

# 基于朴素贝叶斯方法和权值分析方法的电机轴承故障诊断

李万清

(河南煤化集团 洛阳龙门煤业公司, 河南 洛阳 471000)

**摘要:** 为了分析电动机轴承故障类型,首先采用小波包分析对轴承振动信号进行了高低频分解及重构,然后以各频带的能量值构成了轴承振动信号的特征向量,最后通过朴素贝叶斯网和提出的权重分析两种方法进行了电机轴承故障分类;采用朴素贝叶斯网对已知电机轴承故障类型的样本数据进行训练,获得参数后识别未知样本的故障类型,利用权重分析法计算未知与已知类型的电机轴承振动样本的相关系数,然后构建权重,并按照权值的大小获取未知样本的故障类型。仿真结果表明,朴素贝叶斯网能较好地实现电机故障诊断,所提出的权重分析方法也能较好地对电机故障进行诊断。

**关键词:** 电机故障;小波包;朴素贝叶斯;权值分析

中图分类号: TM343;TH133.3

文献标志码:A

文章编号:1001-4551(2012)04-0390-04

## Fault diagnosis of motor bearing based on naive Bayes and weight analysis methods

LI Wan-qing

(Longmen Coal Company, Henan Coal Chemical Industry Group, Luoyang 471000, China)

**Abstract:** In order to analyze the types of motor bearing faults, wavelet package analysis was firstly used to decompose and reconstruct the signals of the ball bearings into different frequency bands, then the values of energies on each bands were used to compose feature vectors of ball bearings' signals. Those vectors, which were considered as samples, were used in naive Bayes and weight analysis models respectively to complete the classification of motor bearing faults. Naive Bayes network trained the training samples (whose type are known), and then classified the testing samples (whose type are unknown). In weight analysis model, the Euclidean distances of each testing samples and training samples were computed, and the types of all testing samples were obtained through the constructed weights. The simulation results show that, through wavelet package analysis, Naive Bayes can deal with motor bearing fault well, and the weight analysis method can also analyze the fault signals effectively.

**Key words:** fault of motor; wavelet package; Naive Bayes; weight analysis

## 0 引言

在煤矿生产中,电机被广泛应用。在煤矿事故中,除了瓦斯事故外<sup>[1]</sup>,电机的正常工作对煤矿安全生产过程中的高效、优质及低耗运行意义也非常重大。电机出现故障时不仅会损坏电机本身,还会影响整个系统的正常运行,甚至会危及人身安全,造成巨大的经济损失。因此,尽量减少电机故障的发生与有效的识别

电机故障类型在煤矿生产中具有积极意义。

煤矿生产中所用的电机大多数为三相异步电动机。采区任务的运输通常由三相异步电动机驱动刮板输送机,通过旋转运动实现。电动机的故障形式主要有以下5种:轴承损坏;绕组接线、引线和端子烧坏;绕组匝间或相间局部击穿;绕组烧毁;转子断轴。其中,轴承故障发生的比例最大,约为40%<sup>[2]</sup>,轴承故障将引起电机异常振动,使电机无法正常运行,严重时会对煤

矿生产和人身安全带来极大危害。因此有必要研究电机轴承故障诊断。

在电动机的故障诊断中,经常用到的智能方法有神经网络(ANN)、支持向量机<sup>[3]</sup>、聚类等方法。中国矿业大学的张建文教授等人<sup>[4]</sup>采用神经网络对电机的转子断条故障进行诊断。首先用一种扩展到Park矢量方法提取故障特征,然后利用反向传播(BP)神经网络进行异步电机转子断条故障诊断,BP网络是神经网络的一种,其缺点在于收敛速度慢、易于陷入局部极小点。付光杰等人<sup>[5]</sup>也采用BP神经网络对电机转子故障进行诊断,同时为了克服BP网络的缺点,他们研究者采用粒子群优化方法训练BP网络,加快了模型的收敛速度,并克服了BP网络易于限于局部极小点的缺点等。

本研究首先采用振动检测的方法对电机轴承振动进行检测,以获得轴承振动信号;通过小波包变换对振动信号进行高低频分解及重构,以获取故障特征向量。为了避免诊断模型陷入局部极小点的情况,本研究采用朴素贝叶斯网以及提出的权值分析法进行模式识别,诊断出故障。最后,通过仿真验证本研究提出方法的有效性。

### 1 故障特征向量提取

电动机轴承故障的信号是非平稳的,为了较好地反映其频域特性,本研究采用小波包分析。小波包分析是对小波变换的一种改进,它可以将小波变换中没有细分的高频部分做进一步分解,提高信号通频带的频率分辨能力<sup>[6]</sup>。本研究采用小波包分解方法从轴承振动信号中提取特征向量。

