文章编号: 1000−8152(2009)10−1105−06

电力电子装置故障波形相似性度量的小波矩阵变换法

胡志坤[1], 徐 飞[1], 桂卫华[2], 阳春华[2]

(1. 中南大学 物理科学与技术学院, 湖南 长沙 410083; 2. 中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘要: 提出一种基于小波矩阵变换的时序序列相似度量方法, 并对该方法应用于电力电子装置故障波形相似性度量进行了抗噪性、灵敏度及相似值准确性分析. 方法首先采用小波变换将时序序列压缩到小波子空间, 再由K-L变换(Karhunen-Loeve transformation)提取样本时序序列的特征向量和正交基, 然后将分析时序序列通过内积变换映射到正交基中得到分析特征向量, 最后计算两个特征向量之间的欧式距离以判定时序序列的相似度. 以电力电子装置故障波形的相似度量为例, 实验表明该方法特征向量维数低, 抗噪性好于直接小波法30倍, 灵敏度是直接小波法1/3, 相似值准确性好于小波奇异值法. 该方法对于大规模时序序列的相似匹配和检索具有潜在的应用价值.

关键词: 小波变换; 奇异值分解; 内积变换; 相似时序序列; 电力电子装置

中图分类号: TP18 **文献标识码**: A

Wavelet-matrix transforming method for similarity measurement of fault waveform of electronic power devices

HU Zhi-kun[1], XU Fei[1], GUI Wei-hua[2], YANG Chun-hua[2]

(1. School of Physics Science and Technology, Central South University, Changsha Hunan 410083, China;
2. School of Information Science and Engineering, Central South University, Changsha Hunan 410083, China)

Abstract: Based on the wavelet and the matrix transformation, we propose a method for measuring the time series similarity for application in the fault waveform similarity of electronic power devices. The noise-rejection ability, the sensitivity and the accuracy of this method are discussed. By using the wavelet transformation, we compress the time-series sequence into the wavelet subspace. The sample's feature vector and the orthogonal basis of the sampled time-series sequence are obtained by K-L transformation(Karhunen-Loeve transformation). By taking the inner-product, the analyzed time-series sequence is projected into the orthogonal basis, and the analyzed feature vector is thus obtained. Finally, the similarity value is calculated by the Euclid distance between the sample's feature vector and the analyzed feature vector. In the measurement of the similarity between the fault waveforms in electronic power devices, the experimental results show that the dimension of feature vectors is low by the proposed method. In addition, the noise-rejection ability of the proposed method is 30 times higher than that of the plain wavelet method, the sensitivity of the proposed method is 1/3 of that of the plain wavelet method, and the accuracy of similarity value of the proposed method is higher than that of the wavelet singular-value-decomposition method. The proposed method has potential value in similarity matching and indexing for lager time-series databases.

Key words: wavelet transform; singular-value-decomposition; inner-product transform; time series similarity

1 引言(Introduction)

由于故障录波器的广泛使用, 大量电力电子装置的故障波形得以保存, 采用基于信号处理的方法将当前故障波形和历史故障波形进行相似性匹配实现故障诊断, 其中故障波形的相似性度量是关键. 时序序列的相似性一般直接采用欧氏距离度量, 但直接采用时序序列作为特征向量, 则维数过高, 不适合大数据量的检索. 文献[1,2]提出了基于傅立叶变换的相似匹配算法, 将最大的几个谐波系数作为特征向量计算时间序列的相似性, 但它是一种频率特征分析方法, 仅反映全周期波形的频率特性. 文献[3,4]采用了正交小波变换进行时序序列相似匹配, 但小波系数与时序序列的长度相等, 其特征维数也过高不利于大数据量的检索. 文献[5]采用了小波奇异值法, 它将时序序列的小波系数矩阵进行奇异值分解得到奇异值向量. 其实质是用双重降维将原始时序序列

收稿日期: 2008−05−16; 收修改稿日期: 2008−12−09.
基金项目: 国家自然科学基金资助项目(60634020, 60904077); 深圳市科技基础研究项目资助项目(JC200903180555A).

映射为特征向量, 有利于大数据量检索. 但奇异值向量只代表某正交空间的能量分布, 不同时序序列分解后的正交空间不同. 因此, 奇异值向量代表不同空间的特征向量, 不同小波系数矩阵的奇异值向量之间不具有相似对比性.

