

# 谱聚类结合 LIF 在矿井突水水源类型识别中的应用

周孟然, 宋红萍\*, 胡 锋, 来文豪, 王锦国

安徽理工大学电气与信息工程学院, 安徽 淮南 232001

**摘 要** 突水事故威胁井下人员的生命安全和造成财产损失, 因此准确检测出突水水源类型具有重大意义。使用水化学分析法检测水源类型耗时长、过程复杂。激光诱导荧光(LIF)技术具有快速、灵敏、干扰小等优点, 将 LIF 技术结合智能算法建立突水水源识别模型可以准确检测出突水水源的类型。目前这类模型一般需要对荧光光谱进行去噪、降维、波段选取等处理, 过程繁琐, 并且模型都是在均匀分组的突水水源荧光光谱上建立的, 并没有讨论不均匀分组对模型的影响, 也没有针对不均匀分组建立模型。在实际工程应用中, 采集的样本数量是有很大概率呈现不均匀的, 因此本文提出一种飞蛾扑火(MFO)算法结合谱聚类(SC)的方法实现对不均匀分组的突水水源荧光光谱的识别。实验中, 首先从淮南煤矿获取 5 种实验水样, 使用激光诱导荧光实验设备采集所有水样的荧光光谱, 五种水样的组数分别为 75, 80, 80, 30 和 135。其次, 建立 MFO-SC 水样识别模型, 通过对比后标签映射方式选择 K-Means、相似矩阵的计算方式选择高斯核函数和划分准则选择  $ncut$ , 用 MFO 对高斯核函数的参数寻优得到  $\sigma$  的值为 1.745 并且固定模型的初始聚类中心。随后, 分别建立 K-Means, SVM 和 MFO-SVM3 种水样识别模型。对比 MFO-SC 模型与 K-Means 模型, 得到 MFO-SC 模型的最优准确率为 100% 且平均准确率也为 100%, K-Means 模型的最优准确率为 99.75%, 而平均准确率为 79.57%; 再分别计算 SVM 模型和 MFO-SVM 模型的训练集准确率和测试集准确率, SVM 模型训练集准确率为 80%, 测试集准确率为 80%; MFO-SVM 模型训练集准确率为 100%, 测试集准确率为 95.625%。最后, 使用 4 种模型对其他三个不均匀分组的突水水源荧光光谱进行识别, 研究结果表明将 MFO-SC 算法用于突水水源类型的识别上是有效的, 可以准确地检测出突水水源的类型, 对煤矿生产安全有重要意义。

**关键词** 激光诱导荧光; 矿井突水水源; 谱聚类; MFO

**中图分类号:** O657.3    **文献标识码:** A    **DOI:** 10.3964/j.issn.1000-0593(2021)02-0435-06

## 引 言

随着煤矿开采深度的增加, 矿井突水问题越来越严重<sup>[1]</sup>, 一旦发生矿井突水事件, 带来的不只是经济损失, 更严重的是造成井下人员伤亡。因此, 建立突水预警模型和水样识别模型已经成为治理突水的关键。目前突水水源类型识别方法主要利用水化学分析法<sup>[2-3]</sup>和智能算法<sup>[4-5]</sup>等。利用水化学方法来识别突水水源类型, 需要检测 pH 值和电导率等, 获取这些变量通常需要较长的时间, 虽然可以比较准确地检测出矿井突水的水源类型, 但是不适合在现场建立水样识别模型。激光诱导荧光技术具有分析精度高、速度快等特

点, 在化工、医学、环境等领域有很多的应用<sup>[6-8]</sup>。近些年来, 也有许多人将激光诱导荧光技术应用于突水水源类型的识别, 具有很好的应用效果。如: 文献<sup>[9]</sup>使用激光诱导荧光技术、间隔偏最小二乘法(interval PLS, iPLS)结合粒子群(particle swarm optimization, PSO)联合支持向量分类算法(PSO-SVC)对突水水源的类型进行识别, 先用 iPLS 对水样的荧光光谱进行波段选取, 然后使用 PSO-SVC 对选取的波段进行识别, 实现对突水水源类型的识别。此外还有使用主成分分析法<sup>[10]</sup>(principal component analysis, PCA)等算法建立水样识别模型对突水水源的荧光光谱进行识别, 然而这些模型都是在突水水样荧光光谱进行均匀分组的基础上进行识别, 并没有对不均匀分组进行讨论, 也没有针对不均匀分组

