

基于卡方检验的玻璃文物的成分分析与鉴别的研究

温柯轩 程新宇 陈俊宏 王 哲 陈休豪
重庆交通大学 重庆 400074

摘要: 古代玻璃易受环境的影响而产生风化, 导致内部元素比例发生变化, 从而使判别玻璃的类型成为一大难题。本文通过对已有数据利用卡方检验分析了玻璃文物表面风化与纹饰、类型和颜色的关系, 并对卡方检验进行了效应量化分析, 检验了 Cramer's V 系数, 结果显示玻璃风化与纹饰不存在相关性, 与颜色和类型存在相关性; 同时, 又对比分析了玻璃文物风化前后, 各种化学成分的变化情况, 找出了不同类型玻璃风化前后变化的主要化学成分, 从而根据这些变化相对较大的化学成分建立起一种简易的玻璃文物分类模型。

关键词: 卡方检验; 效应量化分析; Cramer's V 系数; 频率分布直方图

Research on composition analysis and identification of glass relics based on Chi-square test

Kexuan Wen, Xinyu Cheng, Junhong Chen, Zhe Wang, Xiuhao Chen
Chongqing Jiaotong University, Chongqing, 400074

Abstract: Ancient glass is susceptible to weathering due to environmental influences, which can result in changes in the proportions of internal elements, making it difficult to identify the type of glass. In this paper, the relationship between surface weathering of glass artifacts and their patterns, types, and colors was analyzed using chi-square test based on existing data. The effect size was quantitatively analyzed using Cramer's V coefficient. The results showed that there was no correlation between glass weathering and patterns, but there was correlation with color and type. Furthermore, the changes in various chemical components before and after weathering of glass artifacts were compared and analyzed, and the main chemical components that changed significantly in different types of glass before and after weathering were identified. Based on these significant changes in chemical components, a simple model for classifying glass artifacts was established.

Keywords: Chi-square test; Quantitative analysis of effect; Cramer's V coefficient; Frequency distribution histogram

一、玻璃文物表面风化与纹饰、类型和颜色的关系

在探究玻璃文物表面风化与纹饰、类型和颜色的相关性时有众多方法, 其中卡方检验、Pearson 相关系数计算、Spearman 相关系数计算以及 Kendall 的 tau-b (K) 相关系数计算, 由于所掌握的数据不符合正态分布, 故不采用 Pearson 相关系数计算来分析, 因此, 本文使用卡方检验来分析玻璃文物的表面风化与类型、纹饰、颜色的关系。

卡方检验^[1]是一种用途非常广泛的假设检验方法, 该检验是通过统计样本的实际观测值和理论推断值之间的偏离程度, 实际观测值与理论推断值之间的偏离程度就决定卡方值的大小, 如果卡方值越大, 二者偏差程度越大; 反之, 二者偏差越小; 若两个值完全相等时, 卡方值就为 0, 表明理论值完全符合。

我们假定表面风化为变量 X, 纹饰、类型、颜色为变量 Y, 利用 SPSS 软件的交互分析功能, 得到卡方检验表, 由卡方检验表可以得出: 表面风化和纹饰的显著性 P 值为 0.056*, 接受原假设, 因此不存在显著性差异;

表面风化和类型的显著性 P 值为 0.020**, 拒绝原假设, 即存在显著性差异; 表面风化和颜色的显著性 P 值为 0.507, 拒绝原假设, 即存在显著性差异。

在此基础上进行卡方检验的效应量化分析, 其包含 Phi、Cramer's V、列联系数以及 lambda 用于分析表面风化和其余三个指标的相关程度。卡方检验效应量化分析表如表 1 所示:

表 1 卡方检验效应量化分析

字段名 / 分析项	Phi	Cramer's V	列联系数	lambda
纹饰	0.326	0.326	0.310	0.000
类型	0.316	0.316	0.302	0.000
颜色	0.341	0.341	0.323	0.000

由于输入数据类型为 m*n 的形式, 故我们采用 Cramer's V 值表示卡方检验的效应量。从结果可以看

出纹饰、类型、颜色的 Cramer's V 值为 0.326、0.316、0.314, 与表面风化的差异程度为中等程度差异。

二、玻璃风化前后化学成分含量变化分析

首先, 我们已知数据的基础上对高钾玻璃和铅钡玻璃等不同类型的玻璃风化前后化学含量进行统计, 根据统计结果, 我们挑选出高钾玻璃和铅钡玻璃风化前后的二氧化硅 (SiO₂) 含量、氧化铝 (Al₂O₃) 含量、氧化铅 (PbO) 含量、五氧化二磷 (P₂O₅) 含量和氧化钾 (K₂O) 含量五种化学成分制作频率分布直方图, 借助 Python 编程求解, 结果如下图 1 所示:

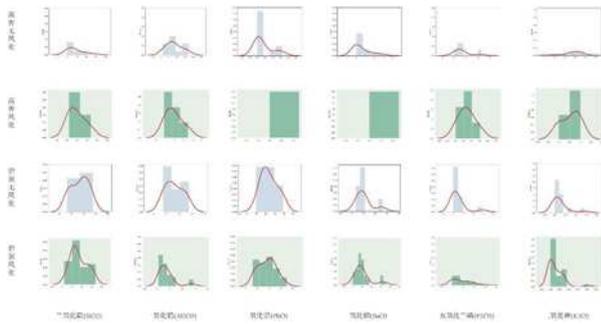


图 1 各成分频率分布直方图

由频率分布直方图可以明显观察到高钾玻璃风化后主要化学成分呈下降趋势, 其中氧化钡 (BaO)、氧化铅 (PbO) 和氧化钾 (K₂O) 含量下降较为明显; 铅钡玻璃在风化后主要化学成分含量呈上升趋势, 氧化铝 (Al₂O₃)、氧化钡 (BaO) 和氧化铅 (PbO) 含量增加较为明显。

