; $E(N)$ 为窗能量,

$$E(N) = \sum_{k=0}^{N-1} |\omega(k)|^2; \omega(k) \text{ 为窗函数。}$$

当 $Q(N)$ 取得最小值时,则 STFT 的时频聚集性为最好,此时的窗长为最优窗长。此准则的特点在于引入了对数窗能量因子,因此称之为对数窗能量准则。

利用这一准则对前面的信号 $x(t)$ 进行最优窗长选择,窗长选择范围同样为 $8 \sim 256$,选择曲线如图 5 所示。可见,采用对数窗能量后, $Q(N)$ 不再像图 2 中的 $H(N)$ 那样单调下降,而是先下降后增长,然后再略为下降,有效地改变了 $H(N)$ 原来的那种单调下降性,在短窗和长窗之间取得了良好的折衷,这种结果证实了上面的分析。当窗长 $N=70$ 时, $Q(N)$ 取得最小值,因此最优窗长为 $N=70$ 。

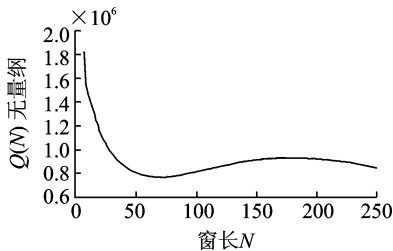


图 5 对数窗能量准则的窗长选择曲线

Fig. 5 Window length selection curve of logarithm window energy rule

3.3 不同准则的时频聚集性效果

对于本例所分析的仿真信号 $x(t)$,从上面可知,对数窗能量准则确定的最优窗长是 $N=70$, Daubechies 最大模值准则得到的最优窗长是 $N=209$,而其余所有准则得到的最优窗长都是 $N=256$ 。现分别取窗长 $N=70, 209, 256$ 对信号 $x(t)$ 进行短时傅里叶变换,结果如图 6 所示。

对于本例所分析的仿真信号 $x(t)$,由于其时域表达式是已知的,故其时频关系可以通过计算得到,该信号前半部分为余弦调制,其时频关系为

$$f_1(t) = \frac{1}{2\pi} \frac{d}{dt} [6\cos(3\pi t) + 60\pi t] = -9\sin(3\pi t) + 30$$

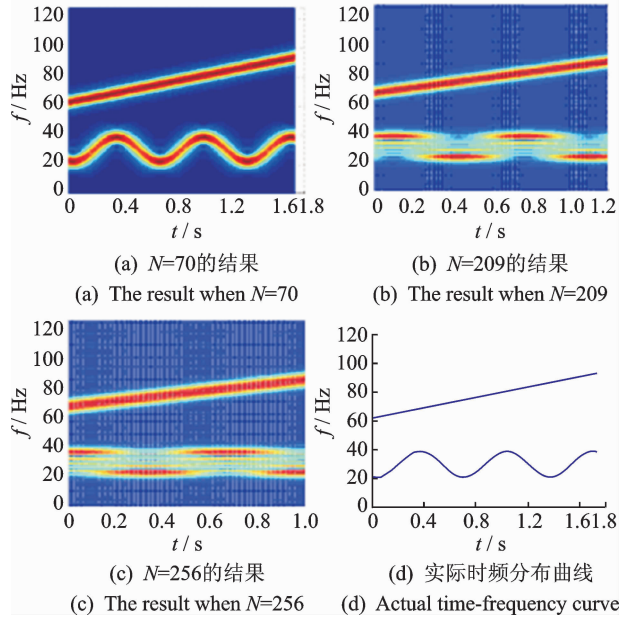


图 6 不同窗长下 STFT 的时频聚集性效果

Fig. 6 Time-frequency concentration effect of STFT under different window lengths

后半部分为线性调制,其时频关系为

$$f_2(t) = \frac{1}{2\pi} \frac{d}{dt} (18\pi t^2 + 120\pi t) = 18t + 60$$

根据以上两个时频关系可以画出 $x(t)$ 的实际时频分布。因为在 STFT 所得到的时频关系图中,其时间轴都是以第一个窗的中心所对应的时间为零时刻,因此在绘制 $x(t)$ 的实际时频分布时,也采用这样的零时刻。以窗长 $N=70$ 为标准,它的中心所对应的时间是 $t=35/256=0.137$ s,其中 $1/256$ 是对原始信号 $x(t)$ 的采样周期。因此从 $t=0.137$ s 开始绘制 $x(t)$ 的实际时频分布,并将该时刻标识为零时刻,得到 $x(t)$ 的实际时频分布曲线如图 6(d) 所示。与图 6 的其他各图对比可见,当 $N=70$ 时 STFT 所得到的时频关系与实际的时频关系是一致的;而 N 取 209 和 256 时 STFT 所得到的时频关系则与实际的时频曲线已经有相当大的误差了,而且 $N=256$ 的效果比 $N=209$ 还要更糟。