收稿日期: 2020-01-05, 修订日期: 2020-04-27

基金项目: 国家重点研发计划项目(2018YFC0604503), 国家“十二五”科技支撑计划重点项目(2013BAK06B01), 国家安全监管总局安全生产重特大事故防治关键技术科技项目(anhui-0010-2018AQ), 安徽省青年科学基金项目(1808085QE157)资助

作者简介: 周孟然, 1965 年生, 安徽理工大学电气与信息工程学院教授    e-mail: mrzhou8521@163.com

\* 通讯作者    e-mail: shp7420@163.com

建立水样识别模型。只有在理想情况下,每种样本的数量才会是均匀分组的。在实际的工程应用中,采集的每种样本数量有很大可能是不均匀的,因此本文在不均匀分组上建立模型更具有工程应用指导意义。

谱聚类<sup>[11]</sup>(spectral clustering, SC)属于无监督学习,可以实现对任意形状的数据集的聚类,计算量较小、实现简单、聚类效果好并能很快收敛于全局最优解。文献[12]是使用模糊估计谱聚类实现对癌症的检测。飞蛾扑火算法<sup>[13]</sup>(moth-flame optimization, MFO)是近几年提出来的优化算法,具有参数少、计算精度高等特点。

本文提出一种将激光诱导荧光与 MFO-SC 算法相结合建立水样识别模型,先建立 MFO-SC 水样识别模型,然后与其他模型进行性能对比,最后在其他不均匀数据上进行验证。MFO-SC 水样识别模型在不需较多的先验知识前提下,不需要对突水水源荧光光谱进行去噪、降维、波段筛选等处理,过程简单,可以直接对不均匀分组的突水水样荧光光谱进行识别,准确率高,具有很强的泛化能力。

## 1 实验部分

### 1.1 材料

根据淮南谢一矿突水水源的特点,科学的选取了危害性较大的老空水、砂岩水以及按一定体积比的老空水和砂岩水的混合水作为实验的研究对象,并对水样进行密封和遮光处理,带回实验室。以 2019 年 7 月 7 号在淮南某矿区采集到的老空水、砂岩水并按一定体积比混合后得到 5 种实验样本,依次为老空水与砂岩水体积比为 2:1 的混合水(以下简称“混合水 1”)、老空水与砂岩水体积比为 1:1 的混合水(以下简称“混合水 2”)、老空水与砂岩水体积比为 1:2 的混合水(以下简称“混合水 3”)、老空水和砂岩水。

### 1.2 仪器

实验所选用的光谱仪是 USB2000+ 光谱仪,光谱仪检测荧光光谱的波长范围为 340~1 021 nm,激光由 405 nm 激光器提供。激光通过石英光纤由 FPB-405-V3 型荧光探头垂直浸入实验水样中,激光诱导水样发出的荧光由荧光探头接收传送给光谱仪。实验采集荧光光谱的实验示意图如图 1 所示。

### 1.3 算法原理

#### 1.3.1 SC 算法

谱聚类(SC)是一种无监督学习算法,不需要样本的标签信息、不需要对样本进行去噪、降维、波段选取等处理且能在任意形状的样本空间上聚类且收敛于全局最优解。谱聚类将聚类转化为图的最优划分,主要思想是将样本看作图中的点,用边将点连接组成图,计算两点之间的边权重,最后对图进行分割,分割后不同子图间的边权重和尽可能的低,子图内的边权重和尽可能的高,从而完成聚类。

谱聚类的算法流程:

输入:样本集  $D=(X_1, X_2, \dots, X_n)$ ,相似矩阵的计算方式,划分准则,标签映射的方法,类别数  $K$ 。

过程:

