基于 CGHMM 的轴承故障音频信号诊断方法

陆汝华[1],段　盛[1],杨胜跃[2],樊晓平[2]
LU Ru-hua[1],DUAN Sheng[1],YANG Sheng-yue[2],FAN Xiao-ping[2]

1.湘南学院 计算机系,湖南 郴州 423000
2.中南大学 信息科学与工程学院,长沙 410075
1.Department of Computer Science,Xiangnan University,Chenzhou,Hunan 423000,China
2.School of Information Science & Engineering,Central South University,Changsha 410075,China
E-mail:luruhua658520@163.com

LU Ru-hua,DUAN Sheng,YANG Sheng-yue,et al.Continuous Gaussian mixture HMM based acoustic fault diagnosis scheme for bearings.Computer Engineering and Applications,2009,45(11):223-225.

Abstract: Plentiful significant information about the operation status of bearings,which is potential for the fault diagnose after processed properly,is contained in their acoustic signals.In this paper,a new fault diagnosis scheme using acoustic signals is proposed for the bearings by introducing Continuous Gaussian mixture Hidden Markov Model(CGHMM) method,in which the data processing error due to vector quantization is avoided,and therefore the diagnosis precision is improved.Besides,a clustering algorithm and a scaled coefficient algorithm are introduced for parameters initiation and the forward and backward algorithms to simplify the complexity in the computation and improve the training and recognizing speed and diagnosis precision.At last,experiment results of a diagnosis precision achieve to 98.75% and demonstrate the feasibility and potential for applications of the presented scheme.
Key words: bearing;fault diagnosis;Continuous Gaussian mixture Hidden Markov Model(CGHMM);acoustic signal

摘　要:轴承音频信号包含其运行状态的重要信息,通过分析这些信息就能对轴承故障进行有效诊断。率先引入基于连续高斯混合密度隐马尔可夫模型的轴承故障音频诊断方法,避免矢量量化带来的数据处理误差,提高了系统诊断精度;引入基于聚类算法的模型参数初始化方法和标定系数的前向-后向算法,简化系统复杂度,加快了训练和诊断速度,进一步提高了诊断精度。实验结果表明,诊断精度达到98.75%,具有很好的应用前景。
关键词:轴承;故障诊断;连续高斯混合密度隐马尔可夫模型;音频信号
DOI:10.3778/j.issn.1002-8331.2009.11.067　**文章编号**:1002-8331(2009)11-0223-03　**文献标识码**:A　**中图分类号**:TP18

1　引言

轴承运行正常与否直接影响整台机器的性能,因而轴承故障诊断成了重要的研究课题和目前的研究热点[1-2]。在轴承故障诊断研究中,通常是对其工作时产生的振动信号[1]或音频信号[2]进行分析,以判断轴承运行状态。振动信号法通过安装在轴承座或箱体适当地方的加速度传感器获取轴承振动信号,并对其进行分析与处理,进而判断轴承是否运行正常。此方法的不足在于需要将加速度传感器固定在待检测的设备上,增加成本,使用不便。音频信号的采集属于非接触式,只需要利用麦克风作为声音传感器,使用方便,成本低廉,具有振动信号不可代替的优势。文献[2]研究表明,当轴承运行状态发生变化,音频信号特性也会随之变化,因而对音频信号分析是一种有效、可行的轴承故障诊断方法。目前,基于音频信号的轴承故障诊断方法主要有:小波分析、神经网络和盲源分离方法等。

HMM 是一种描述随机过程统计特性的概率模型,在系统表象与内在状态之间建立了一种概率描述。能够对多个观察样本进行有效融合而构成一个模型,具有较好的抗噪能力。HMM 建模方法在交通监测、图像识别、语音识别[3-5]以及基于振动信号的故障诊断[1]等领域中都得到了较好的应用,也是目前为止最有效的语音识别方法。

本文率先将 HMM 方法应用于轴承故障音频诊断,采用 CGHMM 对输出概率给予更合理的描述,同时避免矢量量化带来的数据处理误差,提高系统诊断精度;进一步引入基于聚类算法的模型参数初始化方法和标定系数的前向-后向算法,简化模型复杂度,加快训练和诊断速度,同时提高系统诊断精度。