小波包分解提取特征向量的步骤如下所示:

(1) 由下式对信号进行小波包分解:

$$\begin{cases} d_m^{j-1,2n} = \sum d_l^{j,n} h_{j-2m} \\ d_m^{j-1,2n+1} = \sum d_l^{j,n} g_{j-2m} \end{cases} \quad (1)$$

小波包分解每次得到的结果都是两个序列,对轴承振动信号进行3层小波包分解,则第3层一共可获得8个频率段的信号特征。本研究提取8个频率成分的信号特征向量  $X_{ij}$ , 其中,  $i=0,1,2,3$ ,  $j=0,1,2,3,4,5,6,7$ 。 $i$  和  $j$  分别表示小波分解的层数和结点数。

(2) 由下式对小波包分解系数重构,提取8个频带范围的信号:

$$d_k^{j,n} = \sum (d_m^{j-1,2n} h_{k-2m} + d_m^{j-1,2n+1} g_{k-2m}) \quad (2)$$

本研究以  $S_{ij}$  为  $X_{ij}$  的重构信号,分别对变换后的第3层小波包所有结点进行分析,则总信号  $S$  可以表

示为  $S = S_{30} + S_{31} + S_{32} + S_{33} + S_{34} + S_{35} + S_{36} + S_{37}$ 。

(3) 求各频带信号的总能量。本研究设  $S_{3j}$  ( $j=0,1,\dots,7$ ) 对应的能量为  $E_{3j}$  ( $j=0,1,\dots,7$ ), 则有:

$$E_{3j} = \int |S_{3j}(t)|^2 dt = \sum_{k=1}^n |x_{jk}|^2 \quad (3)$$

式中:  $x_{jk}$  ( $j=0,1,\dots,7; k=1,2,\dots,n$ )—重构信号  $S_{3j}$  离散点的幅值。

(4) 构造特征向量。当电机轴承出故障时<sup>[7]</sup>,对各频带内信号的能量有较大的影响,因此,能够以能量为标度构造特征向量。当能量较大时,  $E_{3j}$  是一个较大的数值,为了分析的方便,笔者对所构造的特征向量进行归一化处理:

$$E = \left( \sum_{j=0}^7 E_{3j} \right)^{1/2} \quad (4)$$

$$T = [T_{30}, T_{31}, \dots, T_{37}] = [E_{30}/E, E_{31}/E, \dots, E_{37}/E]$$

式中:  $T$ —归一化后的特征向量。

## 2 朴素贝叶斯以及权值分析方法的电机故障诊断

电机故障诊断模型如图1所示。

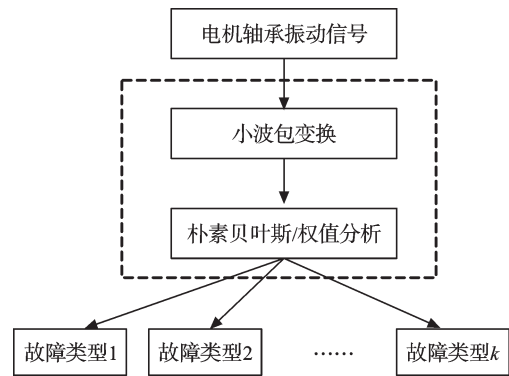


图1 电机故障诊断模型

### 2.1 基于朴素贝叶斯分类的故障诊断

贝叶斯学习方法中实用性很高的是朴素贝叶斯<sup>[8]</sup>。朴素贝叶斯分类器同ANN以及SVM一样都是监督学习方法,贝叶斯分类器从提供的训练样本中训练模型然后对未知样本进行分类。

假定数据集中有  $k$  个类  $c_1, c_2, \dots, c_k$ ,  $k$  表示电机故障的类别。对任意的一个未知的故障数据样本  $x$ , 分类器将具有最高后验概率(条件  $x$  下)的类标号赋予  $x$ , 即朴素贝叶斯分类将未知的样本分配给类  $c_i$ , 其中  $P(c_i/x) > P(c_j/x)$ ,  $1 \leq j \leq k$  且  $j \neq i$ 。