三、玻璃文物分类简易规律

根据上述分析可得, 玻璃文物在风化前后主要的化学成分发生了明显的变化, 故我们根据玻璃文物中主要化学成分的含量可以对玻璃文物的类型进行大致划分。

(一) Excel 可视化处理

将已有数据利用 Excel 分别将未风化的高钾玻璃和铅钡玻璃的化学成分进行可视化, 如图所示:

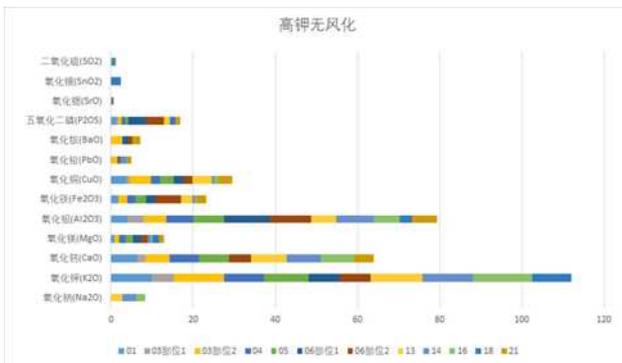


图 2 无风化高价玻璃化学成分占比柱形图

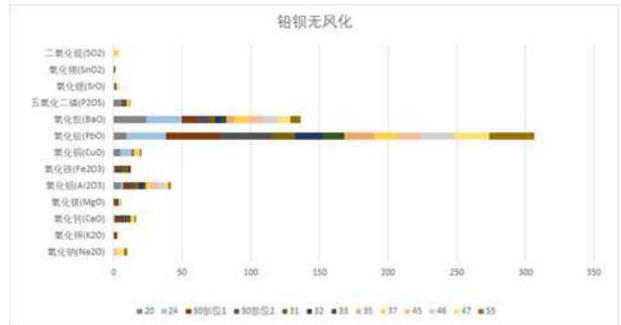


图 3 无风化铅钡玻璃化学成分占比柱形图

由图可知, 在表面无风化的情况下, 对于高钾玻璃, 玻璃中的氧化钾 (K₂O) 和氧化铝 (Al₂O₃) 含量相对较高, 故可依据玻璃中的氧化钾 (K₂O) 和氧化铝 (Al₂O₃) 含量来分类是否为高钾玻璃; 而铅钡玻璃的氧化铅 (PbO) 和氧化钡 (BaO) 含量显著, 故可依据玻璃中氧化铅 (PbO) 和氧化钡 (BaO) 含量占比是否最大来分类铅钡玻璃。根据这几种含量在玻璃中的含量作为两种玻璃的分类依据, 从而形成未风化玻璃分类模型。

(二) 随机森林模型

现采用随机森林分类 [2], 以玻璃类型为因变量, 其余化学成分作为自变量, 得到特征重要值结果如下图 4 所示:

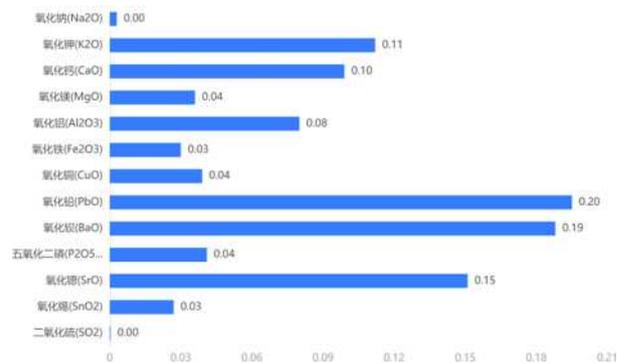


图 4 各化学成分特征重要值

观察得知氧化铅 (PbO) 和氧化钡 (BaO) 是分类两种玻璃的主要依据。继续观察测试结果与预测结果 (详细结果见附件) 可知, 氧化铅 (PbO) 和氧化钡 (BaO) 的含量较其他化学成分显著高时, 即为铅钡玻璃; 氧化钾 (K₂O) 含量较其他化学成分显著高, 氧化铝 (PbO) 和氧化钡 (BaO) 较低时, 即为高钾玻璃。

四、总结

通过以上对玻璃文物不同方面的分析和研究, 我们发现玻璃文物表面风化与玻璃的纹饰是不相关的, 与玻璃的颜色和类型是相关的; 接着, 我们又对玻璃文物风化前后主要化学成分进行了分析, 找出来其主要变化的几种化学成分, 从而开发出另一种和建议的玻璃分类模型, 这种玻璃分类模型对于玻璃文物初期的检验是具有一定的帮助的。

参考文献:

[1] 李濛,包蕾,胡毅,成嵩,胡晓波,高鹰.基于卡方检验的随机数在线检测方法的实现[J].微电子学,2022,52(03):388-392.DOI:10.13911/j.cnki.1004-3365.210329.

[2] 丁炜,徐毅,金有杰等.基于随机森林的大坝潜在风险预测方法[J].水利信息化,2023,No.172(01):46-50. DOI:10.19364/j.1674-9405.2023.01.009.