第 36 卷第 6 期
2016 年 12 月

振动、测试与诊断
Journal of Vibration, Measurement & Diagnosis

Vol. 36 No. 6
Dec. 2016

doi:10.16450/j.cnki.issn.1004-6801.2016.06.027

自适应邻域构造流形学习算法及故障降维诊断*

张晓涛， 唐力伟， 王 平， 邓士杰
（军械工程学院火炮工程系 石家庄,050003）

摘要 针对流形学习算法中近邻构造问题,提出一种自适应邻域构造方法,该方法基于马氏距离计算样本间相似系数,由相似系数均值确定初始近邻数,根据样本高斯核概率密度估计调整近邻数,并将自适应邻域构造方法用于改进的主成分分析联合局部保持投景(principal component analysis-locality preserving projections,简称 PCA-LPP)流形学习算法中。通过齿轮箱故障类型识别对其特征降维性能进行验证,结果表明,自适应邻域 PCA-LPP 方法比传统的 k 近邻方法及原始无处理的特征识别率都高,可以达到 94.67%。

关键词 自适应；邻域；高斯核概率估计；流形学习；故障诊断
中图分类号 TH165；TN911

引 言

流形学习算法是一类非常有效的非线性数据降维算法[1],在很多具体的流形学习算法实现中,诸如局部保持投影(locality preserving projection,简称 LPP)[2]、拉普拉斯映射(Laplacian eigenmap,简称 LE)[3]都需要先构造数据样本近邻图。传统的近邻图构造主要是 k 近邻和 ε 近邻两种方法[4],其主要缺点在于各个样本的近邻数需要人为经验确定,且所有样本采用相同的近邻数,不能针对样本自身的邻域分布情况进行数据调整。自适应邻域的思想主要是基于样本周围一定范围数据分布的情况,按照预设规则自动确定邻域大小。目前常用的自适应邻域构造思想主要基于欧氏距离或者余弦夹角距离计算相似系数,并根据相似系数的分布设定阈值自适应求取近邻数,这样的方法多用于图像处理及人脸识别研究中。Yang 等[5]基于欧氏距离指数函数计算样本相似系数。刘凤连等[6]基于图像欧氏距离导数计算样本相似系数。黄璞等[7]采用样本夹角余弦距离作为相似系数。三种方法在人脸识别研究中取得了良好的分类效果。李城梁等[8]将基于样本切空间距离的自适应邻域方法应用于机械故障信号特征提取,故障分类正确率得到明显提升。

自适应邻域构造方法能够有效避免人为选择近邻范围的随意性,具有更好的样本局部非线性流形特征表达能力。笔者提出一种自适应邻域构造方法,基于马氏距离相似系数均值及样本高斯核概率密度估计调整自适应获得样本近邻数,并将自适应邻域构造方法应用于 PCA-LPP 流形学习算法,通过齿轮箱故障信号特征向量降维识别验证了自适应邻域 PCA-LPP 方法的有效性。

1 自适应邻域构造方法

设高维空间 R^D 中存在数据集 $\boldsymbol{X}=[x_1,x_2,\cdots,x_N]$,基于近邻图构造相似矩阵 $\boldsymbol{S}$ 能够用于多种流形学习算法,实现数据的降维处理。

传统的 k 近邻构造方法,基于样本间的欧氏距离计算样本相似性,每个样本近邻值相同,相似系数具体表达如下

$$s_{ij}=\begin{cases}\alpha_{ij} & (x_i \text{ 和 } x_j \text{ 互为近邻}) \\ 0 & (\text{其他})\end{cases} \tag{1}$$

其中:$\alpha_{ij}=\exp(-\|x_i-x_j\|^2/\beta)$,$\|x_i-x_j\|^2$ 为样本间欧氏距离;β 为所有样本之间欧氏距离均值的平方。

k 近邻法存在两点不足:a. 所有样本采用相同的近邻数,不能很好地适应每个样本的局部流形结构特征;b. 近邻数的选择没有成熟的算法,常使用的交叉验证方法存在效率低下等弊端。

1.1 马氏距离相似度衡量

马氏距离是一种基于数据协方差的距离描述,其定义为

* 国家自然科学基金资助项目(50775219)
收稿日期:2014-11-12;修回日期:2015-03-27

$$d_{ij}=\sqrt{(x_i-x_j)\boldsymbol{C}^{-1}(x_i-x_j)^{\mathrm{T}}} \tag{2}$$

其中：$\boldsymbol{C}$ 为数据集 $\boldsymbol{X}$ 的协方差矩阵。

马氏距离对一切非奇异性线性变换具有不变性，不受特征量纲选择的影响。马氏距离与欧氏距离的对比如图 1 所示，图中虚线圆为欧氏距离等距线，实线椭圆为马氏距离等距线。由图 1 可以看出，马氏距离的描述更符合数据真实的分布情况，仅当协方差矩阵 $\boldsymbol{C}$ 为单位矩阵时，马氏距离与欧氏距离才会相等[9]。

