

一种基于监督流形学习的头部姿态估计方法

刘勇 厦门大学计算机科学系 福建厦门 361005

【文章摘要】

我们提出一种监督 Laplacian LLE 算法,并结合正则化的最小二乘方法来有效地解决头部姿态估计问题,比传统的流形学习算法能更有效的保持数据的局部几何结构,并且能获得显式的直接映射来处理样本外扩展问题。在 FacePix 数据集的头部姿态估计实验结果表明,我们的算法是有效的,对于训练数据和测试数据,我们提出的算法的性能明显高于其他对比算法的性能。

【关键词】

流形学习;拉普拉斯特征映射;局部线性嵌入;头部姿态估计

人脸识别一直以来都是计算机视觉和模式识别研究中的热点问题,作为生物特征识别的关键技术之一,其特定条件下的人脸识别取得了很大的进展,但头部姿态变化是其中的瓶颈。为了实现对不同头部转动姿态的人脸识别,估计人脸头部姿态因而具有重要的研究意义和实用价值。流形学习自从 2000 被提出以来,到现在已经有众多的方法用来解决人脸识别、姿态估计还有维数约简等方面的实际问题。流形学习理论假设数据是采样于一个高维欧氏空间中的低维流形上,并从高维观测数据中发现并找到潜在的低维流形结构,并构造高维空间到低维空间嵌入的非线性映射或者线性映射,以实现维数约简或数据可视化。目前提出的流行学习方法中比较有影响力的,包括等距特征映射算法 (ISOMAP)、局部线性嵌入算法 (LLE)、拉普拉斯特征映射算法 (LE) 等。而具有头部转动姿态的人脸图像很自然地看成一个具有特定低维流形结构的高维空间,因此采用流形学习进行人脸头像的头部姿态估计获得了比较大的关注,比如 Balasubramanian 等人将流形嵌入扩展到监督学习领域并进行头部姿态估计; BenAbdelkader 等人提出了监督流形学习框架用于头部姿态估计的方法。

本文融合 LE 和 LLE 算法,提出监

督 Laplacian LLE 流形学习算法,针对传统流形学习算法无法获得直接的显式映射的问题,提出用正则化最小二乘来学习直接的显式映射。而且实验验证了我们算法比对比的算法能更加有效的提高头部姿态估计的准确率。

一、基于监督的 Laplacian LLE 流形学习算法

如果存在一组具有嵌入头部姿态流形的数据集,观测空间中的每个样本点被假定是其邻域内的近邻数据点的加权组合,同时针对数据集的局部近邻关系,数据集的头部姿态样本近邻需要得到保持,因此融合 LLE 和 LE 的算法思想,提出 Laplacian LLE 算法,如下:

Step 1: 给定数据集 $\{(\mathbf{x}_i, t_i) | \mathbf{x}_i \in \mathbb{R}^D, t_i \in \{1, \dots, n\}\}$ 其中 n 为样本总数, t_i 为类别标记, 搜索数据集中每个的 k 个最近邻;

Step 2: 定义 $\phi(\mathbf{x}) = \sum_i \|\mathbf{x} - \mathbf{x}_i\|$, 考虑约束 $\sum_i \alpha_i = 1$, 按照最小方差准则求解用于 LLE 项的权矩阵;

Step 3: 定义局部近邻的高斯核权值用于 LE 项:

$$(1) \quad w_{ij} = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), & \mathbf{x}_i \in N_k(\mathbf{x}_j), i, j = 1, \dots, n \\ 0, & \text{others} \end{cases}$$

Step 4: 定义 Laplacian LLE 的目标函数:

$$(2) \quad \Phi_{\text{Laplacian LLE}}(\mathbf{Y}) = \sum_i \|\mathbf{y}_i - \sum_j \alpha_j \mathbf{y}_j\|^2 + \sum_{i,j} w_{ij} \|\mathbf{y}_i - \mathbf{y}_j\|^2$$