$P(c_i/x)$  最大的类  $c_i$  成为最大后验假定, 计算如下:

$$P(c_i/x) = \frac{P(c_i)P(x/c_i)}{P(x)} \quad (5)$$

由于  $p(x)$  对于所有类为常数,为获得最大  $P(c_i/x)$ , 只要计算最大  $P(c_i)P(x/c_i)$  即可。其中  $P(c_i) = C_i/C$ ,  $C_i$  是类  $c_i$  中的训练样本数,  $C$  是训练样本总数。根据朴素贝叶斯的类条件独立假设(朴素假定),属性之间相互独立,则有:

$$P(x/c_i) = P(x_1, x_2, \dots, x_n/c_i) = \prod_{k=1}^n P(x_k/c_i) \quad (6)$$

由式(2)可以看出,朴素贝叶斯的朴素假定使得在对新实例预测类标号时,大大地简化了联合概率  $P(x_1, x_2, \dots, x_n/c_i)$  的计算,节省了系统的开销。概率  $P(x_k/c_i)$  可由训练样本计算,即:

$$P(x_k/c_i) = \frac{C_{mk}}{C_i} \quad (7)$$

式中:  $C_{ik}$  一类  $c_i$  中在属性  $x_k$  上具有值  $m$  的训练样本数。

### 2.2 基于权值分析法的电机故障诊断

朴素贝叶斯诊断模型建立的前提是后验假定,而在实际问题中,并不是所有的数据集都能满足后验假定,因此,本研究提出权值分析法的电机故障诊断模型。该模型以欧式距离作为测量标度,通过权重的构造对未知故障类型的数据进行诊断。数据集记为  $X = (x_j^{i'})_J^I$ , 其中  $J$  和  $I$  分别表示故障样本中的属性个数与样本个数,  $x_j^{i'}$  为第  $i'$  个样本在第  $j'$  个属性下的表达值。对任意未知故障类型的样本  $o$ , 由下式计算  $o$  与已知故障类型样本  $q$  的欧式距离  $d$  :

$$d_{oq} = \sqrt{(x_o - x_p)(x_o - x_p)^T} \quad (8)$$

然后由下式构造权重:

$$w_{oq} = \frac{1/d_{oq}}{\sum(1/d)} \quad (9)$$

式中:  $\sum(1/d)$  一样本  $o$  与所有已知故障类型样本间的欧式距离的倒数之和。

对  $k$  类电机故障样本,每一类都选出一些样本作为已知样本,将构造好的权重进行分类,对样本  $o$  与同一类故障类型中的已知样本间欧式距离权重求和:

$$w_o^{k'} = \sum w_{oq_{k'}} \quad (10)$$

式中:  $w_o^{k'}$  一样本  $o$  属于第  $k'$  类故障的总权重,  $w_{oq_{k'}}$  一样本  $o$  与属于第  $k'$  类的所有已知样本  $q_{k'}$  的权重。

最后把样本  $o$  归为最大的  $w_o^{k'}$  对应的第  $k'$  类。

## 3 仿真研究

### 3.1 仿真数据集

本研究以电机滚动轴承获得的振动信号为研究对象,分别对电机轴承故障中的正常情况、内圈故障、外圈故障和滚动体故障进行仿真。首先,本研究对不同工况下的电机滚动轴承振动信号进行3层小波包分解,然后进行信号重构,提取经归一化处理的8个频带成分的特征信号构成数据集,数据集与文献[9]中数据集相同。

12组样本数据及各状态对应的频带如表1、表2所示。

表1 样本数据

状态	样本标号	$T_{30}$	$T_{31}$	$T_{32}$	$T_{33}$	$T_{34}$	$T_{35}$	$T_{36}$	$T_{37}$
正常	1	0.126	0.128	0.036	0.057	0.149	0.089	0.147	0.034
正常	2	0.089	0.176	0.104	0.082	0.113	0.102	0.066	0.078
正常	3	0.146	0.086	0.132	0.104	0.076	0.032	0.025	0.103
内圈故障	4	0.495	0.956	0.578	0.496	0.102	0.147	0.189	0.176
内圈故障	5	0.962	0.473	0.524	0.412	0.256	0.224	0.065	0.234
内圈故障	6	0.526	0.544	0.676	0.378	0.104	0.112	0.201	0.077
外圈故障	7	0.204	0.304	0.523	0.614	0.626	0.628	0.301	0.12
外圈故障	8	0.109	0.104	0.476	0.576	0.948	0.234	0.104	0.289
外圈故障	9	0.292	0.107	0.553	0.702	0.992	0.587	0.305	0.137
滚动体故障	10	0.206	0.307	0.108	0.109	0.676	0.973	0.894	0.579
滚动体故障	11	0.182	0.201	0.304	0.065	0.772	0.0865	0.973	0.672
滚动体故障	12	0.086	0.083	0.205	0.243	0.676	0.581	0.764	0.724

表2 各状态对应的频带

状态	$T_{30}$	$T_{31}$	$T_{32}$	$T_{33}$	$T_{34}$	$T_{35}$	$T_{36}$	$T_{37}$
频带/kHz	0~0.625	0.625~1.25	1.25~1.875	1.875~2.5	2.5~3.125	3.125~3.75	3.75~4.375	4.375~5