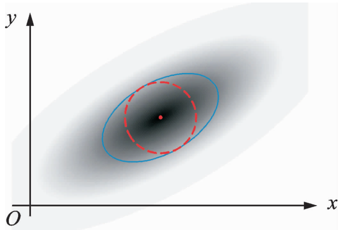


图 1　马氏距离与欧氏距离

Fig. 1　Mahalanobis distance and Euclidean distance

基于马氏距离数据分布特征描述方面的优势，提出基于马氏距离的样本相似矩阵构造方法，相似矩阵元素表达如下

$$s_{ij}=\begin{cases}\alpha_{ij} & (\alpha_{ij}>M_i=(\sum_{j=1}^{N}\alpha_{ij})/N)\\ 0 & (\text{其他})\end{cases} \tag{3}$$

其中：$\alpha_{ij}=\exp(-d_{ij}/\beta)$；$\beta$ 为所有样本之间马氏距离均值的平方。

在式(3)中，样本的近邻由样本 x_i 与所有样本的相似系数均值 M_i 进行界定，$M_i=(\sum_{j=1}^{N}\alpha_{ij})/N$，当样本 x_i 与 $x_j\,(j=1,2,\cdots,N)$ 的相似系数 α_{ij} 大于 M_i 时，x_j 是 x_i 的近邻，否则二者不是近邻。基于相似系数均值 M_i 确定样本近邻是一种自适应方法，该方法可以得到每一个样本的近邻数 k_i，不同样本的近邻数 k_i 一般是不相等的。

1.2　核密度估计调整

样本近邻构造时，空间中密集分布的样本局部特征往往较为相似。对样本 x_i 来讲，其附近区域其他样本出现的概率密度越大，表示与其具有相似局部特征的样本越多，则 x_i 对应的样本近邻数应该大一些。据此提出将高斯核密度估计(Gauss kernel density estimation，简称 GKDE)[10]用于近邻图构造，对基于相似系数均值 M_i 的自适应近邻结果进行修正调整。

高斯核密度估计是一种非参数概率密度估计方法，能够在样本分布先验知识未知的情况下，根据样本自身信息估计总体分布概率密度。高维空间 R^D 中样本 x_i 的邻域概率密度高斯核估计为

$$p(x_i)=\frac{1}{k_i^D}\sum_{j\in k_i}\left[\frac{1}{N}\frac{1}{(\sqrt{2\pi})^D}\exp\left(-\frac{d_{ij}^2}{2k_i^2}\right)\right] \tag{4}$$

根据式(4)可得到所有样本的邻域概率密度估计值 $p(x_i)$，由此计算数据集 $\boldsymbol{X}$ 全部样本的平均邻域概率密度 $\bar{p}=[\sum_{j=1}^{N}p(x_i)]/N$，基于邻域分布概率的样本 x_i 近邻数调整如下

$$k_i^p=\mathrm{floor}(k_i p(x_i)/\bar{p}) \tag{5}$$

其中：floor 表示数据向负无穷大方向取整。

通过式(5)可知，当 x_i 附近数据样本分布密度大时，该点的近邻数自动调整增大，当其附近数据分布稀疏时，近邻数自动调整减小。

1.3　自适应邻域构造流程及特性

首先，基于马氏距离 d_{ij} 计算样本间相似系数，通过样本相似系数均值 M_i 确定每个样本的初始近邻数 k_i；其次，根据高斯核密度估计 GKDE 得到样本的近邻分布概率密度 $p(x_i)$，并依据 $p(x_i)$ 和 $\bar{p}$ 对初始近邻数调整修正；最后，得到最终的样本近邻数 k_i^p，从而构造最终的相似矩阵 $\boldsymbol{S}$。自适应邻域构造流程见图 2。

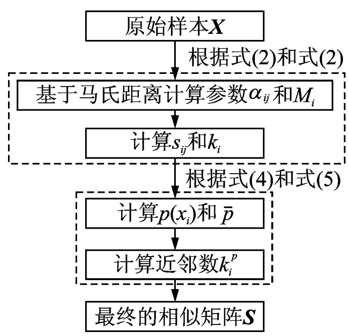


图 2　自适应邻域构造流程

Fig. 2　Flow chart of adaptive neighborhood selection

自适应邻域构造方法最终得到的相似矩阵 $\boldsymbol{S}$ 一般情况下是非对称的，主要包含以下几种情况：

1）当样本协方差矩阵为单位阵，各样本分布概率密度相等时，自适应邻域构造方法得到的相似矩阵与 k 近邻方法相同，样本 x_i 和 x_j 互为近邻，且近邻数 $k_i^p=k_j^p$，相似系数 $\alpha_{ij}=\alpha_{ji}$，相似矩阵 $\boldsymbol{S}$ 对称；