最后说明,文中 STFT 的计算采用快速傅里叶变换 (fast Fourier transform, 简称 FFT) 来实现。由于窗长最大选择范围为 256,因此计算时采用 256 点的 FFT,当所选窗长小于 256 时,则将加窗信号补零至长度为 256 后再进行 FFT。

4 对转子振动信号的冲击特征提取

在 INV1612 型多功能柔性转子实验台上进行转子振动检测实验,实验时在转子的配重盘上加两

个螺钉,这两个螺钉并未彼此平衡,如图 7 所示。转子转速为 2 000 r/min,利用电涡流传感器检测转子的振动位移,采样频率为 2000 Hz,采样点数为 1 024 点,得到的转子振动位移信号及其幅值谱如图 8 所示,其频率成分如表 2 所示。可见,主要有 3 个频率成分,其中 33.203 1 Hz 为转子的转频,而其他两个频率分别为转频的二次谐波和三次谐波。根据转子振动理论,转频伴随着其二次谐波、三次谐波的出现意味着转轴组件对中不良,从普通傅里叶变换只能得到这一结论。

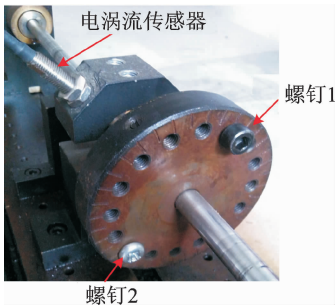


图 7 转子配重盘上添加的两个螺钉

Fig. 7 Two screws appended in weight plate of rotor

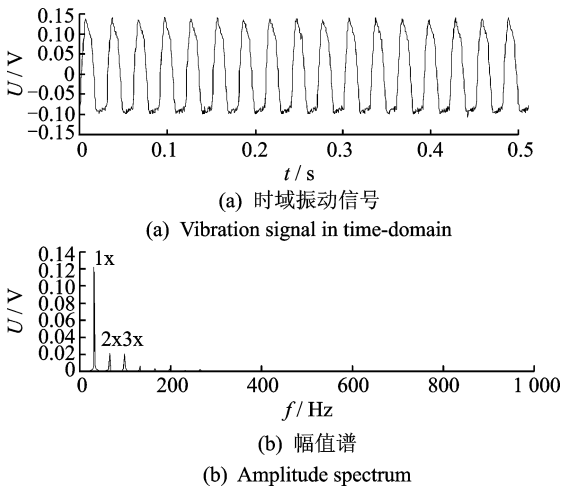


图 8 转子原始振动信号及其幅值谱

Fig. 8 Original vibration signal of rotor and its spectrum

表 2 转子的振动频率

Tab. 2 Vibration frequencies of rotor

f/Hz	U/V
33.203 1	0.121 3
66.406 2	0.021 3
99.609 4	0.020 6

现利用 STFT 来分析这一信号,窗函数选为 Hanning 窗,首先来确定合适的窗长。采用对数窗能量准则以及目前的其他 5 种窗长选择准则来选择窗长,窗长选择范围定为 8~256,取这一范围是因为太短的窗会使频率分辨率很粗糙,而太长的窗又会使 STFT 失去“短时”的意义。各种准则的窗长

选择曲线如图 9 所示。从图 9 可见,除对数窗能量准则和 Daubechies 最大模值准则外,其余 4 种准则的选择曲线都是单调下降或上升的。各种准则得到的最优窗长分别为:对数窗能量准则, $N=24$; Daubechies 最大模值准则, $N=172$; 而 Stankovic 准则、最大峭度准则、Renyi 熵准则、归一化 3 阶 Renyi 熵准则得到的最优窗长都是 $N=256$ 。