### 3.2 仿真结果及分析

本研究以标号1、2、4、5、7、8、10、11的样本为已知样本,其余的为测试样本。

采用朴素贝叶斯和权重分析法的仿真结果如表3~4所示。

表3 朴素贝叶斯仿真结果

样本序号	概率			
	正常	内圈故障	外圈故障	滚动体故障
3	0.972 8	0.005 6	0.012 8	0.008 8
6	0.017 1	0.898 9	0.048 9	0.035 1
9	0.033 9	0.028 9	0.913 1	0.024 1
12	0.047 2	0.062 0	0.104 8	0.786 0

表4 权值分析法仿真结果

样本序号	概率			
	正常	内圈故障	外圈故障	滚动体故障
3	0.703 9	0.108 7	0.109 7	0.077 7
6	0.199 8	0.500 1	0.148 2	0.151 9
9	0.120 1	0.099 8	0.680 2	0.099 9
12	0.092 1	0.120 0	0.103 0	0.684 9

由表3和表4可以看出,样本3、6、9、12的诊断类别都分别为正常、内圈故障、外圈故障、滚动体故障,符合实际情况。

为了更好地说明方法的准确性,本研究采用从各类样本中随机的抽取两个样本为已知样本进行仿真,仿真中已知样本共为8个,测试样本为3个,仿真次数为 $C_3^2 \cdot C_3^2 \cdot C_3^2 \cdot C_3^2$ 即81次,以总的正确率为指标,在81次仿真中,被正确诊断出的样本数记为 $f'$ ,则总正确率为 $f'/81 \times 100\%$ 。

仿真结果如表5所示。

表5 两种方法的仿真结果

方法	正确率/(%)
朴素贝叶斯方法	98.77
权重分析方法	99.07

由表5可以看出,朴素贝叶斯方法和权值分析法都能较好地对4种类型的电机轴承振动信号进行故障诊断,其中朴素贝叶斯方法对未知类型故障的诊断正确率相对本研究提出的权值分析方法稍差一些,但是朴素贝叶斯对样本属于每一类的概率区分性强于本

方法。总体来讲,朴素贝叶斯方法和本研究提出的权值分析法都能很好的诊断电机故障。

## 4 结束语

本研究利用朴素贝叶斯诊断模型和权值分析模型对经过小波包变换及信号重构后的电机故障信号数据进行了分析,通过仿真验证了朴素贝叶斯模型在电机轴承故障诊断中的有效性,以及本研究提出的权值分析模型的可行性。

由于朴素贝叶斯方法和权值分析方法都属于监督学习方法,需要实现选取已知样本,已知样本选取过多会造成信息冗余,过少则会导致信息不足,并且已知样本选取的好坏直接影响到诊断的效果,因此,今后的研究应在如何选取已知样本方面做工作。

随着主动学习研究热潮的兴起,将监督学习与主动学习结合对电机故障进行诊断将会是一个新的研究方向。

### 参考文献(References):

- [1] 王雨虹,郝毅魁. 有指导学习在建立瓦斯灾害信息提取模型中的应用[J]. 工矿自动化,2008(2):8-12.
- [2] CALS H, CAKLR A. Rotor bar fault diagnosis in three phase induction motors by monitoring fluctuations of motor current zero crossing instants[J]. **Electric Power Systems Research**,2007(77):385-392.
- [3] 杨俊,吴建华. 基于支持向量机的异步电机故障诊断系统[J]. 机电工程,2008,25(1):72-74.
- [4] 张建文,左官芳,于江,等. 基于BP网络的异步电机转子断条故障自动识别方法[J]. 工矿自动化,2006(6):8-11.
- [5] 付光杰,李云鹏,杨秀菊. 基于改进粒子群优化算法的电机故障诊断研究[J]. 科学技术与工程,2010,10(4):1001-1004.
- [6] 吴国华,吴秦寒,周晓栋. 一种小波变换的音频零水印算法[J]. 机电工程,2009,26(9):16-19.
- [7] 周将坤,陆森林. 基于EMD平均能量法的滚动轴承故障诊断[J]. 轻工机械,2010,28(2):36-40.
- [8] MOORE A W, ZUER D. Internet Traffic Classification Using Bayesian Analysis Techniques[C]. Proc. ACM Sigmetrics,2005:50-60.
- [9] 王国锋,王子良,秦旭达,等. 基于小波包和径向基神经网络轴承故障诊断[J]. 北京科技大学学报,2004,26(2):184-187.

[编辑:张翔]