本文以电力电子装置故障输出波形的相似匹配为对象, 利用小波变换具有很好的局部时频表征能力、矩阵奇异值分解分量具有正交性和内积空间变换具有很好函数逼近性的特点, 提出小波矩阵变换法.

2 小波矩阵变换法(Wavelet matrix transform)

2.1 离散正交小波变换(Discrete orthogonal wavelet transform)

离散正交小波变换是一种非冗余的小波变换. 它在离散框架的基础上构造正交基, 将离散数据映射到相互正交的子空间中, 分解公式如下:

$$\begin{cases} c_k^{j+1} = D(h_0 * c_k^j), \\ d_k^{j+1} = D(h_1 * c_k^j). \end{cases} \tag{1}$$

其中: c_k^j为逼近信号, d_k^j为细节信号, j为分解层数, k为平移因子, h_0为低通滤波器, h_1为带通滤波器, D为抽样算子, 过程如图1所示[6].

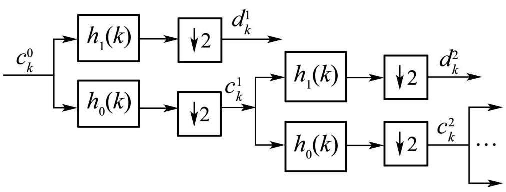


图 1 离散正交小波变换

Fig. 1 Discrete orthogonal wavelet transform

设长度为N的时序序列

$$T = (t_1, t_2, \cdots, t_N),$$

经L级离散正交小波变换后, 得到小波系数

$$C_w = (c_L, d_L, d_{L-1}, \cdots, d_1).$$

其中c_L是L层的近似系数, d_k是k层$(k = 1, 2, \cdots, L)$的细节系数. 将各级系数按尾数补零方式拓延成长度为$N/2$的序列, 以行优先顺序构造矩阵$M_T \in \mathbb{R}^{(L+1)\times(N/2)}$.

$$M_T = \begin{pmatrix} c_L' \\ d_L' \\ d_{L-1}' \\ \vdots \\ d_1' \end{pmatrix}. \tag{2}$$

该矩阵代表序列T所有的时频信息, 因此时序序列的相似度量可在小波子空间中进行.

2.2 矩阵的分解和变换提取正交基(Extracting orthogonal basics by decomposition and transform of matrix)

小波系数矩阵的维数过高不利于大型数据的相似检索, 因此需对小波系数矩阵进行特征值提取. 本文利用矩阵的分解和变换提取正交基, 为时序序列特征向量的提取提供统一且有效的空间.

引理 1 对于任何矩阵$A \in \mathbb{R}^{m\times n}$, 其中$(m > n)$都存在公式(3)的分解[7,8]:

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^{\mathrm{T}}. \tag{3}$$

其中: U为m阶酉矩阵, V为n阶酉矩阵,

$$\delta = \mathrm{diag}\{\delta_1, \delta_2, \cdots, \delta_r\}$$

由引理1的矩阵奇异值分解可得推论1.

推论 1 矩阵$A \in \mathbb{R}^{m\times n}$经过奇异值分解, 表示为$r$个秩为1的$\delta_i \bar{A}_i$子矩阵之和:

$$A = \sum_{i=1}^{k} \delta_i u_i v_i^{\mathrm{T}} = \sum_{i=1}^{k} \delta_i \bar{A}_i. \tag{4}$$

其中: r为A的秩, u_i和v_i为矩阵U和V的第i列向量, $\bar{A}_i$为向量u_i和v_i的乘积.

由引理1中V为n阶酉矩阵, 即v_i之间是两两正交, 可得出推论2.

推论 2 矩阵$A \in \mathbb{R}^{m\times n}$经过奇异值分解后, 各个子矩阵$\bar{A}_i \in \mathbb{R}^{m\times n}$之间两两相互正交:

$$\bar{A}_i \cdot \bar{A}_j = u_i v_i^{\mathrm{T}} v_j u_i^{\mathrm{T}} = \begin{cases} \Omega, \ i = j, \\ 0, \ \ i \neq j. \end{cases} \tag{5}$$

其中Ω为实数.