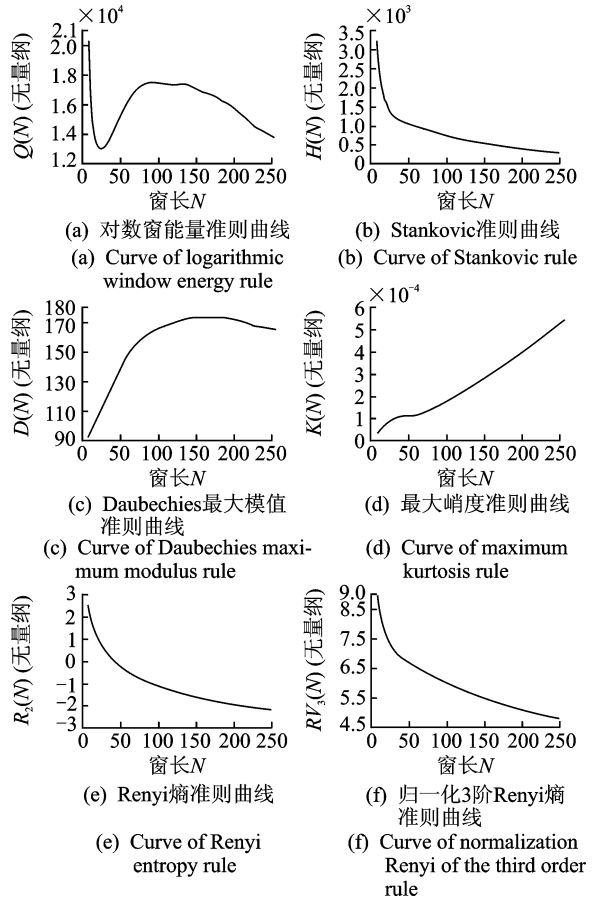


图 9 不同准则对转子振动信号的窗长选择曲线

Fig. 9 Window length selection curves of different rules for rotor vibration signal

首先利用对数窗能量准则得到的最优窗长 $N=24$ 来对原始振动信号做 STFT 分析,在这一窗长下 STFT 得到的三维时频结果如图 10(a) 所示,可见时频图中显示出在均匀的时间间隔上、在这些时刻的较大频率范围都产生了平行于频率轴的成分。在时频图中与频率轴相平行的时频特征是冲击的典型特征。理想脉冲函数的频谱在各个频率段上都存在,并且在所有频段上都是等强度的,是一种均匀谱。但实际中绝不可能产生如此理想的冲击,实际冲击的频谱总是只发生在一定频率范围内,且强度逐渐衰减,图 10(a) 的时频谱符合这一特征。画出此三维时频图在 $[0, 200]\text{Hz}$ 范围内的投影,如图 10(b) 所示,它可以更清楚地显示出各个时刻平行于

频率轴的冲击谱。从图 10(b)可见,冲击发生的间隔非常均匀,每个周期内有两个冲击,一个稍大,一个略小。这种结果毫无疑问地表明转子受到了周期性的冲击。

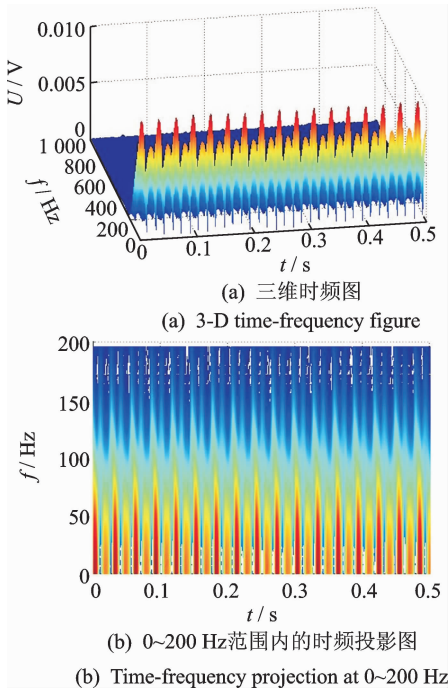


图 10 窗长 $N=24$ 时转子振动信号的 STFT 时频图
Fig. 10 STFT time-frequency figure of rotor vibration signal when window length is $N=24$

为了更清楚地显示出冲击发生的时刻以及分析冲击发生的来源,可以平行于时间轴截取时频图在冲击频率范围内的一个横截面,这里不妨截取其中频率为 31.2500 Hz 处的截面,结果如图 11 所示。它反映了 31.2500 Hz 这一频率的幅值随时间而变化的情况,实际上平行于时间轴截取冲击频率范围内的其他任何频率横截面都可得到类似的结果。

从图 11 可以看到每周期内大小相间的两个脉冲,而每两个脉冲之间的时间间隔相当均匀,通过分析脉冲峰值的位置很容易计算出每 60 个数据点产生一大一小两个脉冲,其采样频率为 2000 Hz ,所以 60 个数据点的时间间隔为 $T=60/2000=0.03\text{ s}$,