2）与第 1 假设相同，当样本 x_i 和 x_j 互相不为近邻时，相似系数 $\alpha_{ij}=\alpha_{ji}=0$，相似矩阵 $\boldsymbol{S}$ 对称；

3）当样本协方差矩阵不是单位阵，各样本分布概率密度不相等时，若样本 x_i 和 x_j 互为近邻，其近

邻数 $k_i^p = k_j^p$，但相似系数 $\alpha_{ij} \neq \alpha_{ji}$，此时的相似矩阵 $\boldsymbol{S}$ 为非对称矩阵。

4) 与第 3 假设相同，当样本 x_j 是 x_i 的近邻，而 x_i 不是 x_j 的近邻时，其近邻系数 $k_i^p \neq k_j^p$，相似系数 $\alpha_{ij} \neq \alpha_{ji}$，此时相似矩阵 $\boldsymbol{S}$ 为非对称。

自适应邻域构造方法与传统 k 近邻方法都是为了表达数据样本局部结构的相似性，主要不同为：

1) k 近邻方法基于欧氏距离计算样本相似系数，自适应邻域构造方法基于马氏距离计算相似系数，能够更好地考虑样本分布特性；

2) k 近邻方法所有样本近邻数需人为设置，自适应近邻构造方法根据相似系数均值 M_i 确定样本初始近邻数，并采用样本分布概率密度调整近邻数；

3) k 近邻方法所有样本近邻数均相等，自适应近邻构造方法每个样本的近邻数不一定相等，且一般都不相等；

4) k 近邻方法得到的相似矩阵 $\boldsymbol{S}$ 是对称的，自适应近邻构造方法得到的相似矩阵 $\boldsymbol{S}$ 一般不对称。

1.4 自适应邻域构造实例

采用 26 个随机二维数据样本，说明自适应邻域构造方法的特点。采用自适应邻域构造方法寻找样本近邻，以样本点 $A(1,4)$ 和点 $B(8.5,5)$ 的近邻求解进行说明，并给出 k 近邻方法的对比，设定 $k=8$。表 1 为自适应近邻求解过程中的参数及结果。图 3 中 4 幅图所示为原始 26 个样本点及点 A 和点 B 的 3 种近邻求解结果。

表 1 近邻求解参数及结果

Tab. 1 Solution paramters and results of neighborhood

求解情况	点 A	点 B
初始近邻数 k_i	8	11
GKDE 调整参数 $p(x_i)/\bar{p}$	0.510 0	1.136 1
最终近邻数 k_i^p	4	12

由图 3(a)可知，原始 26 个样本点非均匀分布，其中点 A 处数据样本分布稀疏，而点 B 处数据样本分布稠密。图 3(b)为 $k=8$ 时 k 近邻方法得到的近邻结果，点 A 和点 B 的近邻数相等，无法反映数据样本分布的疏密情况。图 3(c)为依据式(3)得到的初始近邻分布，从图中可以看到，点 A 近邻样本数为 8，点 B 近邻样本数为 11。点 A 和点 B 的近邻样本数已经考虑到样本分布的疏密程度。采用高斯核密度估计方法，得到点 A 和点 B 的近邻调整参数 $p(x_i)/\bar{p}$ 分别为 0.51 和 1.136 1，由调整参数可以明显看到点 A 和点 B 处样本分布密度的不同。经过调整后得到的最终近邻分布如图 3(d)所示，其中点 A 和点 B 近邻样本数分别为 4 和 12。图 3(c)到图 3(d)的变化说明了自适应邻域构造方法能够根据样本分布密度对近邻数进行调整，当数据样本分布密度大时，近邻数自动调整增大，反之亦然。

