

机器学习与数据挖掘

化学结构图中化学键信息的自动提取

孙兰兰^{1,2}, 李存华^{2*}, 管燕²

1. 中国矿业大学计算机科学与技术学院, 徐州 221116; 2. 淮海工学院计算机工程学院, 江苏 连云港 222005

摘要: 在化学结构图中, 拐点信息异常重要, 如果能对化学结构的拐点进行有效的判别, 就可使化学键信息提取的有效性得到很大改善。本研究通过对化学结构的图形分析及特征总结, 权衡了Hough变换等提取方法的优势和不足, 发现了结构图边缘点偏移值的变化具有很强的规律性, 其对图形拐点的提取以及化学键键型的判断有着重要作用。基于以上分析, 提出了利用化学结构图偏移值变化的规律性进行化学键信息提取的方法。为了论证算法的有效性, 实验时选取了100幅化学结构图, 对100幅图形的化学键信息进行了提取并进行了统计分析, 得出化学键信息的正确提取率为73.83%, 从而通过实验数据证实了算法的可行性。

关键词: 化学结构图 化学键 直线段提取 拐点

The automatic extraction of the chemical bonds information in the chemical structure images

SUN Lan-lan^{1,2}, LI Cun-hua^{2*}, GUAN Yan²

1. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China; 2. School of Computer Engineering, Huaihai Institute of Technology, Lianyungang 222005, China

Abstract: The inflection points are very important for chemical structure graphs. If the inflection points of chemical structures could be discriminant effectively, they would make the effectiveness of the extraction of chemical bond information improved greatly. The characteristics of the chemical structure was analyzed and summarized, and the advantages and disadvantages of the Hough method was balanced. It was found that the offset value of the edge points was with strong regularity, which played an important role on the extraction of the inflection points and the judgment of chemical bond type. Finally, the experiments selected 100 BMP figures of the molecular structure to extract and count the chemical bond information. The statistic results showed that the accuracy rate of chemical bonds information extraction was of 73.83%, which proved the effectiveness of this method.

Keywords: chemical structure graphs chemical bonds the line segments extraction inflection points

收稿日期 2012-03-20 修回日期 网络版发布日期

DOI:

基金项目:

江苏连云港市科技攻关资助项目(CG0923)

通讯作者: 李存华(1963-), 男, 江苏徐州人, 教授, 博士, 主要研究方向为数据库理论, 数据挖掘和计算机图形学. E-mail: cli@hhit.edu.cn

作者简介: 孙兰兰(1986-), 女, 江苏盐城人, 硕士研究生, 主要研究方向为图像处理. E-mail: 172960793@qq.com

作者Email:

PDF Preview

参考文献:

本刊中的类似文章

1. 管燕, 李存华*, 仲兆满, 孙兰兰. 化学分子结构图分割算法[J]. 山东大学学报(工学版), 2012, 42(5): 65-70

扩展功能

本文信息

Supporting info

PDF(1494KB)

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

本文关键词相关文章

化学结构图

化学键

直线段提取

拐点

本文作者相关文章

PubMed