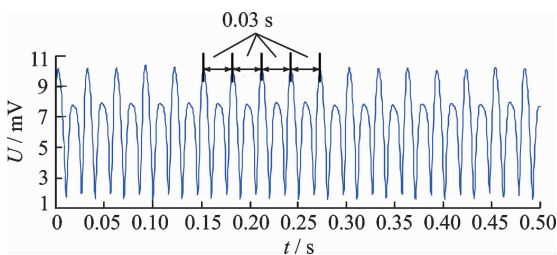


图 11 平行于时间轴截取的时频图的一个横截面
Fig. 11 One cross section intercepted parallel to time axis from time-frequency figure

据此可得冲击频率为 $f=1/T=33.33\text{ Hz}$,这正好是转子的转频,这表明转子旋转一周,分别受到两个冲击一次。根据这一数据以及转子自身的情况,可以确定这种冲击是配重盘上两个螺钉产生的。从图 7 可见,这两个螺钉和转子的轴心这三者并不处在同一条直线上,两个螺钉的质量也不一样,因而转子旋转时这两个螺钉的离心力并不能被对方平衡掉,因此转子每旋转一周,因离心力产生的冲击就分别经过传感器一次,从而被传感器检测到。从图 10(a)以及图 11 的幅值来看,这种冲击的幅度与原始信号的幅值(图 8)相比是很微小的,因而在原始信号中难以显示出来,但在合适的窗长下,STFT 却把它们提取出来了,这一大一小两个冲击的时间间隔和配重盘上两个螺钉的角度是相对应的。

再来看其他准则确定的最优窗长下的 STFT 结果。除 Daubechies 最大模值准则得到的最优窗长是 $N=172$ 外,其他 4 种准则得到的最优窗长都是 $N=256$,现分别利用这两个窗长做 STFT 分析,结果如图 12 所示,可见这两种窗长的结果差别不大,但与图 10 却有显著区别。从图 12 可见,这两种窗长下的 STFT 得到的都是平行于时间轴的 3 个频率成分,这 3 个频率成分分别是转频、转频的 2 次谐波和 3 次谐波,这种时频结果表明原始信号在任何时刻总是只存在这 3 个频率。但这样的结果在前面的普通傅里叶变换中就已经得到了,因此这两种窗长的 STFT 与普通傅里叶变换相比,并不能显示出

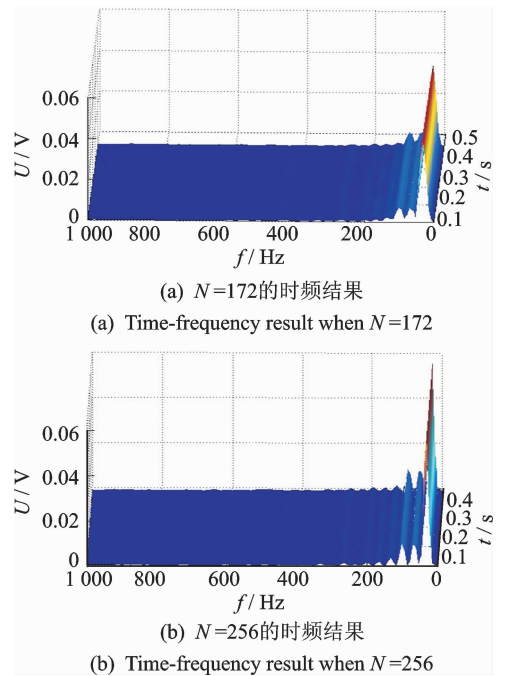


图 12 其他准则确定的最优窗长下的 STFT 时频结果
Fig. 12 STFT time-frequency results under the optimal window lengths obtained by other rules

STFT 的优势,转子受到的冲击在图 12 的两个时频图中没有任何反映。

5 结束语

窗长是决定短时傅里叶变换时频聚集性效果的关键因素,文中分析了现有 5 种窗长选择准则的选择机理,指出其中的归一化 3 阶 Renyi 熵准则和 Stankovic 准则往往只会取到窗长选择范围的最大值,并且指出这是由于此二准则中分子增长速度赶不上分母增长速度而导致选择曲线单调下降。

提出了一种基于对数窗能量的窗长选择准则。与普通的窗能量随窗长线性增长的特性不同,对数窗能量在长窗时的增长速度会变慢,而在短窗时的增长速度却会变快,有效地克服了普通窗能量的增长速度一成不变的缺陷,从而在短窗和长窗之间实现了更好的折衷选择。