其中 $\mathbf{y}_i = \arg \min_{\mathbf{y}} \phi(\mathbf{y})$, 考虑约束 $\sum_i \alpha_i = 0$ 及

$\mathbf{Y}^T \mathbf{Y} = \mathbf{I}$ 求解 $\mathbf{y}^* = \arg \min_{\mathbf{y}} \Phi(\mathbf{y})$, 式 (2) 可重写为求解 $\mathbf{Y}(\mathbf{M} + \lambda \mathbf{L})\mathbf{Y}^T$ 的最小化问题。 \mathbf{Y}^* 的最优解是求矩阵 $(\mathbf{M} + \lambda \mathbf{L})$ 的一组最小本征向量。由约束可知, 本征值为零时, 需要去除, 此时数据集会坍缩至一个点, 因此获得的低维方法是计算矩阵 $(\mathbf{M} + \lambda \mathbf{L})$ 的最小 $(l+1)$ 个本征向量, 并舍弃 $(l+1)$ 中的最小本征值对应的本征向量。

上述提出的 Laplacian LLE 算法能够探索和保持数据集本身具有的特定的局部几何结构, 对于监督学习来说, 要加入监督的类别标签信息来进一步提高头部姿态估计准确率。根据 SML 框架的原理, 加入 $\sum_i \|\mathbf{x}_i - \mathbf{x}_i\|$, 用于监督学习问题中, 其中,

$$\lambda = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), & \mathbf{x}_i \in N_k(\mathbf{x}_j) \\ 0, & \text{others} \end{cases}$$

因此监督 Laplacian LLE 算法描述如下:

$$(3) \quad \Phi_{\text{Supervised Laplacian LLE}}(\mathbf{Y}) = \sum_i \|\mathbf{y}_i - \sum_j \alpha_j \mathbf{y}_j\|^2 + \sum_{i,j} w_{ij} \|\mathbf{y}_i - \mathbf{y}_j\|^2 + \sum_i \|\mathbf{x}_i - \mathbf{x}_i\|^2$$

上式可推导为: $\mathbf{Y}^T \mathbf{M} \mathbf{Y} + \lambda \mathbf{Y}^T \mathbf{L} \mathbf{Y} + \alpha \mathbf{Y}^T \mathbf{L}_s \mathbf{Y}$, 其中, $\mathbf{M} = (\mathbf{I} - \mathbf{Y} \mathbf{Y}^T)$ 。此外, 针对欧式空间的距离度量仅仅定义为样本之间的欧式距离, 由于噪声和冗余信息等因素会造成欧式距离近邻的点并非真正的同类别近邻, 我们加入一个偏差系数来引导欧式距离近邻项同类别近邻靠近, 根据 BME[5] 的思想, 我们对确定局部 K 近邻的欧式距离加入一个系数, $f(t, j) = \frac{t_{\text{class}(t, j)} + \epsilon}{\max_{k \in N_k(t)} t_{\text{class}(t, k)} + \epsilon}$, 其中, $t_{\text{class}(t, j)} = \beta |t - t_j|$ 表示样本之间的类别标记距离, β 都为控制参数。这样局部 K 近邻 N_k 的距离定义为 $\bar{D}(t, j) = D(t, j) f(t, j)$, 其中 $D(t, j)$ 为欧式空间度量。

监督 Laplacian LLE 嵌入能生成一个仅仅定义在训练数据样本上的低维嵌入, 但是监督分类来说则需要一个显式的映射 f 能很好的处理样本外扩展 (out-of-sample extension)。考虑一个线性映射, 我们通过线性变换 $\mathbf{y} = \mathbf{a}^T \mathbf{x}$ 来获得映射 f , 这是线性流形学习采用的方法, 如 LPP 等。实际上, 这个线性映射可能并不存在。因此我们把我们的算法分为两步, 第一步使用监督 Laplacian LLE 求得低维嵌入, 第二步则可以采用 Tikhonov 正则化来拟合从高维输入数据到低维嵌入数据:

$$(5) \quad \min \sum_i \|\mathbf{y}_i - \mathbf{a}^T \mathbf{x}_i\|^2 + \gamma \|\mathbf{a}\|^2$$