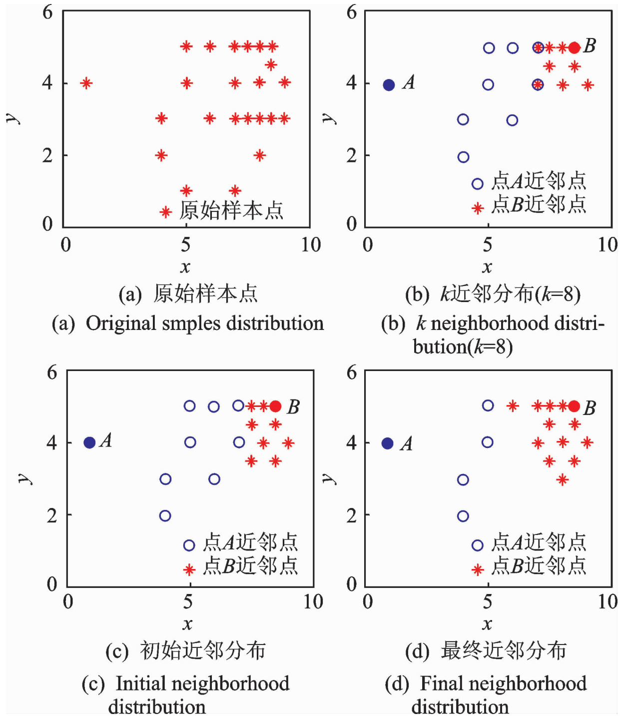


图 3 点 A 和点 B 近邻结果分布

Fig. 3 Neighborhood for point A and point B

2 自适应邻域 PCA-LPP 流形学习算法

自适应邻域构造方法能够用于多种基于近邻图构造的流形学习算法，笔者将其应用于一种结合主元分析与局部保持投影的 PCA-LPP 改进流形学习算法中，标准局部保持投影算法中相似矩阵 $\boldsymbol{S}$ 为对称矩阵，其对称性对算法的求解带来很大便利。但自适应邻域构造方法得到的相似矩阵 $\boldsymbol{S}$ 一般为非对称矩阵，其在 PCA-LPP 中的求解推导如下。

PCA-LPP 流形学习目的在于将高维空间 R^D 中的数据集 $\boldsymbol{X}=[x_1,x_2,\cdots,x_N]$，通过投影矩阵 $\boldsymbol{W}$ 转换投影到低维投影空间 R^d ($d<D$)中的低维数据集 $\boldsymbol{Y}=[y_1,y_2,\cdots,y_N]$，且满足运算 $\boldsymbol{Y}=\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}$。该算法的目标函数构造在于寻找求解投影矩阵 $\boldsymbol{W}$。

PCA 方法寻找数据分布方差最大的坐标系，其全局目标函数为

$$J_g(\boldsymbol{W})=\max_{\boldsymbol{W}}\sum_{i=1}^{N}\|y_i-y_j\|^2= \max_{\boldsymbol{W}}\sum_{i=1}^{N}\boldsymbol{W}^{\mathrm{T}}(x_i-\bar{x})(x_i-\bar{x})^{\mathrm{T}}\boldsymbol{W}= \max_{\boldsymbol{W}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{W} \tag{6}$$

其中：$\bar{x}=(\sum_{i=1}^{N}x_i)/N$;$\boldsymbol{C}=(\sum_{i=1}^{N}(x_i-\bar{x})(x_i-\bar{x})^{\mathrm{T}})/N$。

LPP 是一种非线性局部流形保持投影算法,能够保持投影前后数据样本间的局部结构相似,其局部目标函数为

$$\begin{aligned}J_l(\boldsymbol{W})=&\min_{\boldsymbol{W}}\sum_{i=1}^{N}\|y_i-y_j\|^2 s_{ij}=\\&\min_{\boldsymbol{W}}\sum_{i,j=1}^{N}(y_i^2+y_j^2-y_iy_j-y_jy_i)s_{ij}=\\&\min_{\boldsymbol{W}}(\sum_{i=1}^{N}y_i^2\boldsymbol{D}_{ii}+\sum_{j=1}^{N}y_j^2\boldsymbol{D}_{jj}-\\&\sum_{i,j=1}^{N}y_iy_js_{ij}-\sum_{i,j=1}^{N}y_jy_is_{ij})=\\&\min_{\boldsymbol{W}}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D}_{ii}+\boldsymbol{D}_{jj}-\boldsymbol{S}-\boldsymbol{S}^{\mathrm{T}})\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W})=\\&\min_{\boldsymbol{W}}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{D}'-\boldsymbol{S}')\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W})=\\&\min_{\boldsymbol{W}}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{L}')\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W})\end{aligned}\tag{7}$$

其中：$\boldsymbol{L}'$ 为拉普拉斯矩阵；$\boldsymbol{D}_{ii}$ 和 $\boldsymbol{D}_{jj}$ 为对角阵,其元素为 $\boldsymbol{D}_{ii}=\sum_j s_{ij}$,$\boldsymbol{D}_{jj}=\sum_i s_{ij}$;$\boldsymbol{D}'=\boldsymbol{D}_{ii}+\boldsymbol{D}_{jj}$;$\boldsymbol{S}'=\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}}$ ；s_{ij} 为相似系数,构成相似矩阵 $\boldsymbol{S}$;W 为投影矩阵。