对数窗能量的另一个显著优点是它的增长速度与窗型无关,而仅由窗长决定,因而对不同的窗型具有鲁棒性。信号处理结果表明,对数窗能量准则确定的窗长能够使短时傅里叶变换取得优良的时频聚集性,其效果优于现有的 5 种窗长选择准则。

参 考 文 献

- [1] Cabal-Yepez E, Garcia-Ramirez A G, Romero-Troncoso R J. Reconfigurable monitoring system for time-frequency analysis on industrial equipment through STFT and DWT [J]. IEEE Transactions on Industrial Informatics, 2013, 9(2): 760-771.
- [2] 丛晓妍,王增才,王保平,等. 基于 EMD 与峭度滤波的煤岩界面识别[J]. 振动、测试与诊断, 2015, 35(5): 950-954.
Cong Xiaoyan, Wang Zengcai, Wang Baoping, et al. Application of filtering method based on EMD and kurtosis in coal-rock interface recognition [J]. Journal of Vibration, Measurement & Diagnosis, 2015, 35(5): 950-954. (in Chinese)
- [3] 程军圣,杨怡,杨宇. 基于 LMD 的能量算子解调机械故障诊断方法[J]. 振动、测试与诊断, 2012, 32(6): 915-919.
Cheng Junsheng, Yang Yi, Yang Yu. Mechanical fault diagnosis of energy operator demodulating approach based on LMD [J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(6): 915-919. (in Chinese)
- [4] Astuti W, Sediono W, Akmeliawati R, et al. Investigation of the characteristics of geoelectric field signals prior to earthquakes using adaptive STFT techniques [J]. Natural Hazards and Earth System Sciences,

2013, 13(6): 1679-1686.

- [5] Sivasankari K, Thanushkodi K. An improved EEG signal classification using neural network with the consequence of ICA and STFT [J]. Journal of Electrical Engineering & Technology, 2014, 9(3): 1060-1071.
- [6] 虎乐乐,赵加祥,倪虹,等. 一种基于短时傅里叶变换的心音分裂时间计算方法[J]. 生物医学工程与临床, 2009, 13(1): 18-22.
Hu Lele, Zhao Jiayang, Ni Hong, et al. A method for calculating the heart sound splitting time based on STFT [J]. Biomedical Engineering & Clinical Medicine, 2009, 13(1): 18-22. (in Chinese)
- [7] Goncalves I B, Leiria A, Moura M M. STFT or CWT for the detection of Doppler ultrasound embolic signals [J]. International Journal for Numerical Methods in Biomedical Engineering, 2013, 29(9): 964-976.
- [8] McCallum M, Guillemin B. Stochastic-deterministic MMSE STFT speech enhancement with general a priori information [J]. IEEE Transactions on Audio Speech and Language Processing, 2013, 21(7): 1445-1457.
- [9] Randall R B, Antoni J. Rolling element bearing diagnostics-a tutorial [J]. Mechanical Systems and Signal Processing, 2011, 25(2): 485-520.
- [10] Yao Shuai, Fang Shiliang, Wang Xiaoyan. Parameter estimation for HFM signals using combined STFT and iteratively reweighted least squares linear fitting [J]. Signal Processing, 2014, 99(6): 92-102.
- [11] Daubechies I, Planchon F. Adaptive Gabor transforms [J]. Applied and Computational Harmonic Analysis, 2002, 13(1):1-21.
- [12] Jones D L, Parks T W. A high resolution data-adaptive time-frequency representation [J]. IEEE Transaction on Signal Process, 1990, 38 (12): 2127-2135.
- [13] Richard G, Patrick F, Augustus J, et al. Measuring time-frequency information content using the Renyi entropies [J]. IEEE Transaction on Information Theory, 2001, 47(4): 1391-1409.
- [14] Sang T H, Williams W J. Renyi information and signal dependent optimal kernel design[C]// Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Detroit, USA: IEEE Piscataway, 1995.2:997-1000.
- [15] Stankovic L. A measure of some time-frequency distributions concentration [J]. Signal Processing, 2001, 81(4): 621-631.



第一作者简介:赵学智,男,1970年11月生,博士、教授、博士生导师。主要研究方向为振动测试与信号处理。曾发表《奇异值分解对连续 Morlet 小波变换的压缩和提纯》(《机械工程学报》2015年第51卷第16期)等论文。
E-mail: mezhaoxz@scut.edu.cn