将矩阵A和$\bar{A}_i$, 其中$i = 1, 2, \cdots, r$, 分别序列化后得向量v_A和$v_{\bar{A}i}$. 对于向量v_A和$v_{\bar{A}i}$, 由推论1、推论2可推导出定理1.

定理 1 矩阵$A \in \mathbb{R}^{m\times n}$经奇异值分解和序列化后, 可被表示为长度相等的子序列的叠加, 且子序列之间相互正交:

$$\begin{cases} v_A = \delta_1 \cdot v_{\bar{A}1} + \cdots + \delta_r \cdot v_{\bar{A}r}, \\ v_{\bar{A}i} \perp v_{\bar{A}j}, \ i \neq j. \end{cases} \tag{6}$$

其中: 向量v_A为矩阵A的序列化, $v_{\bar{A}_i}$为$\bar{A}_i \in \mathbb{R}^{m\times n}$的序列化, $i = 1, 2, \cdots, r$, r为矩阵A的秩.

证 由推论1可知, 矩阵A可表示为式(7):

$$A = \sum_{i=1}^{k} \delta_i \bar{A}_i. \tag{7}$$

分别对A和$\bar{A}_i$进行矩阵的序列化可得下式:

$$v_A = (x_1, x_2, \cdots, x_m). \tag{8}$$

其中x_i为矩阵A的第i行向量.

$$v_{\bar{A}i} = (x_{i,1}, x_{i,2}, \cdots, x_{i,m}). \tag{9}$$

其中: $x_{i,j}$为矩阵$\bar{A}_i$的第j行向量, 且$i = 1, 2, \cdots, r$, $j = 1, 2, \cdots, m$.

又因为$x_j = \sum\limits_{i=1}^{r} x_{i,j}$ 由式(7)～(9)可得出结论:

$$v_A = \delta_1 \cdot v_{\bar{A}1} + \cdots + \delta_r \cdot v_{\bar{A}r}. \tag{10}$$

由推论2可知, 经奇异值分解后的子矩阵相互正交, 将矩阵序列化可推出式(11):

$$\begin{pmatrix} x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,m} \end{pmatrix} (x_{j,1}^{\mathrm{T}} \ \ x_{j,2}^{\mathrm{T}} \ \ \cdots \ \ x_{j,m}^{\mathrm{T}}) = 0. \tag{11}$$

由式(11)可推出式(12):

$$\begin{cases} x_{i,k} \cdot x_{j,l}^{\mathrm{T}} = 0, \ i \neq j, \\ k, l = 1, 2, \cdots, m. \end{cases} \tag{12}$$

对于式(10)中的$v_{\bar{A}i}$序列, 当$i \neq j$时可得

$$v_{\bar{A}i} \cdot v_{\bar{A}j} = \sum_{p=1}^{m} x_{i,p} \cdot x_{j,p}. \tag{13}$$

再由公式(11)可得出结论:

$$v_{\bar{A}i} \cdot v_{\bar{A}j} = \begin{cases} \varOmega, \ i = j, \\ 0, \ \ i \neq j, \end{cases} \tag{14}$$

即: $v_{\bar{A}i} \perp v_{\bar{A}j}$, $i \neq j$.

证毕.

该定理的物理意义是为时序序列相似特征向量的提取提供统一且有效的正交基. 该正交基对于样本时序序列是完备的, 可直接应用于内积变换, 提取样本和分析时序序列的特征向量.

2.3 时序序列的相似度量(Time series similarity measurement)

对于样本时序序列T_{sa}和待相似分析时序序列T_{an}, 其特征向量提取方法并不相同.

对于T_{sa}, 其特征向量的提取是将T_{sa}小波变换得小波系数矩阵M_{sa}, 再将M_{sa}奇异值分解, 得到奇异值向量

$$\delta_{\mathrm{sa}} = (\delta_1^{\mathrm{sa}}, \delta_2^{\mathrm{sa}}, \cdots, \delta_r^{\mathrm{sa}})$$

和正交子矩阵M_i^{sa}. δ_{sa}是T_{sa}的特征向量, M_i^{sa}的序列化v_i^{sa}为特征向量提取的正交基, 其中$i = 1, 2, \cdots, r$.

