

机器学习

ISOMAP算法参数的递增式选取

邵超

河南财经学院 信息学院, 郑州 450002

收稿日期 2008-4-30 修回日期 2008-5-29 网络版发布日期 2008-7-17 接受日期

摘要 ISOMAP算法能否被成功应用依赖于其唯一参数——邻域大小的选取是否合适,然而,如何高效地选取一个合适的邻域大小目前还是一个难题。当邻域大小变得不合适时,短路边将会出现在邻域图中,从而严重破坏与之相关的最短路径距离对测地距离的逼近能力。和非短路边不同,短路边的两个端点虽然在欧氏空间中相距较近,但在流形上却相距甚远。基于短路边的这一特点,采用序来近似度量一条边的两个端点在流形上的远近程度,因而能够递增式地对邻域大小进行合适的选取。和基于残差的参数选取方法不同,该方法只需递增式地运行广度优先搜索算法,而无需就每一个可能的邻域大小分别运行整个ISOMAP算法,从而具有比较高的运行效率。最终的实验结果证实了该方法的可行性。

关键词 [数据可视化](#) [ISOMAP](#) [邻域大小](#) [残差](#) [短路边](#) [序](#) [广度优先搜索](#)

分类号

Incremental parameter selection for ISOMAP algorithm

SHAO Chao

School of Information, Henan University of Finance and Economics, Zhengzhou 450002, China

Abstract

The success of the ISOMAP algorithm depends greatly on the suitable selection of its only one parameter, i.e.the neighborhood size; however, it's an open problem how to do this efficiently. When the neighborhood size becomes unsuitable, shortcut edge can be introduced into the neighborhood graph and destroy the approximation ability of the involved shortest-path distances to the corresponding geodesic distances greatly. Unlike non-shortcut edge, shortcut edge links two endpoints lying close in Euclidean space but far away on the manifold, which can be measured approximately by its order presented in this paper. Based on the observation, this paper presents an efficient method to find a suitable neighborhood size, which only requires running the breadth-first search algorithm incrementally, but doesn't need to run the whole ISOMAP algorithm for every possible neighborhood size as those methods based on residual variance do. Finally, the feasibility of this method can be verified by experimental results.

Key words [data visualization](#) [ISOMAP](#) [neighborhood size](#) [residual variance](#) [shortcut edge](#) [order](#) [breadth-first search](#)

DOI: 10.3778/j.issn.1002-8331.2008.21.033

通讯作者 邵超 sc_flying@163.com

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(996KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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](#)
- ▶ [复制索引](#)
- ▶ [Email Alert](#)
- ▶ [文章反馈](#)
- ▶ [浏览反馈信息](#)

相关信息

- ▶ 本刊中 [包含“数据可视化”的相关文章](#)
- ▶ 本文作者相关文章
- [邵超](#)