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.60774023);湖南省自然科学基金(the Natural Science Foundation of Hunan Province of China under Grant No.07JJ6107);湖南省教育厅科研项目(No.07C723)。
作者简介:陆汝华,女,主要研究方向为智能信息处理;段盛,男,副教授;杨胜跃,男,博士,副教授,主要研究方向为智能信息处理、先进控制理论;樊晓平,男,博士,教授,博士生导师,主要研究方向为机器人系统建模与控制、智能信息处理、智能交通系统。
收稿日期:2008-02-28　**修回日期**:2008-04-01

2 理论基础

2.1 连续高斯混合密度马尔可夫模型

根据输出序列的分布特点,HMM 模型可分为离散型和连续型两大类。离散 HMM(Discrete HMM,DHMM)通过矢量量化技术将输出序列空间限定为有限个码本,带来了量化误差,会在一定程度上影响系统的诊断精度。本文采用 CGHMM 进行建模与诊断,其输出序列不存在量化处理,能够比较精确地表示原始信号,有利于提高诊断精度;利用高斯混合密度函数对输出概率进行描述,能够减少模型存储空间,降低运算复杂度。在引入本文模型之前给出如下定义:

定义 1 HMM 状态数为 N,记为 $S=S_1,S_2,\cdots,S_N$。

定义 2 t 时刻 Markov 链所处的状态记为 q_t, 显然,$q_t\in(S_1,S_2,\cdots,S_N)$。

定义 3 t 时刻观察值(或输出序列)记为 O_t。

有了如上定义,CGHMM 可描述为:

$$\lambda=(\pi,\boldsymbol{A},\boldsymbol{B}) \tag{1}$$

式中,$\pi=\{\pi_i|1\leqslant i\leqslant N\}$为初始概率分布,$\pi_i=P(q_1=s_i)$表示 Markov 链从状态 i 开始的概率, 显然 $\sum_{i=1}^{N}\pi_i=1$ 且 $\pi_i\geqslant 0$;$\boldsymbol{A}=\{a_{ij}|1\leqslant i,j\leqslant N\}$为状态转移概率矩阵,$a_{ij}=P(q_{t+1}=S_j/q_t=S_i)$表示从状态 i 变化到状态 j 的转移概率,显然 $\sum_{i=1}^{N}a_{ij}=1$ 且 $a_{ij}\geqslant 0$;$\boldsymbol{B}=\{b_j(O_t)|1\leqslant j\leqslant N,1\leqslant t\leqslant T\}$为观察值概率矩阵,$b_j(O_t)=P(O_t/q_t=S_j)$表示进入状态 j 时输出为 O_t 的概率,满足:

$$\begin{cases} b_j(O_t)=\sum_{m=1}^{M}\omega_{jm}N(O_t,\boldsymbol{\mu}_{jm},\boldsymbol{\sigma}_{jm}^2) \\ \sum_{m=1}^{M}\omega_{jm}=1 \end{cases} \tag{2}$$

其中,$N(O_t,\boldsymbol{\mu}_{jm},\boldsymbol{\sigma}_{jm}^2)$多维高斯概率密度函数,$M$ 为混合概率密度函数的个数,T 为观察次数。$\boldsymbol{\mu}_{jm}$、$\boldsymbol{\sigma}_{jm}^2$ 和 ω_{jm} 分别表示状态 j 时第 m 个混合高斯概率密度函数的均值矢量、协方差矩阵和权值。

2.2 CGHMM 的训练与诊断

HMM 模型训练是指从同类故障的大量音频信号样本中提取统计信息,利用恰当的训练算法对模型参数反复修正直至收敛,最后得到模型的初始状态概率 π_i、状态转移概率矩阵 $\boldsymbol{a}_{ij}$、混合高斯函数权值 ω_{jm}、均值$\boldsymbol{\mu}_{jm}$与协方差矩阵 $\boldsymbol{\sigma}_{jm}^2$ 等参数。

