

2023年第1卷第1期 1

XRF 结合化学计量学对纸张快递文件袋的分类研究(二号黑体)

周贯旭 1 姜红 2，3\* 周飞翔 1 满吉 4(四号仿宋)

(1. 中国人民公安大学 侦查学院 北京 100038；

2. 甘肃警察职业学院 刑事侦查系 甘肃 兰州 7300463；

3. 食品药品安全防范山西省重点实验室 山西 太原 030006；

4. 北京华仪宏盛技术有限公司 北京 100024)

摘要(小五，黑体)： 本文旨在建立一种快速无损的识别纸质快递文件袋的分析方法。利用手持式 X 射线荧光光谱仪对 50 个纸质 快递文件袋样品进行检验，通过分析样品的 X 射线光谱数据的元素种类和含量对样本进行分类。对归一化后的光谱 数据分别建立贝叶斯判别、RBF 神经网络和 K- 近邻三种分类模型，对分类结果进行分析验证。50 个样品可被分为 四类，贝叶斯判别模型、RBF 神经网络和 K- 近邻分类模型的分类准确率分别为 96% 、98% 和 100%。通过对分类 准确率的比较，K- 近邻分类模型更适合此类样本的 XRF 光谱数据，XRF 与化学计量学相结合可以对纸质快递文件 袋实现有效的识别。该方法简单快速且无损样品，可为快递文件袋类物证鉴定提供科学依据。

关键词(小五，黑体)： X 射线荧光光谱；纸质快递文件袋；化学计量学；贝叶斯判别； K- 近邻算法(小五，宋体)

中图分类号(小五，黑体)： D918.9 ； O657.33 (小五，Times New Roman)； TB99 文献标识码： A

**Research** **on** **the** **Classification** **of** **Paper** **Express** **Document** **Bags** **by** **XRF** **and** **Chemometrics**(小三，加粗)

Zhou Guanxu1 Jiang Hong2,3\* Zhou Feixiang1 Man Ji4(小四，姓大写)

(1. Collage of Investigation ，People's Public Security University of China ， Beijing 100038，

China;

2. Criminal Investigation Department，Gansu Police Vocational College ，Gansu Lanzhou 730046 ，China;

3. Shanxi Key Laboratory of Food and Drug Safety Prevention and Control ，Shanxi Taiyuan 030006 ，China;

4. Beijing Huayi Hongsheng Technology Co. ，Ltd. ，Beijing 100024 ，China)(五号)

**Abstract**t(小五，加粗)**:** To establish a fast and non-destructive analysis method for identifying paper express document bags . 50 samples of paper delivery document bags were inspected by handheld X-ray fluorescence, and the samples were classified by analyzing the element types and contents of the X-ray spectrum data of the samples . Establish Bayesian discriminant analysis, RBF neural network, and K-nearest neighbor classification models for normalized spectral data, and analyze and verify the classification results. The 50 samples can be divided into four categories. The classification accuracy of Bayesian Discriminative model, RBF neural network and K-nearest neighbor classification model is 96%, 98% and 100% respectively. By comparing the classification accuracy, the K-nearest neighbor classification model is more suitable for XRF spectral data of such samples . The combination of XRF and Chemometrics can effectively identify paper express document bags. The combination of XRF and Chemometrics can effectively identify paper express document bags. This method is simple, fast, and non-destructive for samples. It can provide scientific basis for the identification of physical evidence in express delivery document bags.(小五)

**Keywords(小五，加粗):** X-ray fluorescence spectroscopy; Paper express delivery document bag; chemometrics; Bayesian discrimination; K-nearest neighbor algorithm

基金项目： 食品药品安全防范山西省重点实验室开放课题资助(202204010931006)（六号宋体，此处为脚注，和正文分开）

第一作者：周贯旭(2000 年 - )，男，在读硕士生，研究方向：物证检验， E-mail：jiangh2001@163.com。

通讯作者：姜红(1963 年 - )，女，教授，硕士生导师，研究方向：微量物证分析， E-mail：jiangh2001@163.com。



2

实验与分析

SHIYAN YU FENXI

2023年第1卷第1期



(正文：五号，宋体)在社会的不断发展中，快递业务涵盖社会的各个 方面。在不同的案件类型中，经常能够提取到快递文 件袋类物证，对该类物证进行检验分析，对其进行快 速分类并区分，有助于对该物证来源等其他信息进行 判断，为侦查破案提供线索，缩小侦查范围 [1]。