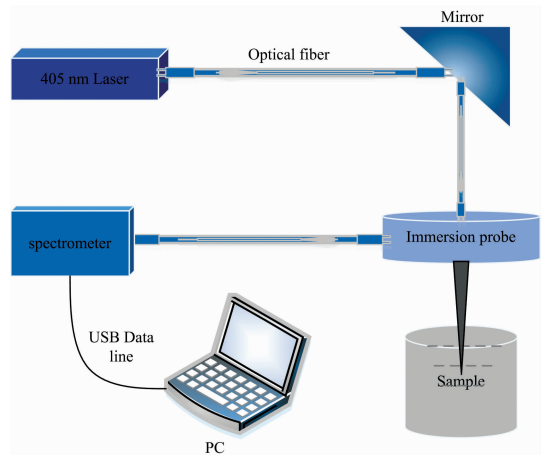


图 1 激光诱导荧光实验示意图

Fig. 1 Schematic diagram of laser induced fluorescence experiments

Step 1: 根据输入的相似矩阵的计算方式构建样本的相似矩阵  $S$ 。

Step 2: 根据相似矩阵  $S$  构建邻接矩阵  $W$ , 构建度矩阵  $D$ 。

Step 3: 构建拉普拉斯矩阵  $L$ , 根据输入的划分准则对  $L$  进行规范化得到矩阵  $L'$ 。

Step 4: 计算矩阵  $L'$  最大的  $K$  个特征值各自对应的特征向量  $f$ 。

Step 5: 将各自对应的特征向量  $f$  组成矩阵  $F$ ,  $F$  按行标准化得到特征矩阵  $F'$ 。

Step 6: 将  $F'$  的每一行作为一个新样本,用输入的聚类方法进行聚类,聚类的类别数为  $K$ 。

Step 7: 得到簇划分  $C(C_1, C_2, \dots, C_k)$ 。

输出: 簇划分  $C(C_1, C_2, \dots, C_k)$ 。

#### 1.3.2 MFO 算法

飞蛾扑火算法(MFO)是一种元启发式优化算法,它的灵感来源于自然界中飞蛾的横向往导方法。根据飞蛾与火焰保持固定的飞行角度的行为建立合理的数学模型。每个飞蛾都是潜在的最优值,飞蛾在解空间中的位置是所求问题的解,飞蛾围绕火焰寻优并通过迭代的方式更新位置,直至得到问题解的最优值。使用 SC 算法对矿井突水水源荧光光谱数据建立模型后,使用 MFO 算法对荧光光谱数据进行寻优的实验框图如图 2 所示。

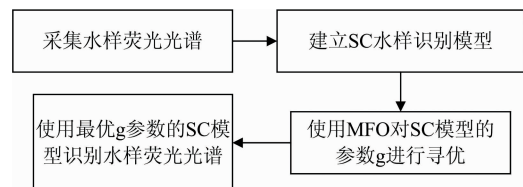


图 2 MFO 优化 SC 的实验框图

Fig. 2 Experimental block diagram of MFO optimized SC

## 2 结果与讨论

### 2.1 采集水样原始荧光光谱

利用激光诱导荧光实验设备采集实验水样的荧光光谱, 实验在遮光、同等温度和湿度的环境下进行, 5 种水样采集的组数各不相同, 混合水 1 采集 75 组、混合水 2 采集 80 组、混合水 3 采集 80 组、老空水采集 30 组以及砂岩水采集 135 组。

图 3 为 400 组的实验水样的原始荧光光谱图, 1—75 为混合水 1、76—155 为混合水 2、156—235 为混合水 3、236—265 为老空水、266—400 为砂岩水。从图 3 上可以很明显的看到不同实验水样的荧光光谱有所不同, 差异主要集中在 400~650 nm 之间, 并且老空水荧光光谱与其他实验水样的差异比较明显, 混合水 1、混合水 2、混合水 3 和砂岩水荧光光谱的差异相对较小。因此, 很难通过观察去区分实验水样的类别, 需要研究实验水样的荧光光谱, 利用智能算法建立水样识别模型对实验水样进行识别。