由定理1可知, δ_i^{sa}为v_i^{sa}系数, 正交基$\{v_i^{\mathrm{sa}}\}$为v_A的完备正交基, δ_i^{sa}可被表示为

$$\delta_i^{\mathrm{sa}} = < v_{\mathrm{sa}}, v_i^{\mathrm{sa}} > . \tag{15}$$

其中: $< \cdot >$为内积变换, v_{sa}为M_{sa}的序列化.

由式(15)可知, δ_{sa}可完全表征v_{sa}的信息, 因此δ_{sa}可作为T_{sa}的特征向量.

对于T_{an}, 其特征向量的提取是将T_{an}小波变换得小波系数矩阵M_{an}, 再对M_{an}序列化得向量v_{an}, 最后将v_{an}投影到正交基$\{v_i^{\mathrm{sa}}\}$中得T_{an}的特征向量

$$\delta_{\mathrm{an}} = (\delta_1^{\mathrm{an}}, \delta_2^{\mathrm{an}}, \cdots, \delta_r^{\mathrm{an}}),$$

表达式如下:

$$\delta_i^{an} = < v_{\mathrm{an}}, v_i^{\mathrm{sa}} > . \tag{16}$$

由于v_{sa}和v_{an}都是在空间$\{v_i^{\mathrm{sa}}\}$中提取, 因此两特征向量具有相似对比性. 虽然$\{v_i^{\mathrm{sa}}\}$并非T_{an}的完备正交基, 但是它的特征向量对于相似对比仍有效. 这是因为当T_{an}与T_{sa}比较相似时, 它们完备正交基很相似, 在$\{v_i^{\mathrm{sa}}\}$空间中提取的特征向量能用于相似比较; 当T_{an}与T_{sa}不相似时, 它们之间的完备正交基必然不同, 由内积变换提取的特征向量差异也会很大, 因此对于相似比较也非常适合.

分析时间序列与样本时间序列的相似性度量的详细步骤如下:

1) 将样本时序序列$T_{\mathrm{sa}} = t_1, t_2, \cdots, t_n$和分析时序序列$T_{\mathrm{an}} = t_1', t_2', \cdots, t_n'$ 分别进行离散正交小波变换得到小波系数矩阵M_{sa}和M_{an}.

2) 对M_{sa}进行奇异值分解和序列化, 得奇异值向量$\delta_{\mathrm{sa}} = (\delta_1^{\mathrm{sa}}, \delta_2^{\mathrm{sa}}, \cdots, \delta_r^{\mathrm{sa}})$ 和正交集合$\{v_i^{\mathrm{sa}}\}$ $(i = 1, 2, \cdots, r)$.

3) 将M_{an}序列化后的v_{an}与$\{v_i^{\mathrm{sa}}\}$进行内积运算$\delta_i^{\mathrm{an}} = < v_{\mathrm{an}}, v_i^{\mathrm{sa}} >$, 得$T_{\mathrm{an}}$的特征向量

$$\delta_{\mathrm{an}} = (\delta_1^{\mathrm{an}}, \delta_2^{\mathrm{an}}, \cdots, \delta_r^{\mathrm{an}}).$$

4) 由

$$d_{\mathrm{an}} = \sqrt{\sum_{i=1}^{r} (\delta_i^{\mathrm{sa}} - \delta_i^{\mathrm{an}})^2}$$

计算T_{sa}和T_{an}的相似值.

5) 设定合理的阈值Y, 当$d_{\mathrm{an}} \leqslant Y$时, 则$T_{\mathrm{sa}}$与$T_{\mathrm{an}}$相似, 反之则不相似.

时序序列相似度量算法的流程图如图2所示.

3 仿真分析(Simulation and analysis)

以电力电子装置最典型的主电路—逆变器模型为例, 如图3所示. 研究逆变器中IGBT开路故障的诊断. 假设逆变器中T_1和T_6号IGBT同时出现开路故障, 图3中B点与N点之间的电压U_{BN}为故障波形, 如图4所示.