目前，用于检验纸张物证的方法主要有红外光 谱法、拉曼光谱法、扫描电镜 / 能谱法、X 射线荧

光光谱法、原子发射光谱法等 [2-5] 。红外光谱和拉曼 光谱检验过程简单快速，且无损样品，此前笔者利 用红外光谱结合主成分分析对纸质快递文件袋进行 了检验，实现了快递文件袋的初步分类。陈维娜 [6] 等利用显微共聚焦拉曼光谱结合化学计量学对静电 复印纸类物证进行了鉴别分类。但利用红外光谱和 拉曼光谱检测样品中含量较少的成分时，其吸收特 征峰可能会被含量高的、吸收峰相近的有机成分所 掩盖， X 射线荧光光谱针对样品中的元素进行测定， 可以有效弥补上述两种光谱法的缺点。姜红 [7] 等利 用 X 射线荧光光谱结合种属差异性对塑钢窗进行分 类研究，可以将塑钢窗样本进行完全的分类。孙家 政 [8] 等利用 X 射线荧光光谱结合 Fisher 判别函数对 牛皮纸类物证进行了分类研究，并取得了较好的分 类效果，表明利用 X 射线荧光光谱对纸质类物证进 行分类研究可以取得较好的效果。利用 XRF 结合化 学计量学对纸质快递文件袋进行分类研究的相关研 究尚未见报道。所以本文利用手持式 X 射线荧光光 谱仪对来自不同快递公司的 50 个纸质快递文件袋样 品进行测试，依据元素含量对样品进行分类，依据 贝叶斯判别、RBF 神经网络和 K- 近邻三种算法，分 别建立样品分类模型，进而优选出最适合此类样品 的分类算法。

1 实 验（4号，宋体）

1.1 实验仪器及条件(二级标题：五号，黑体)

X-MET8000Expert 型 X 射线荧光光谱仪，英国

牛津公司，铑(Rh)阳极靶，测试电压为 50 kV，

电流设置为 200 μA，工作温度为 -25℃~ 50℃。

1.2 实验样品

不同品牌的纸质快递文件袋样品50 个(样品表略)。

1.3 实验方法

分别剪取约 0.5 cm ×0.5 cm 大小的样品，利用 X 射线荧光光谱仪在上述条件下以及最优测量时间 下进行检测，对同一样品进行三次测试后取平均值 作为该样品 XRF 光谱数据。

优选测量时间实验：随机选取 1# 样品，利用上 述仪器在不同的测试时间下分别进行测试，比较在 不同的时间下，仪器所测元素含量差异，测量时间 分别为： 40s、50s 、60s 、70s、80s。

重现性实验：随机选取 3# 样品，利用 X 射线光 谱仪对同一部位进行 6 次平行实验。

2 结果与讨论

2.1 优选时间

表 1 为每种元素在不同测量时间所测的元素含 量，通过比较发现，在 40s 到 50s 中，多种元素含 量均有较大的上升，在 50s 至 60s 时元素含量相对 较为稳定，故确定利用 X 射线荧光光谱仪对纸质快 递文件袋的最佳测量时间为 60s。

表 1 不同测试时间下元素含量 (mg/kg)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 测量时间(s ) | Ca | Fe | Ti | Zn | Sn |
| 40 | 93139 | 1398 | 152 | 17 | 85 |
| 50 | 98681 | 1504 | 170 | 22 | 87 |
| 60 | 99514 | 1573 | 188 | 21 | 88 |
| 70 | 100590 | 1616 | 190 | 26 | 86 |
| 80 | 101985 | 1632 | 189 | 25 | 89 |