设有 L 个训练样本,经 Mel 频率倒谱系数(Mel-Frequency Cepstrum Coefficients,MFCC)特征参数提取[5],得到 L 个观察值序列集合$\{O^{(1)},O^{(2)},\cdots,O^{(L)}\}$,每个观察值序列的长度都为 T,那么,$O^{(l)}=\{O_1^l,O_2^l,\cdots,O_T^l\}$,$l=1,2,\cdots,L$。对任一样本观察序列 $O^{(l)}$,根据前向-后向算法[3]计算出前向概率变量 $\alpha_t^{(l)}(t)$,$1\leqslant t\leqslant T$,$1\leqslant i\leqslant N$ 和后向概率变量 $\beta_t^{(l)}(j)$,$1\leqslant t\leqslant T$,$1\leqslant j\leqslant N$ 然后得出过渡概率 $\varepsilon_t^{(l)}(i,j)$和混合输出概率 $\gamma_t^{(l)}(j,m)$:

$$\begin{cases} \varepsilon_t^{(l)}(i,j)=P(q_t=S_i,q_{t+1}=S_j/O^{(l)},\lambda) \\ \gamma_t^{(l)}(j,m)=\dfrac{\alpha_t^{(l)}(j)\beta_t^{(l)}(j)}{\sum_{i=1}^{N}\alpha_t^{(l)}(i)\beta_t^{(l)}(i)}\,\dfrac{\omega_{jm}N(O_t^{(l)},\boldsymbol{\mu}_{jm},\boldsymbol{\sigma}_{jm}^2)}{\sum_{k=1}^{M}\omega_{jk}N(O_t^{(l)},\mu_{jk},\sigma_{jk}^2)} \end{cases} \tag{3}$$

相应地,基于 L 个观察值序列的 B-W 参数修正算法为:

$$\bar{\pi}_i=\sum_{l=1}^{L}\sum_{j=1}^{N}\varepsilon_1^{(l)}(i,j) \tag{4}$$

$$\bar{a}_{ij}=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T-1}\varepsilon_t^{(l)}(i,j)}{\sum_{l}^{L}\sum_{t=1}^{T-1}\sum_{k=1}^{N}\varepsilon_t^{(l)}(i,k)} \tag{5}$$

$$\bar{\mu}_{jm}=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T}\gamma_t^{(l)}(j,m)O_t^{(l)}}{\sum_{l=1}^{L}\sum_{t=1}^{T}\gamma_t^{(l)}(j,m)} \tag{6}$$

$$\bar{\sigma}_{jm}^2=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T}\gamma_t^{(l)}(j,m)(O_t^{(l)}-\mu_{jm})(O_t^{(l)}-\mu_{jm})'}{\sum_{l=1}^{L}\sum_{t=1}^{T}\gamma_t^{(l)}(j,m)} \tag{7}$$

$$\bar{\omega}_{jm}=\frac{\sum_{l=1}^{L}\sum_{t=1}^{T}\gamma_t^{(l)}(j,m)}{\sum_{l=1}^{L}\sum_{t=1}^{T}\sum_{k=1}^{M}\gamma_t^{(l)}(j,k)} \tag{8}$$

训练完成之后将模型参数存储,系统就具备了对训练故障诊断的能力。

在诊断环节,离线或在线读入待检测轴承音频信号,经过预处理和 MFCC 特征参数提取后,得到观察值序列 $O=\{O_1,O_2,\cdots,O_T\}$。然后采用前向-后向算法计算出各 HMM 模型下输出 O 的概率 $P(O/\lambda)$,一般情况下,概率最大的模型即为诊断结果。为了进一步提高系统的诊断精度,可在后处理阶段辅以必要的拒识算法,比如设定适当的概率阈值,如果最大概率小于这个阈值,则诊断为其他状态。