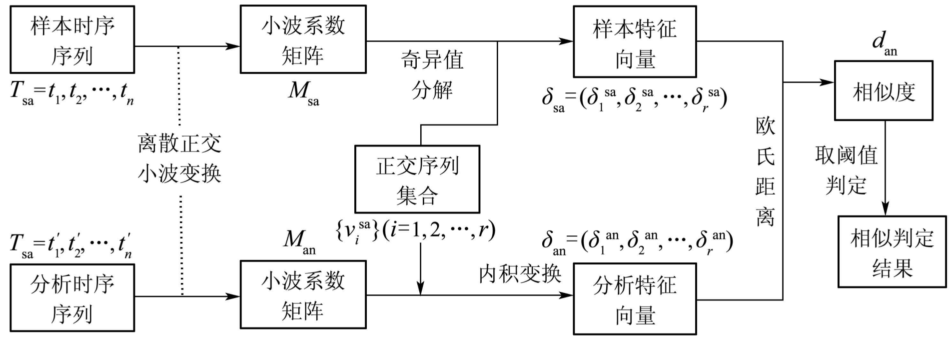


图 2 时序序列相似度量算法的流程图

Fig. 2 Flow chart of time series similarity measurement

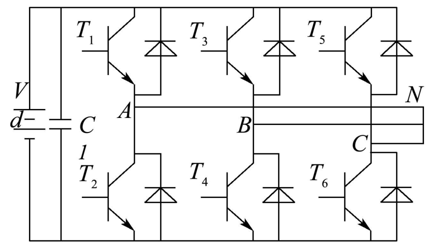


图 3 逆变器模型

Fig. 3 Invert model

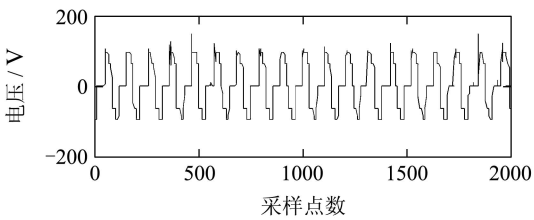


图 4 故障波形图

Fig. 4 Fault wave of Invert

3.1 抗噪能力仿真分析(The simulation and analysis of anti-noise ability)

将故障电压U_{BN}加白噪声, 如式(17)所示:

$$U_{\mathrm{no}}(i) = U_{\mathrm{BN}} + i * N_{\mathrm{o}}. \tag{17}$$

其中: $1 \leqslant i \leqslant 10$, N_{o}为白噪声.

分别用3种方法计算U_{BN}和$U_{\mathrm{no}}(i)$的相似值. 因为DB小波具有N阶消失矩作用, 且在数据压缩中小波函数的消失矩越高则压缩倍数就越大, 但消失矩越大计算量也相应增加. 故本文对3种方法的小波变换均采用“db3”小波, 分解层数为4层. 在噪声影响下3种方法计算的相似值对比图如图5所示.

由由图5可知, 当噪声系数增加时, 直接小波法的相似值以近似45度角的直线增长, 小波奇异值法以15度角增长. 相对于以上两种方法, 小波矩阵变换法的相似值增长特别缓慢, 其增长角度约为1.5度, 因此其抗噪性好于直接小波法约30倍, 好于小波奇异值法约10倍.

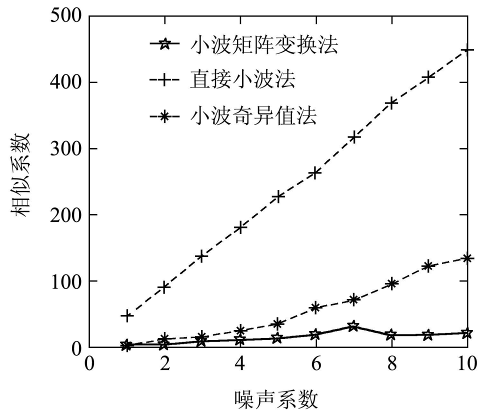


图 5 3种方法抗噪比较

Fig. 5 The comparison of anti-noise ability

噪声信号通常为白噪声, 其均匀分布在各频带中, 故带有噪声的信号经正交小波变换后各级系数都会受噪声影响. 直接小波法叠加了各级小波系数的噪声影响, 所以相似值改变很大. 小波矩阵变换法采用小波系数与正交基的内积运算, 正交基为样本时序序列经奇异值分解后的正交分量, 该正交分量与噪声信号相关性很小, 因此噪声信号经内积运算后对特征向量影响较小, 具有较强的抗噪性能.