特别注意此处的是样本的低维嵌入而不是样本类别标记。(5) 式可以推导为一个闭式解 $\mathbf{a} = (\mathbf{X} \mathbf{X}^T + \gamma \mathbf{I})^{-1} \mathbf{X} \mathbf{y}$, 那么线性映射矩阵 \mathbf{A} 则由 d 个解向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d$ 构成。很自然地, 通过 (5) 式获得的直接的显式线性映射 f 能够充分处理样本外扩展, 同时考虑到训练数据的映射到的低维子空间和测试数据的低维映射子空间的一致性, 我们把训练数据和测试数据使用线性映射都投影至同一低维子空间, 然后再低维子空间的 K 近邻分类算法则能有效的进行头部姿态估计。

二、实验

实验部分主要采用 FacePix 数据集, 示例如 Fig. 1。FacePix 数据集包含 30 个主题, 每个主题包含 181 幅图像, 姿态转动角度为 $[-90, +90]$ 。我们实验使用 15 个主题, 并且使用留一法 (leave-out-one) 进行验证, 即最后一个主题作为测试其余作为训练。

Table 1 显示了我们的算法与 LE+RLS、LLE+RLS 的在不同的头部姿态采样率下的头部姿态估计 MAE 比较, 显

示我们的算法更稳定并且其 MAE(Mean Absolute Error 用来评价估计的头部姿态与真实姿态的绝对偏差)值更低更加有效。Fig. 2 显示了我们的算法与 LE+RLS 和 LLE+RLS 算法在不同的嵌入维数下的 MAE 的比较,结果同样显示在不同的嵌入维数下我们的算法更加鲁棒有效。因此我们提出的 S-Laplacian LLE 算法有效的提高了头部姿态估计的准确率。

三、结论

我们提出了一种两步 (two-step) 的监督 Laplacian LLE 流形学习方法,用于头部姿态估计问题,采用监督 Laplacian LLE 流形嵌入和正则化最小二乘方法结合来有效的处理地维嵌入和样本外扩展问题,实验结果显示我们的算法有效的提高了的头部姿态的准确率。

【参考文献】

[1]M.E and Trivedi, M. M, Head pose

estimation in computer vision: A survey, in: PAMI, 31, 2009, pp. 607-626.
[2]J.Tenenbaum, V. Silva, and J.

Langford, A global geometric framework for nonlinear dimensionality reduction, in: Science, 290, 2000, pp. 2319-2323.

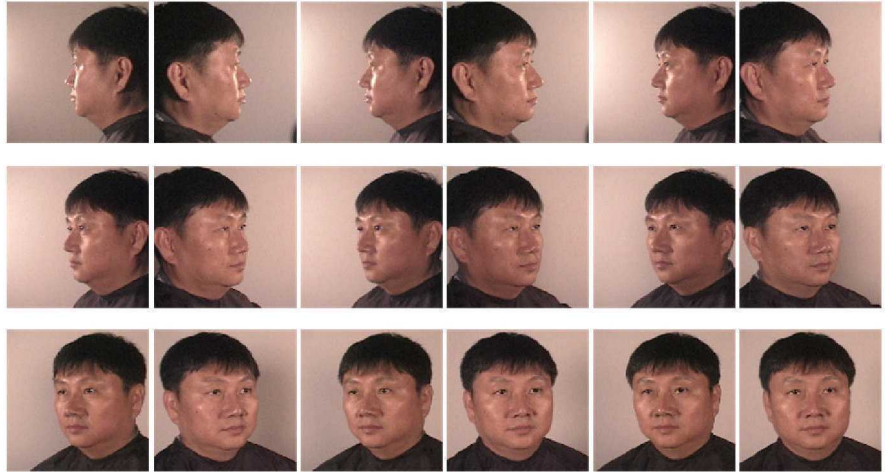


Fig. 1 FacePix 数据集头部姿态图像示例

采样度数间隔	10	8	6	4	2
LE+RLS	4.27	3.63	3.84	3.58	3.83
LLE+RLS	4.11	3.69	3.43	3.14	3.78
S-Laplacian LLE+RLS	3.51	2.79	3.16	2.74	2.70