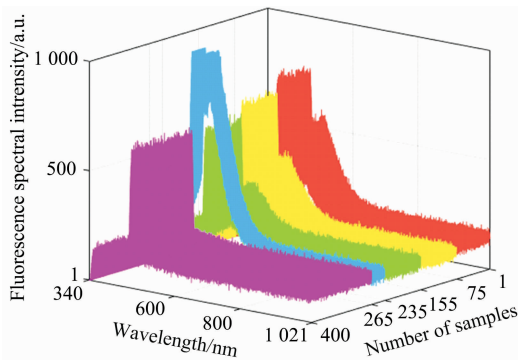


图 3 原始荧光光谱图  
Fig. 3 Original fluorescence spectra

### 2.2 MFO-SC 水样识别模型的建立

本文在建立模型时, 通过大量实验发现标签映射更能影响 SC 算法的性能, 因此建立模型时需要先确定标签映射的方式, 最后确定相似矩阵的计算方式和划分准则。

由于 K-Means 具有原理简单、实现方便和收敛速度快等优点, 所以选用 K-Means 作为标签映射的方法。实验对比了常见的 6 种核函数作为谱聚类的相似矩阵计算方式的性能最终选择高斯核函数作为相似矩阵的计算方式, 模型的划分准则选择 ncut。建立 MFO-SC 模型时高斯核函数的参数  $\sigma$  设为 5, 使用 MFO-SC 模型对实验水样的荧光光谱进行识别, 图 4 是实验 100 次模型准确率的变化图。准确率在 47.75% 到 96.25% 之间来回波动, 虽然均值为 86.29%, 但是波动的幅度较大。原因是模型在进行识别时, 随机选择初始聚类中心导致模型最终的聚类中心不准确从而影响聚类的效果, 因此模型需要固定初始聚类中心, 选择准确率为 96.25% 的聚类中心作为模型的初始聚类中心, 初始聚类中心如表 1 所示。

高斯核函数的参数  $\sigma$  对 MFO-SC 模型的准确率息息相关, 使用 MFO 寻优算法对高斯核函数的参数  $\sigma$  进行寻优。

表 1 MFO-SC 模型初始聚类中心

Table 1 Initial clustering centers of MFO-SC model

	维度 1	维度 2	维度 3	维度 4	维度 5
初始聚类中心 1	-0.008 7	0.113 1	0.485 5	0.850 1	0.169 3
初始聚类中心 2	-0.090 2	-0.600 6	-0.574 7	0.323 1	0.432 1
初始聚类中心 3	0.661 4	0.485 7	-0.292 6	0.017 9	0.447 4
初始聚类中心 4	-0.584 8	0.506 9	-0.139 0	-0.100 7	0.525 5
初始聚类中心 5	0.135 4	-0.371 9	0.578 7	-0.397 2	0.586 8

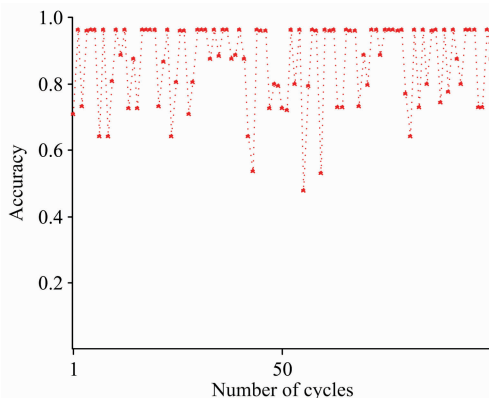


图 4 未固定初始聚类中心的 MFO-SC 模型准确率的变化图  
Fig. 4 Change diagram of accuracy of MFO-SC model without fixed initial clustering center

MFO 进行参数寻优时, MFO 的飞蛾数设为 10, 火焰数设为 10, 选择合理的寻优范围, 最大迭代次数设为 50 次。得到高斯核函数的参数  $\sigma$  最优值为 1.745, 并且模型的准确率达到 100%。

### 2.3 MFO-SC 水样识别模型与之比较

#### 2.3.1 MFO-SC 水样识别模型与 K-Means 水样识别模型对比

使用 MATLAB 软件建立 K-Means 水样识别模型。针对实验水样, 使用 K-Means 水样识别模型和 MFO-SC 水样识别模型对其进行识别, 实验 100 次。两个模型平均准确率、最优准确率和准确率方差如表 2 所示。

从表 2 我们可以看到 K-Means 水样识别模型的最优准确率可以达到 99.75%, 但是平均准确率只有 79.57%, 而 MFO-SC 水样识别模型的最优准确率可以达到 100%, 而且平均准确率也达到了 100%, 通过比较可以看出 MFO-SC 水样识别模型稳定性更好, 识别率更高。

表 2 K-Means 模型和 MFO-SC 模型性能比较

Table 2 Performance comparison of K-Means model and MFO-SC model

识别模型	平均准确率/%	最优准确率/%	准确率方差
K-Means	79.57	99.75	0.028 1
MFO-SC	100	100	0

#### 2.3.2 MFO-SC 水样识别模型与监督学习水样识别模型对比

针对实验水样建立了两种基于监督学习算法的水样识别

模型, 分别是 SVM 水样识别模型和 MFO-SVM 水样识别模型, 在建立模型之前需要合理的将实验水样划分成训练集和测试集, 经过多次实验, 最终将 5 种共 400 组实验水样按 3:2 的比例划分为训练集和测试集, 训练集和测试集的划分结果如表 3 所示。