PCA-LPP 流形学习算法兼顾 PCA 的全局分布方差最大特性及 LPP 的局部流形保持特性,能够全面刻画数据的整体和局部特征,其全局目标函数为

$$\begin{cases}J(\boldsymbol{W})=\max\limits_{\boldsymbol{W}}(J_g(\boldsymbol{W})-J_l(\boldsymbol{W}))=\\\qquad\max\limits_{\boldsymbol{W}}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{W}-\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W})=\\\qquad\max\limits_{\boldsymbol{W}}(\boldsymbol{W}^{\mathrm{T}}(\boldsymbol{C}-\boldsymbol{X}(\boldsymbol{D}'-\boldsymbol{S}')\boldsymbol{X}^{\mathrm{T}})\boldsymbol{W})=\\\qquad\max\limits_{\boldsymbol{W}}(\boldsymbol{W}^{\mathrm{T}}(\boldsymbol{C}-\boldsymbol{X}\boldsymbol{L}'\boldsymbol{X}^{\mathrm{T}})\boldsymbol{W})\\\qquad\mathrm{s.t.}\,\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{D}'\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}=\boldsymbol{I}\end{cases}\tag{8}$$

其中：限定条件 $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{D}'\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}=\boldsymbol{I}$ 可消除随机尺度因子的影响。

在 LPP 局部目标函数中,自适应邻域构造方法得到的相似矩阵 $\boldsymbol{S}$ 是非对称矩阵,但局部目标函数 $J_l(\boldsymbol{W})$ 中的矩阵 $\boldsymbol{L}'$ 是对称矩阵,其证明如下。

证明：$\forall$ 相似矩阵 $\boldsymbol{S}$

$\because$ $\boldsymbol{D}_{ii}=\sum_j s_{ij}$,$\boldsymbol{D}_{jj}=\sum_i s_{ij}$ 均为对角阵

$\therefore$ $\boldsymbol{D}_{ii}$ 和 $\boldsymbol{D}_{jj}$ 为对称矩阵

$\therefore$ $\boldsymbol{D}'=\boldsymbol{D}_{ii}+\boldsymbol{D}_{jj}$,$\boldsymbol{S}'=\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}}$ 为对角对称矩阵

$\therefore$ $\boldsymbol{L}'=\boldsymbol{D}'-\boldsymbol{S}'$ 为对称矩阵。

自适应邻域构造 PCA-LPP 算法的目标函数式(8)可以通过拉格朗日乘子法[11],将其转化为有约束最大值问题,如式(9)所示

$$\boldsymbol{L}(\boldsymbol{W})=\boldsymbol{W}^{\mathrm{T}}(\boldsymbol{C}-\boldsymbol{X}\boldsymbol{L}'\boldsymbol{X}^{\mathrm{T}})\boldsymbol{W}-\lambda(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{D}'\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}-1)\tag{9}$$

式(9)对 $\boldsymbol{W}$ 求导并置零可得

$$\frac{\partial\boldsymbol{L}(\boldsymbol{W})}{\partial\boldsymbol{W}}=2(\boldsymbol{C}-\boldsymbol{X}\boldsymbol{L}'\boldsymbol{X}^{\mathrm{T}})\boldsymbol{W}-2\lambda\boldsymbol{X}\boldsymbol{D}'\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}=0\tag{10}$$

化简式(10)可以得到

$$(\boldsymbol{C}-\boldsymbol{X}\boldsymbol{L}'\boldsymbol{X}^{\mathrm{T}})\boldsymbol{W}=\lambda\boldsymbol{X}\boldsymbol{D}'\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\tag{11}$$

由式(11)可知,投影矩阵 $\boldsymbol{W}$ 的求解,实质是广义特征值求解问题,通过求取前 d 个最大特征值对应的特征向量,从而得到投影矩阵 $\boldsymbol{W}$,投影后低维数据通过 $\boldsymbol{Y}=\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}$ 计算。