Table 1 在不同的采样度数间隔下的算法头部姿态估计 MAE 比较 (4NN 分类器)

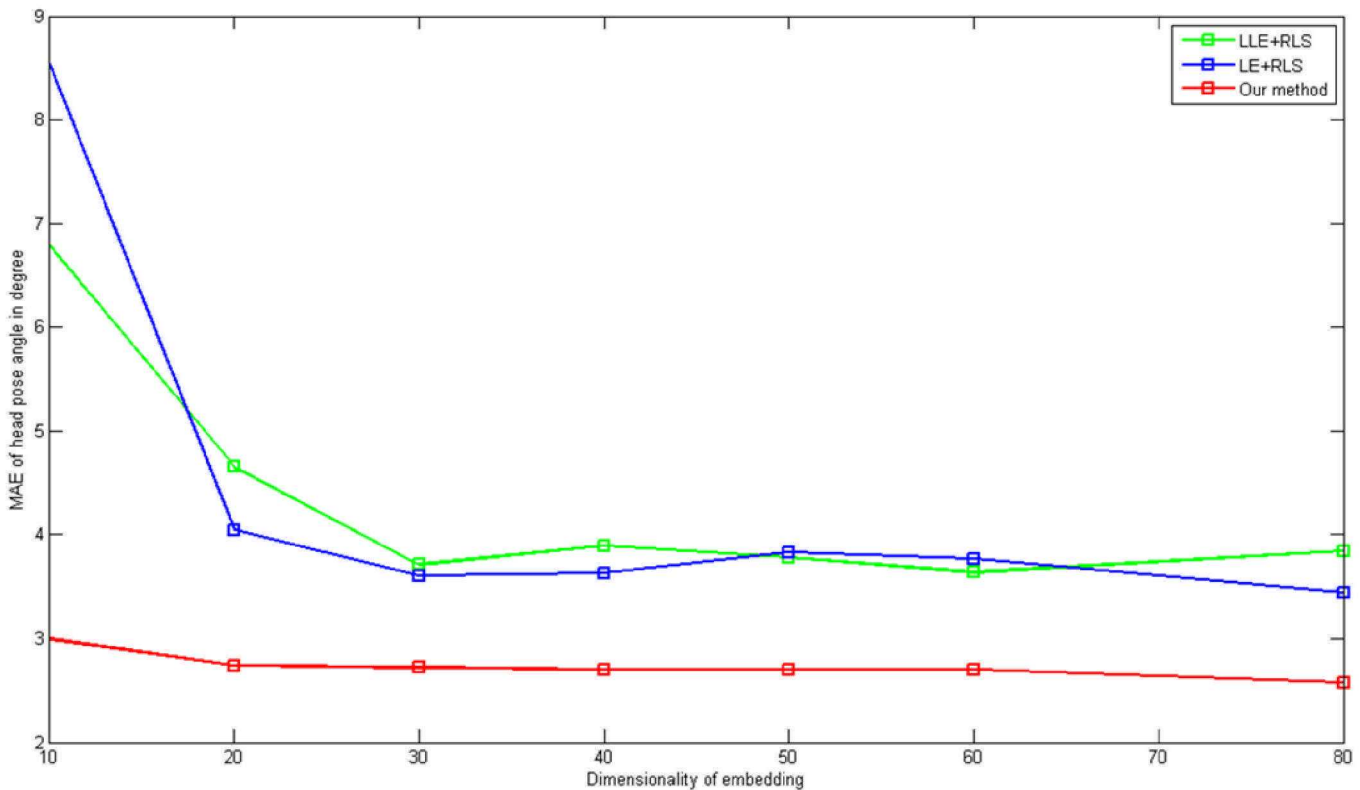


Fig.2 在不同的嵌入维数下算法的性能比对 (4NN 分类器)