表 3 实验水样数据集的划分  
Table 3 Classification of experimental water sample data sets

水样类别	训练集(240)	测试集(160)
混合水 1(75)	45	30
混合水 2(80)	48	32
混合水 3(80)	48	32
老空水(30)	18	12
砂岩水(135)	81	54

先使用 MATLAB 软件的 libsvm 工具箱建立 SVM 水样识别模型, SVM 的惩罚系数和核函数参数  $g$  使用默认值。然后使用 libsvm 工具箱建立 MFO-SVM 水样识别模型, 使用 MFO 寻优算法对 SVM 的  $c$  和  $g$  参数进行寻优, 飞蛾数设为 10, 火焰数设为 10, 最大迭代次数设为 100 次,  $c$  参数的范围设为  $[0.01, 50]$ ,  $g$  的参数范围设为  $[0.01, 50]$ , 得到参数的最优值  $c=0.3742$ ,  $g=0.01$ 。最后使用 SVM 水样识别模型和 MFO-SVM 水样识别模型对实验水样的荧光光谱进行识别, 两种模型的训练集准确率和测试集准确率如表 4 所示。

表 4 SVM 模型和 MFO-SVM 模型的性能比较  
Table 4 Performance comparison of SVM model and MFO-SVM model

识别模型	SVM	MFO-SVM
$c$	1	0.3742
$g$	1/2048	0.01
训练集准确率	80%	100%
测试集准确率	80%	95.625%

从表 4 我们可以看到使用 SVM 水样识别模型对实验水样的荧光光谱进行识别, 训练集的准确率为 80%, 训练集的准确率较低说明模型的参数选择不合理。SVM 水样识别模型的测试集准确率也为 80%, 如图 5 所示模型之所以将不属于砂岩水的实验水样误识别为砂岩水, 是因为 400 组实验水样中砂岩水有 135 组, 训练模型时训练集的 240 组实验水样砂岩水有 81 组, 不均衡分组导致训练好的模型在识别时更加倾向把实验水样分到训练集中实验水样组数更多的类型, 导致测试集的准确率不理想, 建立的模型不合理。而使用 MFO-SVM 水样识别模型对实验水样的荧光光谱进行识别时, 使用 MFO 对参数进行寻优, 使得训练集的准确率达到 100% 同时测试集的准确率也达到了 98.75%。SVM 水样识别模型和 MFO-SVM 水样识别模型在对不均匀分组的实验水样进行识别时都需要合理的划分测试集和训练集, 过程复

杂, 需要较多的先验知识并且两种模型的准确率都小于 MFO-SC 水样识别模型的准确率。

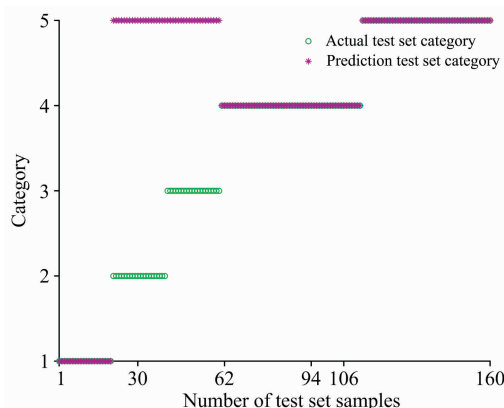


图 5 SVM 水样识别模型测试集的分类结果  
Fig. 5 Classification results of SVM model test set

2.3.3 在其他不均匀分组实验水样上的验证

合理的对采集好的 5 种水样进行分组, 并且保持实验水样的总组数为 400 组, 得到 3 组不同的实验水样即水样 1、水样 2 和水样 3。使用建立好的四种水样识别模型分别对水样 1、水样 2 和水样 3 进行识别, 实验 100 次, 其平均准确率如图 6 所示, 使用 SVM 和 MFO-SVM 建立的水样识别模型的平均准确率是指测试集与训练集分类正确的实验样本与总实验样本数的比值。

