

博士论文

基于Web页面结构和主色调的聚类算法

赵涓涓, 陈俊杰, 李元俊

(太原理工大学计算机与软件学院, 太原 030024)

收稿日期 修回日期 网络版发布日期 接受日期

摘要 针对目前Web聚类准确率不高的问题, 提出一种基于Web页面链接结构和页面中图片主色调特征的聚类算法。通过分析Web页面中的链接结构和Web页面中所显示图片的主色调来比较页面之间的相似度, 对Web站点中的Web页面进行聚类。聚类过程兼顾Web页面结构和页面的主要色彩特征。系统实验结果表明, 该算法能有效提高聚类的准确性。

关键词 [聚类](#); [Web挖掘](#); [链接结构](#); [主色调](#)

分类号 [TP391](#)

DOI:

通讯作者:

作者个人主页: [赵涓涓](#); [陈俊杰](#); [李元俊](#)

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF](#) (355KB)

▶ [\[HTML全文\]](#) (0KB)

▶ [参考文献\[PDF\]](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [引用本文](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

相关信息

▶ [本刊中 包含“聚类; Web挖掘; 链接结构; 主色调”的 相关文章](#)

▶ [本文作者相关文章](#)