由此可以得到自适应邻域 PCA-LPP 流形学习算法流程,如图 4 所示。

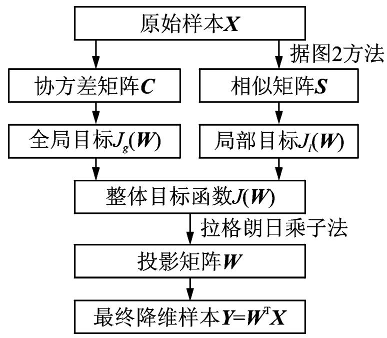


图 4　自适应邻域 PCA-LPP 算法流程

Fig. 4　PCA-LPP algorithm based on adaptive neighborhood

3　齿轮箱故障特征降维分析

流形学习算法在数据可视化以及故障特征降维识别中应用广泛[12-13],下面对自适应邻域 PCA-LPP 流形学习算法的性能进行验证。

采用齿轮箱故障实测声发射信号对 PCA-LPP 的特征降维性能进行验证。试验中齿轮箱故障包括正常状态、轴承内圈故障、外圈故障、内外圈复合故障及齿根裂纹故障 5 种模式,故障轴承为 6206,安装在中间传动轴,故障齿轮为中间传动轴大齿轮,一级传动比为 0.5,齿轮箱结构原理及传感器安装如图 5 所示。试验中齿轮箱空负载运转,其输入轴转速为 1 490 r/min。信号采集设备为北京软岛 DS2-8A 型全息声发射信号分析仪,设置采集仪采样频率为 1 MHz,采集仪硬件滤波参数为 100 kHz～400 kHz 带通滤波,声发射传感器为声华 SR150M 型,匹配 40 dB 前置放大器。每个数据样本长度为 1 s,每种故障模式对应 70 个样本,5 种模式共 350 个样本。

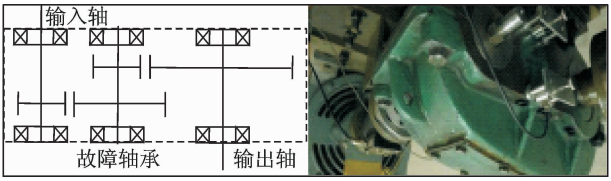


图 5 齿轮箱结构及传感器布置

Fig. 5 Gearbox structure and sensor arrangement

对每个故障信号样本进行 db4 小波包 4 层分解，重构各子频带小波包系数，得到 16 个子频带重构分量，各子频带宽度为 31.25 kHz，计算子频带信号能量，得到能量特征向量 $\boldsymbol{E}=[E_1, E_2, \cdots, E_{16}]$，求取 16 个子频带信号的能量熵作为特征向量[8]，能量熵的计算方法如下。

$$E_{si}=(E_i/\sum_{i=1}^{16}E_i)\log(E_i/\sum_{i=1}^{16}E_i)\quad(i=1,2,\cdots,16) \tag{12}$$

采用自适应邻域 PCA-LPP 方法对所提故障特征向量进行处理，并将降维后特征向量输入支持向量机进行训练识别，对比故障模式的分类识别性能，每个故障 70 个样本中，40 个用于训练支持向量机分类器，30 个用于样本故障分类识别。

3.1 自适应邻域 PCA-LPP 故障特征降维识别分析

齿轮箱故障声发射信号的特征向量原始维度为 16 维，采用自适应邻域 PCA-LPP 对故障特征向量进行降维处理，降维后特征的维度范围为 3～15 维，间隔为 1。不同降维数的特征向量输入支持向量机的整体故障识别率如图 6 所示。由图 6 可以看到，原始的 16 维特征向量经降维处理，不同维度的特征向量对应的故障辨识率有所不同。辨识率在 5 维以后出现明显增高，当降维数 $d=8$ 时，整体故障辨识率最高，达到 94.67%。其中整体故障辨识率指各单项故障识别率的均值，表示 4 种故障类型正确识别的总数占 150 个测试样本的比例。