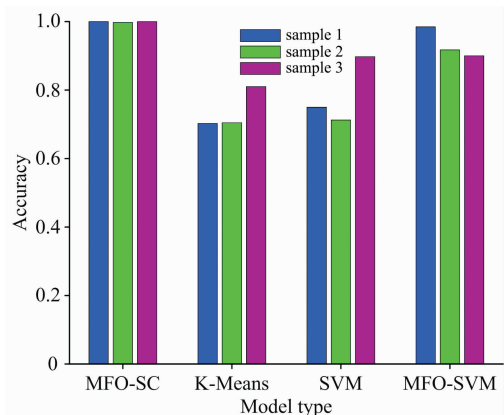


图 6 四种模型平均准确率对比图  
Fig. 6 Comparison chart of average accuracy of four models

从图 6 我们可以看到在 3 组水样数据集上 MFO-SC 水样识别模型的平均准确率都是最高的, K-Means 水样识别模型的平均准确率都是最低的, 并且在水样 1 上 MFO-SVM 水样识别模型的平均准确率高于 SVM 水样识别模型而在水样 3 上 SVM 水样识别模型的平均准确率要高于 MFO-SVM 水样识别模型。针对 3 组水样, 4 种模型的平均准确率都出现波动, MFO-SC 水样识别模型的波动最小, MFO-SVM 水样识别模型次之, K-Means 水样识别模型和 SVM 水样识别模型的平均准确率波动最大, 说明 K-Means 水样识别模型、SVM 水样识别模型和 MFO-SVM 水样识别模型的泛化能力

比 MFO-SC 水样识别模型较弱, MFO-SC 水样识别模型的稳定性更好。

### 3 结 论

根据淮南煤矿突水水源的特点, 选取老空水、砂岩水以及按一定体积比的老空水和砂岩水的混合水作为实验的研究对象, 先选取 MFO-SC 算法针对不均匀分组的水样荧光光谱建立水样识别模型, 然后将 MFO-SC 水样识别模型与另外三种水样识别模型进行了比较。通过实验可以发现: 第一, 建立合理的 MFO-SC 水样识别模型可以很好的识别出不均匀分组的水样荧光光谱, 并且识别率可以达到 100%; 第二, MFO-SC 和 K-Means 都属于无监督学习算法, 都不需要划分数据集, 与 K-Means 水样识别模型相比, 使用 MFO-SC 水样

识别模型的稳定性更好, 准确率更高。使用监督学习算法 SVM 和 MFO-SVM 建立水样识别模型相对于 MFO-SC 水样识别模型来说需要合理划分数据集, 过程复杂, 并且识别的准确率都要低于 MFO-SC 水样识别模型的准确率; 第三, 通过验证, 使用 MFO-SC 建立水样识别模型对其他不均匀分组的水样荧光光谱的识别率均达到 99% 以上, 明显高于其他三种水样识别模型, 说明 MFO-SC 水样识别模型具有更好的泛化能力。实验证明了使用 MFO-SC 算法建立水样识别模型具有可行性, 对识别矿井突水水源以及矿井安全生产有重大意义。本文采用的 MFO-SC 水样识别模型不仅可以用于老空水、砂岩水以及按一定体积比的老空水和砂岩水的混合水的水样识别, 也可以用于其他突水水源荧光光谱的识别, 同时也为激光诱导荧光技术在其他领域上的应用提供了一种简单、有效的方法。

