

数据库、信号与信息处理

## 新的连续自编码网络流形学习研究

胡昭华<sup>1</sup>, 宋耀良<sup>1</sup>, 吴佑林<sup>2</sup>

1.南京理工大学 电子工程与光电技术学院, 南京 210094

2.南京技师学院 电气工程系, 南京 210046

收稿日期 2008-6-12 修回日期 2008-9-8 网络版发布日期 2009-11-6 接受日期

**摘要** 发现高维观测数据空间的低维流形结构, 是流形学习的主要目标。在前人利用神经网络进行非线性降维的基础上, 提出一种新的连续自编码(Continuous Autoencoder, C-Autoencoder)网络, 该方法特别采用CRBM(Continuous Restricted Boltzmann Machine)的网络结构, 通过训练具有多个中间层的双向深层神经网络可将高维连续数据转换成低维嵌套并继而重构高维连续数据。特别地, 这种连续自编码网络可以提供高维连续数据空间和低维嵌套结构的双向映射, 不仅有效解决了大多数非线性降维方法所不具备的逆向映射问题, 而且特别适用于高维连续数据的降维和重构。将C-Autoencoder用于人工连续数据的实验表明, C-Autoencoder不仅能发现嵌入在高维连续数据中的非线性流形结构, 也能有效地从低维嵌套中恢复原始高维连续数据。

**关键词** [连续自编码网络\(C-Autoencoder\)](#) [高维数据](#) [降维](#) [重构](#)

分类号 [TP391](#)

## Manifold learning about novel continuous autoencoder network

HU Zhao-hua<sup>1</sup>, SONG Yao-liang<sup>1</sup>, WU You-lin<sup>2</sup>

1.School of Electronic Engineering and Optoelectronic Technology, Nanjing University of Science and Technology, Nanjing 210094, China

2.Department of Electrical Engineering, Nanjing Teachers University of Technology, Nanjing 210046, China

### Abstract

The main goal of manifold learning is to find a low-dimensional manifold embedded in high-dimensional data space. Based on previous nonlinear dimensionality reduction methods using neural network, a novel Continuous Autoencoder (C-Autoencoder) network is put forward in this paper. The method specially uses Continuous Restricted Boltzmann Machine (CRBM) and converts high-dimensional data to low-dimensional codes by training a neural network with multiple hidden layers, and vice versa reconstructs original high-dimensional data. In particular, the C-Autoencoder provides such a bi-directional mapping between the high-dimensional data space and the low-dimensional manifold space and is not only able to overcome the inherited deficiency of most nonlinear dimensionality reduction methods that do not have an inverse mapping but also especially suitable for dimensionality reduction and reconstruction of high-dimensional continuous data. The experiments on synthetic datasets show that the C-Autoencoder network not only can find the embedded manifold of high-dimensional datasets but also reconstruct exactly the original high-dimension datasets from low-dimensional structure.

**Key words** [C-Autoencoder](#) [high-dimensional data](#) [dimensionality reduction](#) [reconstruction](#)

DOI: 10.3778/j.issn.1002-8331.2009.30.047

通讯作者 胡昭华 [zhaohua\\_hu@163.com](mailto:zhaohua_hu@163.com)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(810KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“连续自编码网络\(C-Autoencoder\)” 的相关文章](#)

▶ [本文作者相关文章](#)

· [胡昭华](#)

· [宋耀良](#)

· [吴佑林](#)