2.2 重现性实验结果分析

为验证仪器的稳定性，确保实验结果准确可靠， 对 3# 样品进行了 6 次重现性实验并计算每种元素的



(小五，宋体)表 2 相对标准偏差值 ( 元素含量 mg/kg)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 测试次数 | Ca | Fe | Ti | Zn | Sn |
| 1 | 94630 | 1701 | 155 | 19 | 95 |
| 2 | 107447 | 1621 | 152 | 21 | 87 |
| 3 | 105972 | 1609 | 155 | 18 | 80 |
| 4 | 111210 | 1692 | 142 | 23 | 85 |
| 5 | 104341 | 1470 | 135 | 21 | 75 |
| 6 | 104722 | 1538 | 156 | 19 | 78 |
| 平均值 | 104720.333 | 1605.167 | 149.167 | 20.167 | 83.333 |
| 标准偏差 | 5533.098 | 89.186 | 8.658 | 1.834 | 7.229 |
| RSD 值/% | 5.28 | 5.54 | 5.77 | 9.09 | 8.6 |



2023年第1卷第1期 周贯旭等： XRF结合化学计量学对纸张快递文件袋的分类研究 3



相对标准偏差。元素含量的相对标准偏差(RSD) 值如表 2 所示。实验结果显示， RSD 值均低于 10%， 表明样品重现性良好、仪器精密性较好。因此，可 以利用该实验方法对纸质快递文件袋进行测试 [9]。

2.3 人工分类

纸质快递文件袋的原料主要是纤维素和半纤维 素，因为快递文件袋表面一般是以白色为主，在生 产过程中，生产厂家常常会添加一定量的涂漆填料。 因此不同样品的 XRF 光谱数据的元素种类和含量会 有所不同， 50 个样品的 XRF 光谱图如图 1 所示。

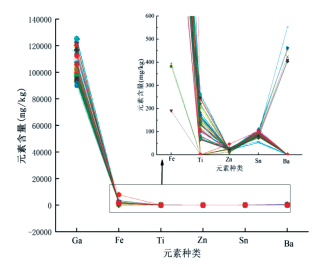


图 1 样品的 XRF 光谱图(小五，宋体)(图中数值带数据清晰，数值带上方要有量名称及单位)

通过图 1 可以看出，所有样品都具有 Ca 、Fe、 Zn 和 Sn 元素，部分样品中含有 Ti 和 Ba 元素中的 一种或者两种。其中，在所有的元素中，元素含量 最多的是 Ca 元素， Ca 元素主要来源于常见的填料 碳酸钙，碳酸钙价格低廉，在填料中增强了涂层耐 磨损和抗腐蚀的能力，在降低成本的同时，可有效 提高产品质量，所以在造纸业中会选择碳酸钙作为 主要的纸张填料；剩下元素的元素含量相近，均在 1000mg/kg 以下，其中含量较多的是 Fe 元素， Fe 元 素主要来自造纸中常用的复盐、加工设备 [10] ；接下 来的元素含量较多的是 Ti 元素， Ti 元素来自纸张中 的二氧化钛填料，二氧化钛可以提高纸张的白度以 及不透明度，钛白粉的耐候性可以提高纸张的耐热 性 [11] ，通过图 1 可以看出并不是所有的样本均含有 Ti 元素，因为二氧化钛的价格较高，性价比较低， 所以在纸质快递文件袋样本中 Ti 元素含量较少并且 使用不太普遍。Zn 元素来源于纸张生产过程中所使 用的漂白剂如硫酸锌；Sn 元素来自纸张中的阻燃剂 和防腐剂；通过光谱图能够看出，在元素含量方面 Ba 元素和 Ti 元素较类似，但和 Ti 元素不同的是 Ba

元素在纸质快递文件袋中使用较少，较多样本的 Ba 元素含量为 0，而 Ti 元素的元素含量为 0 的只有较 少的样本。Ba 元素来自硫酸钡填料，能增强纸张的 抗老化性能和耐候性，纸张不易老化变脆 [12] 。在造 纸业中硫酸钡属于小众化的纸张填料，所以使用并 不普遍 [13] 。本文现依据样品中是否有 Ti 和 Ba 元素 对样品进行人工分类，可将样品分为四类，分类结 果如图 2 所示。