3.2 灵敏度仿真分析(The simulation and analysis of sensitivity)

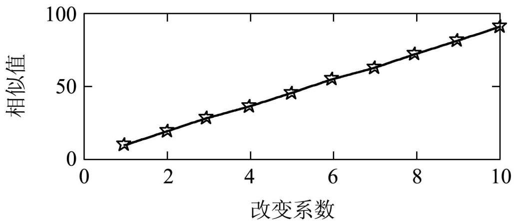
由U_{BN}的傅立叶变换可知, 故障波形能量最高分量是19 Hz, 故少量减少故障波形19 Hz正弦分量, 测试3种方法的灵敏度, 如式(18)所示:

$$U_{se}(i) = U_{\mathrm{BN}} - i * S_{19}. \tag{18}$$

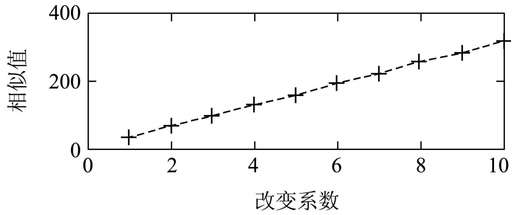
其中: S_{19}为19 Hz正弦分量, $1 \leqslant i \leqslant 8$.

分别用3种方法计算U_{BN}和$U_{se}(i)$的相似值, 灵敏度对比图如图6所示.

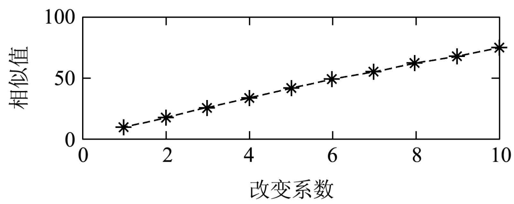
由图6可知, 3种方法都以较大的斜率增长. 小波矩阵变换法与小波奇异值法的斜率大致相等约为9, 直接小波法的斜率约为30. 小波矩阵变换法的抗噪性是直接小波法的30倍, 灵敏度是直接小波法的1/3, 因此小波矩阵法的灵敏度损失较小.



(a) 小波矩阵变换法



(b) 直接小波法



(c) 小波奇异值法

图 6 3种方法灵敏度比较

Fig. 6 The comparison of sensitivity

由于小波矩阵变换法的正交基与噪声不相关, 而与减少的分量相关, 因此小波矩阵变换法在良好的抗噪性同时具有较高的灵敏度.

3.3 相似值的准确性仿真分析(The simulation and analysis of accuracy)

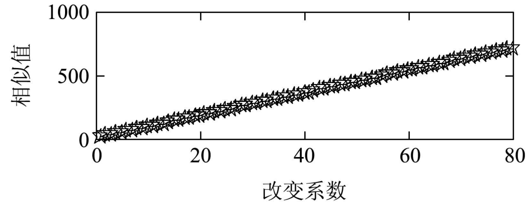
为了验证3种方法相似值的准确性, 将U_{BN}信号进行了较大的改变, 如式(19)所示:

$$U_{\mathrm{ac}}(i) = U_{\mathrm{BN}} - i * S_{19}. \tag{19}$$

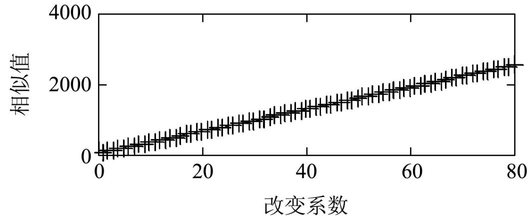
其中: $1 \leqslant i \leqslant 80$, S_{19}为19 Hz正弦分量.

分别用3种方法计算U_{BN}和$U_{\mathrm{ac}}(i)$的相似度, 相似值对比图如图7所示.

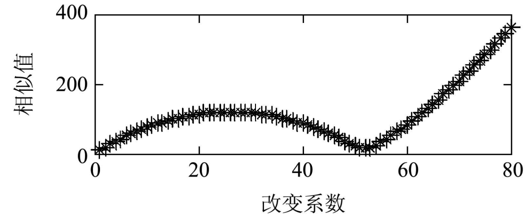
由图7可知, 当信号变化逐渐增大时, 小波矩阵变换法和直接小波法的相似值呈线性增长, 准确地反映了信号的变化. 然而小波奇异值法的相似值曲线呈现波动形态, 其相似值变化与信号的变化不相符, 因此该方法的相似值不够准确.