3 训练与诊断算法的改进

连续 HMM 的主要缺陷是模型运算量较大,复杂度较高,不仅会降低系统运行速度,而且数据的临界问题还会降低系统的诊断精度。因此,对训练与诊断算法做了如下改进。

(1)基于聚类算法的模型参数初始化方法

训练 HMM 模型,首先需要初始化模型参数。一般认为,参数初始化可以随机选取。但由公式(4)~(8)可以看出,重估算法比较复杂,恰当的初值能够减少重估次数,加快训练的收敛速度。另外,B-W 算法是一种局部最优算法,其训练结果与模型初值有很大关系,如果初值选取不当,直接影响到系统的诊断效果。因此,采用聚类方法对 ω_{jm} $\boldsymbol{\mu}_{jm}$ 和 $\boldsymbol{\sigma}_{jm}^2$ 进行初始化估计,具体算法如下:

步骤 1 将所有观察值序列 $O_t^{(l)}$($1\leqslant t\leqslant T,1\leqslant l\leqslant L$)平均分成 N 段,每段对应于一个 HMM 状态,记为 X_i($1\leqslant i\leqslant N$),并给 i 赋初值 $i\leftarrow 1$;设置当前循环数 P 和最大循环数 K,并给 P 赋初值 $P\leftarrow 1$;

步骤 2 在状态 X_i 中,定义一个类空间$\{Y_j^i|1\leqslant j\leqslant M\}$,并记类 Y_j^i的长度为 N_j^i,N_j^i初值为零,根据经验或等间隔选取 M 个点,分别作为类 $Y_1^i,Y_2^i,\cdots,Y_M^i$的中心点 $C_1,C_2,\cdots,C_M$;

步骤 3 将 $O_t^{(l)}$分配到类 Y_k^i;同时,$N_j^i\leftarrow N_j^i+1$;

其中,

$$k=\arg\min_{1\leqslant j\leqslant M}(\|O_t^{(l)}-C_j\|),\quad O_t^{(l)}\in X_i,j=1,\cdots,M \tag{9}$$

$\| O_t^{(l)}-C_j \|$表示观察值序列$O_t^{(l)}$到C_j的距离。

步骤4 将每个类的均值向量作为该类新的中心点：

$$C_j=\frac{1}{N_j^i}\sum_{O_t^{(l)}\in Y_j^i} O_t^{(l)}, 1\leqslant j\leqslant M \tag{10}$$

步骤5 计算X_i中所有观察值序列到$O_t^{(l)}$其对应类中心值的误差平方和：

$$J_e=\sum_{j=1}^{M}\sum_{O_t^{(l)}\in Y_j^i}\| O_t^{(l)}-C_j \|^2 \tag{11}$$

如果J_e基本没有变化，则转入步骤7；

步骤6 如果$P<K$，则$P\leftarrow P+1$，同时转入步骤3；

步骤7 根据：

$$\begin{cases}\boldsymbol{\mu}_{im}=\dfrac{\sum_{O_t^{(l)}\in Y_m^i} O_t^{(l)}}{N_m^i}\\ \boldsymbol{\sigma}_{im}^2=\dfrac{\sum_{O_t^{(l)}\in Y_m^i}(O_t^{(l)}-\boldsymbol{\mu}_{im})^2}{N_m^i}\\ \boldsymbol{\omega}_{im}=\dfrac{N_m^i}{\sum_{j=1}^{M} N_j^i}\end{cases} \tag{12}$$

估计模型参数，其中，$1\leqslant m\leqslant M$；

步骤8 $i+1\rightarrow i$，如果$i\leqslant N$，则转入步骤2；

步骤9 算法结束。

（2）基于标定系数的前向-后向算法

在一般的前向-后向算法中，前向概率变量和后向概率变量通过递归运算求得，每个参与运算的乘积因子都小于1，递归的过程使得变量越来越小，甚至会超出计算动态范围的下限，多次递归可能会使数据无限地趋向于零，造成计算中数据的下溢问题，从而影响了系统的诊断效果。因此，本文引入了标定系数方法，对每一步递归运算的结果除以一个只与时间有关而独立于状态的因子，使得每一步的变量基本保持不变。另外，通过实验发现，引入标定系数能够减小运算的时间复杂度。