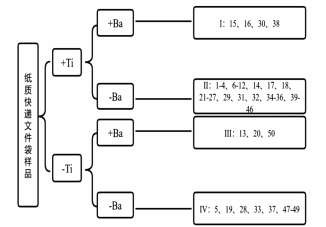


图 2 样品分类结果(图中六号字)

图中“+”表示含有该元素，“- ”表示不含有该 元素。

2.4 数据预处理

通过图 1 可以看出，不同元素之间含量差别较大。 碳酸钙属于造纸业中较为常见的填料，所以在纸张 中 Ca 的含量远远超出其他几种填料，为防止较大的 元素含量差距使计算权重发生变化，故在数据分析 前通过 Matlab 中的 mapminmax 函数对光谱数据进 行归一化处理 [14]。

2.5 贝叶斯判别

贝叶斯判别分析是利用多种因素对预测对象进 行客观分类的统计分析方法，其中常用的统计学分 析方法是 Fisher 判别分析，贝叶斯判别分析使用较 少。本文以 50 个样本的XRF 光谱数据作为研究对象， 通过 MATLAB 中的 ﬁt 函数和 predict 函数建立贝叶 斯分类模型，ﬁt 函数是 y=NaviveBayes.ﬁt(training， class)，其中 training 是输入函数，即经过归一化后 的 XRF 光谱数据，class 是分类标签，即根据元素 含量对 50 个样本做出的人工分类；predict 函数为 c=predict(y，test) ，即根据分类标签对样本做出预测， 并输出预测结果 [15- 16] 。其中，判别函数摘要见表 3。

贝叶斯判别提取了 3 个判别函数，其中第一个





4

实验与分析 SHIYAN YU FENXI

2023年第1卷第1期



表 3 判别函数摘要

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 函数 | 特征值 | 方差 | 累计 | 典型 |
| 百分比 /% | 百分比 /% | 相关性 |
| 1 | 173.985 | 98.3 | 98.3 | 0.997 |
| 2 | 2.842 | 1.6 | 99.9 | 0.860 |
| 3 | 0.167 | 0.1 | 100.0 | 0.378 |

函数相关性最高，其方差百分比也最高，对样本的 分类问题的解释能力也最强，三个判别函数的累积 方差百分比为 100%，可以解释所有的样本分类问题。 故以判别函数 1、2、3 建立样品分布散点图(见图 3)。

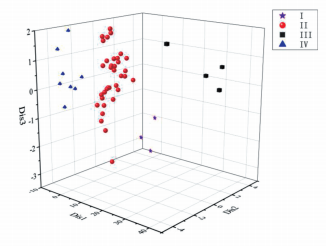


图 3 样品分布散点图

通过图 3 可以看出，四组之间分布较为分散， 同组内的样品分布较为集中。利用留一交叉法对样 品判别结果进行验证，其中第 III 类的 38# 样品距离 第 II 类较近，在最终的分类结果中，31# 样品被误 判为第 II 类；在最终的分类结果中，第 II 类中 31# 样品被误判为第 I 类，通过图 3 可以看出，31# 样品 在空间位置上更靠近第 I 类。经过交叉验证后的样 品总体分类正确率为 96%，说明人工分类方法具有

较强的科学性。

2.6 RBF 神经网络

径向基函数(Radical Basis Function ，RBF)神 经网络是一种有监督的算法模型，模型通过将数据 分为训练集和测试集进而避免过拟合。RBF 模型 包括三个部分，其中输入层中输入外部信息，然后 通过隐藏层进行信息的处理，最后通过输出层将信 息进行下一步的传递。通过 MATLAB 编辑器建立 RBF 分类模型，RBF 模型对于该类样品的分类模型 中按照 72% / 28% 的比例划分训练集与测试集，协 变量以样品的波长作为输入层神经元，选择正态化