- [3] K. S. Lawrence and T. R. Sam, Think globally, fit locally: unsupervised learning of low dimensional manifolds, in: Journal of Machine Learning Research, 4, 2003, pp. 119-155.
- [4] M. Belkin and P. Niyogi, Laplacian eigenmaps for dimensionality reduction

- and data representation, in: Journal of Neural Computation, 15, 2003, 1373-1396.
- [5] V. N. Balasubramanian, J. Ye, and S. Panchanathan, Biased manifold embedding: a framework for person-independent head pose estimation, in:

CVPR, 2007.

- [6] C. BenAbdelkader, Robust Head Pose Estimation Using Supervised Manifold Learning, in: ECCV, 2010, Part VI, LNCS 6316, 518-531.
- [7] X. He and P. Niyogi, Locality preserving projections, in: NIPS, 16, 2004, 100-200.

》接 99 页

倍率下在进行了 80 次的充放电循环后,其容量的保持率还能达到 94.5%。从中可以看出此法合成的 TiO₂ 具有很强的倍率性能。

三、特殊形貌金属氧化物负极材料

Li 等利用模板法合成了 Co₃O₄ 的纳米结构,其中包括纳米管、纳米颗粒以及纳米钉,通过研究得出了如下的结论,形貌不同的 Co₃O₄ 材料,在储锂的性能上存在着较大的差别,其中管状结构比钉状结构的性能要好,颗粒状的性能相对来说最差。张呈乾合成了两种不同形状直径为 1-2 μm 的 Cu₂O 颗粒,对其进行电化学测试后表明,颗粒的形状与电化学性能有一定的关系,正方形颗粒的电化学性能相较于星形颗粒要好,这主要是由于星形颗粒在进行循环时,需要对付内应力的作用,这样就使其分解,而正方形却更加稳

定,这样就保证了其稳定性。

四、总结

随着科学技术的发展,很多电化学性能比较好的金属氧化物的负极材料已经被成功合成出来,这些材料的高容量、稳定循环、优越的倍率性已经逐渐引起人们的关注,而纳米结构的金属氧化物不但拥有上述的优点,而且还有很长的使用寿命,因此一直备受关注。本文主要介绍了三种类型的金属氧化物负极材料,旨在能够对相关的科学研究工作能够起到一定的帮助。

【参考文献】

- [1] 杨同欢,郭永柳,周学酬,刘永梅. 碳包覆纳米 SnO₂ 锂离子电池负极材料合成及其性能研究[J]. 无机材料学报, 2009(01).
- [2] 张玉玺,张晓丽,郝洪河. 锂离子电池负极材料 TiO₂ 的研究进展[J]. 电池, 2009(02).
- [3] 何见超,赵海雷,贾喜娣,王梦微,陈敬波. 锂离子电池 Sn-Co 合金负极材料的研究进展[J]. 化工学报, 2009(05).
- [4] 饶睦敏,黄启明,李伟善. 锂离子电池纳米负极材料的研究进展[J]. 电池工业, 2008(02).

机材料学报, 2009(01).

- [2] 张玉玺, 张晓丽, 郝洪河. 锂离子电池负极材料 TiO₂ 的研究进展[J]. 电池, 2009(02).
- [3] 何见超, 赵海雷, 贾喜娣, 王梦微, 陈敬波. 锂离子电池 Sn-Co 合金负极材料的研究进展[J]. 化工学报, 2009(05).
- [4] 饶睦敏, 黄启明, 李伟善. 锂离子电池纳米负极材料的研究进展[J]. 电池工业, 2008(02).