降维数 $d=8$ 时，不同故障的辨识率如表 2 所示。由表 2 可知，各类故障模式中，齿根裂纹辨识率最高，因为齿根裂纹与其他几种故障具有明显不同，

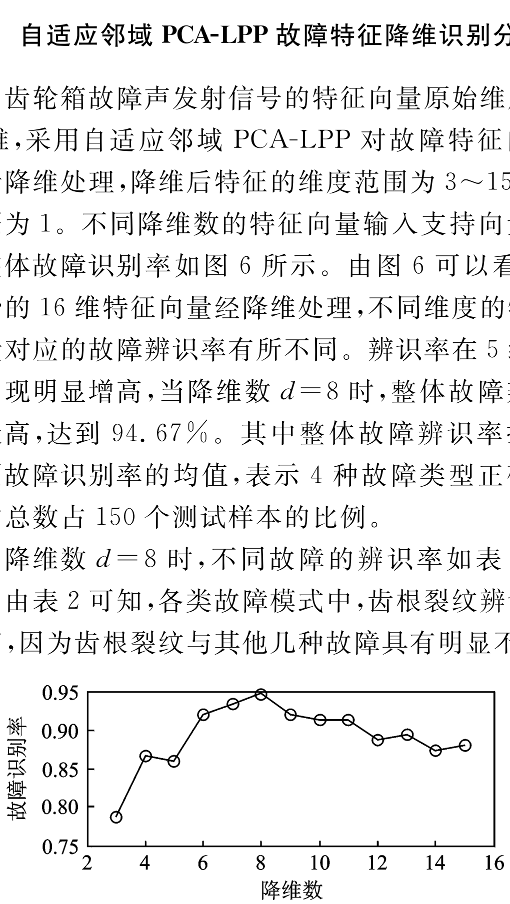


图 6 不同降维数的故障识别率

Fig. 6 Fault identification rate of different dimensions

表 2 自适应邻域故障识别率

Tab. 2 Fault identification rate of adaptive neighborhood

故障类型	齿根裂纹	正常状态	内圈故障	外圈故障	复合故障	整体
识别率/%	100	96.67	93.33	93.33	90.00	94.67

而复合故障辨识率最低，因为复合故障与内、外圈故障具有相似特征成分。

3.2 不同识别方法性能对比

对文中的故障特征向量，将 k 近邻 PCA-LPP 方法降维识别结果、基于欧氏距离的自适应 PCA-LPP 方法降维识别结果以及原始特征向量不做任何处理的识别结果与 3.1 节的识别结果进行对比，k 近邻方法的最优识别结果通过交叉验证获得。设置近邻值 k 的变换范围为 5～30，间隔为 1，降维数 d 的范围为 3～15，间隔为 1，计算后得到 k 近邻方法在 $k=17$，$d=7$ 时获得最高识别率。基于欧氏距离的自适应 PCA-LPP 方法降维识别在降维数 $d=8$ 时取得最高识别率。4 种方法的识别率对比如表 3 所示。

表 3 3 种情况识别率对比

Tab. 3 Fault identification rate of three situations

方法	故障类型					
	齿根裂纹	正常状态	内圈故障	外圈故障	复合故障	整体
自适应	100	96.67	93.33	93.33	90.00	94.67
欧氏距离自适应	93.33	90.00	86.67	90.00	83.33	88.67
k 近邻	93.33	83.33	90.00	86.67	80.00	86.67
原始特征	63.33	56.67	50.00	53.33	43.33	53.33

从表 3 可以看到，原始 16 维特征向量，不经任何处理，直接进行支持向量机分类器的训练和识别，其故障识别率最低，仅为 53.33%，而采用流形学习算法降维之后，故障识别率明显提升。对于同样的 PCA-LPP 流形学习算法，不同的邻域构造方法带来的降维效果差异较大。自适应邻域构造方法降维后特征的故障识别率可达 94.67%，而 k 近邻方法降维后特征的故障识别率为 86.67%。不同相似度衡量方法带来的降维效果也不同，基于马氏距离的相似度衡量最终故障识别率为 94.67%，而基于欧氏距离的相似度衡量最终识别率却为 88.67%。原始特征向量维度高，其中包含故障特征的差异信息，但高维度同时包含过多的冗余信息，因此其故障识别率低。k 近邻方法对所有样本采用相同的近邻数，对各样本局部邻域的描述不能做到分而划之，基于欧氏距离的相似度衡量，对数据样本的分布情况不

能充分考虑，因此两种方法特征向量的降维效果不如自适应邻域构造方法，故其故障识别率低于自适应邻域 PCA-LPP 方法。

4 结束语

针对流形学习算法中近邻图的构造问题，提出一种基于马氏距离的相似矩阵元素计算。根据相似矩阵元素均值确定初始邻域数，并由数据样本邻域高斯核概率密度估计结果调整样本近邻数，解决了自适应构造近邻图的问题。将自适应邻域构造方法应用于兼顾 PCA 与 LPP 二者特性的改进流形学习算法，并给出相应的理论计算模型，证明了求解拉普拉斯矩阵的对称性。通过齿轮箱故障声发射信号 16 维能量熵特征的降维识别对自适应邻域 PCA-LPP 流形学习算法的性能进行验证，结果表明故障特征降到 8 维时识别率最高，可以达到 94.67%。对比研究表明，自适应邻域构造 PCA-LPP 方法的降维识别性能优于 k 近邻 PCA-LPP 方法以及原始特征向量的识别率。

参　考　文　献

[1] 赵振华，郝晓弘. 局部保持鉴别投影及其在人脸识别中的应用[J]. 电子与信息学报，2013，35(2)：463-466.
Zhao Zhenhua, Hao Xiaohong. Linear locality preserving and discriminating projection for face recognition [J]. Journal of Electronics & Information Technology, 2013,35(2): 463-466. (in Chinese)

[2] He Xiaofei, Niyogi P. Locality preserving projections [C] // Neural Information Processing Systems16. Cambridge, USA: MIT Press, 2004:153-160.