径向基函数作为隐藏层激活函数 [17] RBF 分类模

。

型的函数主要是 net = newrbe(p\_train，t\_train，rbf\_ spread)，其 中 p\_train 为 36 训练集样本光谱数据，

t\_train 为 14 个测试集样本光谱数据，rbf\_spread 为 径向基网络扩展速度，本文设置径向基网络扩展速 度为 100 。RBF 模型分类结果见表 4，其中对于训练 集的正确率可以达到 100%，测试集的正确率则只有 92.9%，其总体准确率为 98%。测试集中只有一个 第 IV 类样本被误判为第 I 类。该算法与贝叶斯判别 相比，在对于第 II 类样本和第 I 类样本、第 III 类和 第 II 类之间的区分效果有所提高，通过划分训练集 与测试集的方法建立径向基函数神经网络分类模型， 使得样本的总体分类正确率相较于贝叶斯判别模型 有所提升。

2.7 KNN 分类模型

K- 近 邻(K-nearest neighbor) 算 法， 是 一 种 较简单的有监督的分类算法。通过寻找目标样本中 最合适的 K 个近邻样本，通过对 K 个近邻样本采 取内部投票的方式，最终确定目标样本的所属类别 [18] 。故本文尝试对 XRF 光谱数据建立 KNN 分类模 型，其中经过均一化的数据作为变量，人工分类结



表 4 RBF 模型分类结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 样本 | 实测 | I | II | 预测  III | IV | 正确百分比 |
| 训练 | I  II  III  IV  总体百分比 | 3  0  1  0  8.3% | 0  26  0  0  72.2% | 0  0  2  0  5.6% | 0  0  0  5  13.9% | 100.0%  100.0%  100.0%  100.0%  100.0% |
| 检验 | I  II  III  IV  总体百分比 | 1  0  0  1  14.3% | 0  9  0  0  64.3% | 0  0  1  0  7. 1% | 0  0  0  2  14.3% | 100.0%  100.0%  10000%  66.7%  92.9% |



试验与分析~~-~~内文~~.~~indd 4



2023/8/28 10:31:56





2023年第1卷第1期 周贯旭等： XRF结合化学计量学对纸张快递文件袋的分类研究 5



果作为分类标签，采用交叉验证的方法避免模型过 拟合。 KNN 模型分类结果见表 5 。KNN 模型正确分 类了每一类样本，总体分类准确率为 100%，比贝叶 斯判别模型和 RBF 神经网络模型的分类正确率有所 提升。

表 5 KNN 模型分类结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 实测 | I | II | 预测  III | IV |
| I | 4 | 0 | 0 | 0 |
| II | 0 | 35 | 0 | 0 |
| III | 0 | 0 | 3 | 0 |
| IV | 0 | 0 | 0 | 8 |

图 4 为 KNN 模型对 6 个变量对于分类结果的重 要性评价，可以看出重要性较高的是 Ba 元素含量和 Ti 元素含量。Ba 元素的重要性最高，在 KNN 分类 模型的分类过程中 Ba 元素的含量在样本分类正确发 挥约占 55% 的重要性，同时与本次分类任务的主要 区分点相契合， 所以通过 KNN 分类模型可以对分类 过程中重要性更高的变量进行筛选，同时也可以对 每一类的分类依据进行大致的推测。

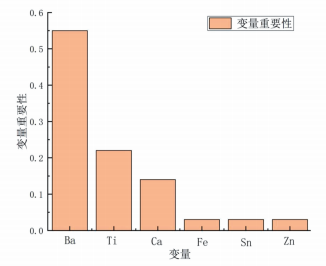


图 4 预测变量重要性

3 结语

利用 X 射线荧光光谱法对纸质快递文件袋样品进 行了测试，根据样品中是否有 Ti 元素和 Ba 元素将 50 个样品划分为四类，对经过归一化的光谱数据建立贝 叶斯判别分类模型、RBF 神经网络和 KNN 分类模型， 其样品分类准确率分别为 96% 、98% 和 100%。对于 未知样品的预测识别， K- 近邻分类模型具有更高的准 确率，并且 K- 近邻分类模型还可以筛选出重要性更 高的特征变量，可以有效确定样品的无机元素组成， 进而大致判断样品的填料成分。

参考文献(五号，黑体)

[1] 姜红 . 刑事案件现场中纸张的检验及应用 [J]. 上海造纸，

2005(03) ：9- 12.