### References

- [ 1 ] FENG Zhi-dong, WANG Tao, GU Qing-hua, et al(冯治东, 王 桃, 顾清华, 等). System Engineering Theory and Practice(系统工程理论与实践), 2017, 37(12): 3289.
- [ 2 ] XUE Liang, YAO Duo-xi, LU Hai-feng, et al(薛 凉, 姚多喜, 鲁海峰, 等). Coal Technology(煤炭技术), 2017, 36(11): 120.
- [ 3 ] LIU Jian-min, WANG Ji-ren, LIU Yin-peng, et al(刘剑民, 王继仁, 刘银朋, 等). Journal of Safety and Environment(安全与环境学报), 2015, 15(1): 31.
- [ 4 ] XU Xing, TIAN Kun-yun, LI Feng-qin, et al(徐 星, 田坤云, 李凤琴, 等). Journal of Southwest University • Natural Science Edition (西南大学学报 • 自然科学版), 2018, 40(4): 170.
- [ 5 ] WANG Yan-bin, SUN Shao-guang, ZHANG Zi-jian(王彦彬, 孙韶光, 张子建). Mathematics in Practice and Theory(数学的实践与认识), 2017, 47(21): 172.
- [ 6 ] WU Wen-fei, LIAO Qing-ling(吴文飞, 廖庆玲). Laser Journal(激光杂志), 2019, 40(10): 36.
- [ 7 ] ZHANG Ling-ling, WEN Long, CHEN Yuan, et al(张玲玲, 文 龙, 陈 媛, 等). Applied Laser(应用激光), 2019, 39(4): 703.
- [ 8 ] ZHANG Wen-xue, ZHANG Xiao-rong, QIN Cheng-bing, et al(张文学, 张晓荣, 秦成兵, 等). Laser & Optoelectronics Progress(激光与光电子学进展), 2019, 56(20): 147.
- [ 9 ] ZHOU Meng-ran, BIAN Kai, HU Feng, et al(周孟然, 卞 凯, 胡 锋, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(7): 2196.
- [10] JU Qi-ding, HU You-biao, ZHANG Shu-ying(琚棋定, 胡友彪, 张淑莹). Coal Engineering(煤炭工程), 2018, 50(12): 90.
- [11] Mashaan Alshammari, Masahiro Takatsuka. Pattern Recognition Letters, 2019, 128: 155.
- [12] Thenmozhi K, Karthikeyani Visalakshi N, Shanthi S. Asian Pacific Journal of Cancer Prevention; APJCP, 2018, 19(7): 1935.
- [13] Seyedali Mirjalili. Knowledge-Based Systems, 2015, 89: 228.

# Application of Spectral Clustering and LIF in Recognition of Mine Water Inrush Source Types

ZHOU Meng-ran, SONG Hong-ping\*, HU Feng, LAI Wen-hao, WANG Jin-guo

College of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China

**Abstract** Water inrush accidents threaten the lives of people and cause property damage. Therefore, it has great significance in accurately detecting the type of water inrush. Hydrochemical analysis method takes a long time and has a complicated process to detect the type of water inrush. Laser-induced fluorescence (LIF) technique has the advantages of fastness, high sensitivity, and low interference. Water inrush source recognition model building with LIF technique and intelligent algorithms can accurately detect the type of water inrush. At present, such models generally require de-noising, dimension reduction, and band selection on the fluorescence spectra, and this process is complicated. The models are built on the fluorescence spectra of the water inrush source which is evenly grouped. The influence of the uneven grouping on the model is not discussed, and the model is not built for the uneven grouping. In practical engineering applications, the number of samples collected is highly likely to be uneven, so Moth-flame optimization (MFO) algorithm combined with spectral clustering (SC) is proposed to realize the uneven grouping of water inrush fluorescence spectrain this paper. In the experiment, firstly, five kinds of experimental water samples were obtained from Huainan coal mine. Laser-induced fluorescence experimental equipment was used to collect fluorescence spectra of all water samples. The number of groups of five water samples is 75, 80, 80, 30 and 135. Secondly, build MFO-SC water sample recognition model. After comparison, K-Means is selected for the label mapping method, the Gaussian kernel function is selected for the calculation method of the similarity matrix, and the  $n_{cut}$  is selected for the partition criterion. The parameters of the Gaussian kernel function were optimized by using MFO to obtain the parameter value of 1.745, and the initial clustering center of the model was fixed. Subsequently, build three water sample recognition models of K-Means, SVM and MFO-SVM, respectively. Comparing the MFO-SC model with the K-Means model, the optimal accuracy of the MFO-SC model is 100%, and the average accuracy is 100%. The optimal accuracy of the K-Means model is 99.75%, and the average accuracy is 79.57%. Then calculate the training set accuracy and test set accuracy of the SVM model and MFO-SVM model respectively. The accuracy of the training set of the SVM model is 80%, and the accuracy of the test set is 80%; the accuracy of the training set of the MFO-SVM model is 100%, and the accuracy of the test set is 95.625%. Finally, four models were used to identify water inrush fluorescence spectra of the other three uneven groups. The research results show that the MFO-SC algorithm is effective in identifying the type of water inrush, and can accurately detect the type of water inrush, which has great significance on the safety of coal mine production.

**Keywords** Laser induced fluorescence; Mine water inrush source; Spectral clustering; MFO

(Received Jan. 5, 2020; accepted Apr. 27, 2020)

\* Corresponding author