[3] 许庆诚，胡建中. 基于改进增量 LE 的压缩机故障特征提取方法[J]. 仪器仪表学报，2013，34(4)：791-796.
Xu Qingcheng, Hu Jianzhong. Fault feature extraction method for compressor based on improved incremental Laplacian eigenmap algorithm[J]. Chinese Journal of Scientific Instrument, 2013, 34(4): 791-796. (in Chinese)

[4] 何强，蔡洪，韩壮志，等. 基于非线性流形学习的 ISAR 目标识别研究[J]. 电子学报，2010，38(3)：585-590.
He Qiang, Cai Hong, Han Zhuangzhi, et al. ISAR target recognition based on non-linear manifold learning[J]. Acta Electronica Sinica, 2010, 38(3): 585-590. (in Chinese)

[5] Yang Bo, Chen Songcan. Sample-dependent graph construction with application to dimensionality reduction[J]. Neurocomputing, 2010, 74(5): 301-314.

[6] 刘凤连，汪日伟，程俊，等. 图像特征提取中领域尺寸和本征维数的自动选择算法[J]. 光电子·激光，2013，24(12)：2416-2420.
Liu Fenglian, Wang Riwei, Cheng Jun, et al. Auto-selected algorithm of the neighborhood size and intrinsic dimension for image feature extraction[J]. Journal fo Optoelectronics · Laser, 2013, 24(12): 2416- 2420. (in Chinese)

[7] 黄璞，唐振民. 无参数局部保持投影及人脸识别[J]. 模式识别与人工智能，2013，26(9)：865-871.
Huang Pu, Tang Zhenmin. Parameter-free locality preserving projections and face recognition[J]. Pattern Recognition and Artificial Intelligence, 2013, 26(9): 865-871. (in Chinese)

[8] 李城梁，王仲生，姜洪开，等. 自适应 Hessian LLE 在机械故障特征提取中的应用[J]. 振动工程学报，2013，26(5)：758-763.
Li Chengliang, Wang Zhongsheng, Jiang Hongkai, et al. Adaptive Hessian LLE in mechanical fault feature extraction[J]. Journal of Vibration Engineering, 2013, 26(5): 758-763. (in Chinese)

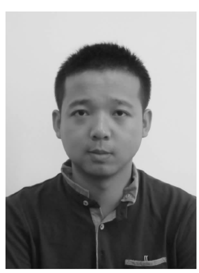
[9] 骆志高，李旭东，赵俊丽，等. 利用马氏距离判别法准确实现对裂纹的识别[J]. 振动与冲击，2013，32(21)：186-188.
Luo Zhigao, Li Xudong, Zhao Junli, et al. Crack identification with mahalanobis distance discrimination method[J]. Journal of Vibration and Shock, 2013, 32(21): 186-188. (in Chinese)

[10] 孙即祥. 现代模式识别[M]. 长沙：国防科技大学出版社，2002：136-140.

[11] 李锋，王家序，杨荣松. 有监督不相关正交局部保持映射故障辨识[J]. 仪器仪表学报，2013，34(5)：1113-1116.
Li Feng, Wang Jiaxu, Yang Rongsong. Fault identification method on supervised uncorrelated orthogonal locality preserving projection[J]. Chinese Journal of Scientific Instrument, 2013, 34(5): 1113-1116. (in Chinese)

[12] 宋涛，汤宝平，李锋. 基于流形学习和 K-最近邻分类器的旋转机械故障诊断方法[J]. 振动与冲击，2013，32(5)：149-153.
Song Tao, Tang Baoping, Li Feng. Fault diagnosis method for rotating machinery based on manifold learning and K-nearest neighbor classifier[J]. Journal of Vibration and Shock, 2013, 32(5): 149-153. (in Chinese)

[13] 刘忠宝，潘广贞，赵文娟. 流形判别分析[J]. 电子与信息学报，2013，35(9)：2047-2050.
Liu Zhongbao, Pan Guangzhen, Zhao Wenjuan. Manifold-based discriminant analysis[J]. Journal of Electronics & Information Technology, 2013, 35(9): 2047-2050. (in Chinese)



第一作者简介：张晓涛，男，1987 年 5 月生，博士生。主要研究方向为机械系统性能检测与故障诊断。曾发表《基于 SVD 与 Fast Kurtogram 算法的滚动轴承故障诊断》(《振动与冲击》2014 年第 33 卷第 10 期)等论文。
E-mail：headic@163.com