[2] Yi Z ， X.T H ，Lindon R, et al.On the selection of the weighting

parameter value in optimizing Eucalyptus globulus pulp yield models based on NIR spectra[J]. Wood Science and Technology，2022 ，56(6):

1835-1850.

[3] Liu P ， Zhang H ， Wang S, et al. Determination of crystallinity

of Chinese handmade papers by means of X-ray diffraction[J] . Restaurator. International Journal for the Preservation of Library and

Archival Material ，2020，41(2): 69-86.

[4] 付钧泽，姜红 . 扫描电镜 / 能谱法结合多元统计学无损检验烟用

接装纸的研究 [J]. 中国造纸， 2021，40(04) ：38-45.

[5] Shinji K, Andrew C, Eugene I, et al . Detecting polystyrene

nanoplastics using filter paper-based surface-enhanced Raman spectroscopy.[J].RSC advances, 2022, 12(32): 20519-20522.

[6] 陈维娜，国中正，李开开，等 . 显微共聚焦拉曼光谱技术结合化

学计量法鉴别静电复印纸 [J]. 光谱学与光谱分析，2022，42(07)：

2033-2038.

[7] 姜红，张岚泽，刘津彤，等 . 基于种属差异性利用 XRF 对塑钢窗

的分类研究 [J]. 中国塑料， 2021 ，35(05) ：72-78.

[8] 孙家政，陆润洲，姜红，等 . XRF 结合化学计量学对牛皮纸物证

的研究 [J]. 中国测试， 2022，48(05) ：77-82.

[9] 马枭，姜红，杨佳琦 . 红外光谱结合 X 射线荧光光谱检验塑料打包

带 ( 绳 ) 的研究 [J]. 化学研究与应用， 2019，31(09) ：1643-1648. [10] 姜红，王欣，徐乐乐，等 . X 射线荧光光谱结合多元统计学检验

纸张灰烬 [J]. 激光技术， 2021，45(03) ：318-321.

[11] 张建平，张川，张千 . 装饰原纸二氧化钛应用及发展趋势 [J]. 造 纸装备及材料， 2018 ， 47(03) ：17-21.

[12] 付钧泽，姜红，刘峰，等 . 差分拉曼光谱结合系统聚类检验香烟 水松纸 [J]. 化学研究与应用， 2020 ， 32(11) ：1973- 1978.

[13] 吴平飞，邱小莹，邓鹏，等 . 常见无机非金属矿物粉体在造纸中 的应用 [J]. 中华纸业， 2023，44(Z2) ：89-93.

[14] Fengyang X, Zhongbing L, Min M, et al. LGWO-SVM geological steering identiﬁcation method for shale gas based on a gamma spectral dataset[J]. Neural Computing and Applications, 2021, 34(15): 12317-12329.

[15] 常凤，徐小华，胡忠旭 . 基于贝叶斯判别法对音乐特征信号的分 类研究 [J]. 电脑知识与技术， 2022, 18(13) ：79-81.

[16] Yong Z Z, Yang A Y, Tong T Y, et al. Bayesian discriminant analysis of yoghurt products based on Raman spectroscopy[J]. Journal of AOAC INTERNATIONAL ， 2020 ， 103(5): 1435- 1439.

[17] 田陆川，杨 俊， 姜 红 . 基于 RBFNN-FDA 模型对塑料拖鞋鞋底

的拉曼光谱研究 [J]. 上海塑料， 2022 ，50(04) ：62-67.

[18] 路敦利，宁芊，杨晓敏 . KNN- 朴素贝叶斯算法的滚动轴承故障 诊断 [J]. 计算机测量与控制， 2018 ，26(06)：21-23+27.