【作者简介】

李亚娟(1983-),女,新疆人,新疆教育学院理学分院,新疆教育学院多功能材料实验室,硕士研究生,主要研究方向为锂离子电池负极材料。

》接 103 页

酷炫效果所付出的最大的代价。Flash 网站会带有一定的用户交互系统、视频播放系统、音乐播放系统这都会给 Flash 网站载入增加不少的时间。

Flash 网站载入时间长与宽带带宽也有关系。因为 Flash 网站数据大,宽带带宽低两者相加就导致了 Flash 网站载入时间较长。这不仅仅是 Flash 网站遇到的问题,其他任何网站或者游戏都会受到宽带因素的影响。

(4) 优化困难

不得不说这是 Flash 网站对于企业来说最大的缺点。导致这种情况的原因主要有两方面,一方面 Flash 网站多以视频、音频、动画元素为主很少有文字表达,而搜索引擎对于这些元素识别困难。另一方面传统 Flash 网站将文字、图片、视频等数据全部集成在 .swf 文件里,搜索引擎根本无法识别,而不容易支持数据库。这是搜索引擎的问题也是 Flash 全站的问题。

二、Flash 网站发展

未来是一个高用户体验的互联网时代,普通的网站已经很难满足日益高涨的用户需求。越来越多的高端企业用户在选择网站时都更加的趋向于 Flash 网站,尤其是在一些对用户体验较高的行业。未来 Flash 网站要想被越来越多的企业接受成为网站技术的主流必须具备以下几个的条件:

1. Flash 网站技术的普及

随着 Flash 网站技术被越来越多的人

熟悉并掌握,flash 网站在未来的发展将会达到一个新的高度。就如同电脑的普及一样,flash 网站将会因为使用人数的增加而取得快速的发展。

2. 网络宽带的提高

随着计算机技术和通信技术的发展,信息传输的手段发生了极大的变化。人们对各种业务的需求,也越来越提高,要求业务的种类越来越多样化,如语音、数据、图像等各种业务,使得多媒体业务的需求迅速上升。为了满足上述业务迅速上升的需求,这就要求网络建设向宽带化、智能化、综合化方向发展。使得宽带网络成为适应上述业务需求急待加快建设的一种网络。

3. 网站优化技术的逐渐成熟

在现有的网络状态下,使用者经常会遇到带宽拥塞,应用性能低下,蠕虫病毒,DDoS 肆虐,恶意入侵等对网络使用及资源有负面影响的问题及困扰,网络优化功能是针对现有的防火墙、安防及入侵检测、负载均衡、带宽管理、网络防毒等设备及网络问题的补充,能够通过接入硬件及软件操作的方式进行参数采集,数据分析,找出影响网络质量的原因,通过技术手段或增加相应的硬件设备及调整使网络达到最佳运行状态的方法,使网络资源获得最佳效益,同时了解网络的增长趋势并提供更好的解决方案。

4. 企业对用户需求的重视

随着人们生活水平的提高,未来人们对于网站将会更加倾向于体验方面,所以

一些企业为了推广自己的产品,必须重视用户的需求,而 flash 网站正好符合一些高端行业的需求。

三、结语

目前,Flash 网站技术已经基本成熟,Flash 网站由于其视觉冲击力强,互动效果好被广泛的应用在房地产行业、汽车行业和奢侈品行业等高端行业。但是 Flash 网站要想有更大的发展还需软件开发人员、计算机专家、语言学家、心理学家、艺术家等通力合作,共同努力,从理论研究和应用开发上有所突破,相信 Flash 网站定会有更加光明的未来。

【参考文献】

- [1] 王川,武寒青;动画前期创意[J];高等教育出版社;2003(2):1-4
- [2] 贺小霞,方宇等;《ActionScript 3.0 编程特效实战》;清华大学出版社;2010.1
- [3] 徐珂,苗彤,赵晖;网页视觉与设计[M];人民邮电出版社;2001
- [4] 李锐;网页设计与制作实训[M];人民交通出版社,2005

【作者简介】

吴准辉(1990-),男,平顶山学院,专业方向:软件工程(数字艺术方向)