(a) 小波矩阵变换法



(b) 直接小波法



(c) 小波奇异值法

图 7 3种方法相似值准确性比较

Fig. 7 The comparison of accuracy

小波奇异值法是对小波系数矩阵进行奇异值分解. 虽然奇异值向量具有良好的特征表达能力, 但奇异值向量只表征矩阵在正交子矩阵中的投影值, 不同矩阵经过奇异值分解后的正交空间不一定相同, 故奇异值向量只能表达矩阵在相应的子空间中的能量分布. 图7中小波奇异值法的曲线呈现波动形态原因是19 Hz分量的波幅由正到零再到负, 相似值先变大然后变小. 当19 Hz正弦信号能量相同符号相反时, 出现相似值为零的误判. 因此, 小波奇异值法对于较相似的时序序列判定有效, 但是对于任意序列其相似值的准确性不高.

4 结论(Conclusion)

本文以研究历史故障波形与现有故障波形的相似度来实现电力电子装置的故障诊断为背景, 提出了时序相似度量的小波矩阵变换法. 通过该方法与直接小波法和小波奇异值法对故障波形相似判定的对比结果, 得出如下结论:

1) 小波矩阵变换法相对于直接小波法具有较低的特征向量维数, 方便大数据库的检索.

2) 小波矩阵变换法的抗噪性是直接小波法的30倍, 是小波奇异值法的10倍.

3) 小波矩阵变换法既具有较高的抗噪性, 也具有较高的灵敏度.

4) 小波矩阵变换法的相似值准确度明显高于小波奇异值法.

小波矩阵变换法是一种有效的研究时序序列相似度量的方法, 可以采用基于案例推理方法对病态运行状态或系统故障的时序序列波形进行匹配和检索, 为实现系统的状态监控和故障诊断奠定基础.

参考文献(References):

[1] AGRAWAL R, FALOUTSOS C, SWAMI A. Efficient similarity search in sequence databases[C] //*Proceedings of the 4th International Conference on Foundations of Data Organization and Algorithms*. New York, USA: Springer, 1993: 69 – 84.

[2] FALOUTSOS C, RANGANATHAN M, MANOLOPOULOS Y. Fast subsequence matching in time series databases[C] //*Proceeding of the 1994 ACM SIGMOD International Conference on Management of Data*. Minneapolis, USA: ACM press, 1994: 419 – 429.

[3] 张海勤, 蔡庆生. 基于小波变换的时间序列相似模式匹配[J]. 计算机学报, 2003, 26(3): 373 – 377.
(ZHANG Haiqin, CAI Qingsheng. Time series similar pattern matching based on wavelet transform[J]. *Chinese Journal of Computers*, 2003, 26(3): 373 – 377.)

[4] FRANKY K, ADA W, CLEMENT Y. Haar wavelets for efficient similarity search of time-series: with and without time warping[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2003, 15(3): 686 – 705.

[5] 张德政. 基于相似模式的知识发现方法的学习和应用[D]. 北京: 北京科技大学, 2002.
(ZHANG Dezheng. *Study and application on knowledge discovery method based on similar pattern*[D]. Beijing: University of Science and Technology Beijing, 2002.)

[6] MALLAT S. A theory for multi resolution signal decomposition: the wavelet representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(7): 674 – 693.

[7] AKRITAS A G, MALASCHONOK G I. Application of singular value decomposition[J]. *Mathematics and Computers in Simulation*, 2004, 67(1): 15 – 31.

[8] VRABIE V D, MARS J I, LACOUME J L. Modified singular value decomposition by means of independent component analysis[J]. *Signal Processing*, 2004, 84(3): 645 – 652.

作者简介:

胡志坤 (1976—), 男, 副教授, 博士, 从事复杂系统的建模、优化与故障诊断研究, E-mail: huzk@mail.csu.edu.cn;

徐 飞 (1983—), 男, 硕士研究生, 从事电力电子变流技术研究, E-mail: ybbdqq@yahoo.com.cn;

桂卫华 (1950—), 男, 教授, 博士生导师, 从事复杂系统的建模和故障诊断研究, E-mail: gwh@mail.csu.edu.cn;

阳春华 (1965—), 女, 教授, 博士生导师, 从事复杂系统建模与优化控制研究, E-mail: ychh@mail.csu.edu.cn.