记前向概率、后向概率分别为$\alpha_t(i)$，$\beta_t(i)$（$1\leqslant t\leqslant T$，$1\leqslant i\leqslant N$）；标定后的前、后向概率分别为$\tilde{\alpha}_t(i)$，$\tilde{\beta}_t(i)$。基于标定系数的前向算法可描述为：

$$\begin{cases}\alpha_1(i)=\pi_i b_i(O_1)\\ \tilde{\alpha}_1(i)=\alpha_1(i)/\sum_{j=1}^{N}\alpha_1(j)\end{cases} \tag{13}$$

$$\begin{cases}\alpha_t(i)=\sum_{j=1}^{N}\tilde{\alpha}_{t-1}(j)a_{ji}b_i(O_t), 2\leqslant t\leqslant T\\ \tilde{\alpha}_1(i)=\alpha_t(i)/\alpha_t(j)\end{cases} \tag{14}$$

相应地，基于标定系数的后向算法为：

$$\tilde{\beta}_t(j)=1/N \tag{15}$$

$$\begin{cases}\beta_t(i)=\sum_{j=1}^{N}a_{ij}b_j(O_{t+1})\tilde{\beta}_{t+1}(j), 1\leqslant t\leqslant T-1\\ \tilde{\beta}_t(i)=\beta_t(i)/\sum_{j=1}^{N}\alpha_{t+1}(j)\end{cases} \tag{16}$$

4 轴承故障诊断实验

在Visual C++ 7.0环境下，自主开发了基于HMM的音频故障诊断平台，本文所有实验均是在此平台上完成。对于本文实验，诊断对象选定为6 202 CM深沟球轴承，轴承转速为1 800 r/m，HMM模型的状态数选为7，混合高斯概率密度函数个数定为3。

在模型训练环节，对于正常、内圈异音、外圈异音以及滚动体异音等四种轴承状态，各采集30组音频数据样本进行训练。为了验证本文改进算法的有效性，先后采用一般的多观测序列B-W算法和改进算法对CGHMM模型进行训练。另外，利用相同的数据样本对离散型HMM（DHMM）模型进行训练[6]。分别得到三类训练过程对于每种故障模型的平均训练时间如表1所示。可以看出，由于DHMM对观测序列进行了量化处理，计算量小，训练速度快；CGHMM的复杂度比较高，采用一般的B-W算法收敛过程长。而采用本文改进措施后，大大缩短了训练时间，只需24.8 s，约为原来的十分之一，相对DHMM也不到四分之一。

表1　训练时间比较

模型与算法		训练时间/s
DHMM		111.6
CGHMM	一般B-W算法	247.3
	改进算法	24.8

在诊断环节，另外采集20组内圈异音、外圈异音、滚动体异音以及10组正常和其他状态音频数据用于诊断。得到基于CGHMM的故障诊断结果如表2所示，共80次诊断只出现1次误诊，正确率达到了98.75%。在诊断过程中采用了标定系数前向-后向算法，诊断时间由采用前的1.62 s降低到1.58 s。

表2　诊断结果

故障类型	诊断次数	诊断结果					诊断精度/(%)
		内圈异音	外圈异音	滚动体异音	正常	其他	
内圈异音	20	20	0	0	0	0	100
外圈异音	20	0	20	0	0	0	100
滚动体异音	20	0	0	19	0	1	95
正常	10	0	0	0	10	0	100
其他	10	0	0	0	0	10	100

另外，还利用上一环节训练好的DHMM模型和一般B-W算法的CGHMM模型完成了诊断实验，其平均诊断精度分别为90%和96.3%，明显低于本文改进算法CGHMM模型的音频故障诊断方法。

5 结束语

将基于音频信号的HMM建模方法率先引入到轴承故障诊断研究中，有效利用了HMM对随机信号超强的建模能力。采用CGHMM能避免矢量量化带来的数据处理误差，提高了系统诊断效果；利用基于聚类算法的模型参数初始化方法和标定系数的前向-后向算法对训练和诊断算法加以改进，简化了系统运算复杂度，加快训练和诊断速度，同时也进一步提高了诊断精度。实验结果表明，该方法对轴承运行状态的诊断精度达到98.75%，诊断时间约为1.58 s。

（下转234页）