

一种基于改进的 LEA 头部姿态估计方法*

李维清, 陈锻生

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

摘要: 局部嵌入分析(LEA)是图嵌入化的局部线性嵌入(LLE)方法。在头姿态估计问题上,选择局部邻域时只考虑属于同一类的姿态,但失去了相邻姿态的几何拓扑信息。为此,提出一种改进的邻域选择方法,充分利用先验姿态信息,使降维后的流形更加平滑,同类姿态互相靠近,不同类姿态之间的距离随着姿态差值变大而增大,且能够使训练及测试样本的低维流形更加靠近,降低了估计误差。在 Facepix 人脸数据库上的实验证明了该方法的有效性。

关键词: 头部姿态估计;流形学习;图嵌入;局部嵌入分析;局部线性嵌入

中图分类号: TP391.4

文献标识码: A

文章编号: 1674-7720(2011)11-0037-04

An improved method of LEA for head pose estimation

Li Weiqing, Chen Duansheng

(Institute of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

Abstract: Local embedding analysis is a graph embedding algorithm of locally linear embedding(LLE). In head pose estimation, LEA selects local neighborhoods only from those of the same pose, but loss of geometrical and topological information among the neighbor poses. The paper presents an improved method of selecting neighborhoods to make full use of the priori pose information. It makes the manifolds more smooth in reduced dimension, and the points of the same pose get closer while the points of more different poses get wider, so as to reduce the pose estimation error. The experimental results in Facepix database show the effectiveness of the proposed method.

Key words: head pose estimation; manifold learning; graph embedding; LEA; LLE

头部姿态估计是计算机视觉和图像模式识别领域中的一个重要研究课题,近年来受到了越来越多的关注^[1]。在人脸识别中,如果得到的人脸图像是非正面的,识别的效果会大大降低,而如果预先估计人脸姿态后选择合适的视角模型进行识别,将会提高非正面人脸的识别率^[2]。由于通过人脸的姿态可以得知人注视的方向,所以姿态估计在理解人的注意力等方面有很高的研究价值。

对于头部姿态估计问题,现有的方法大致可分为基于表现的方法和基于模型的方法两大类^[3]。基于表现的方法是通过含有各种姿态的人脸图像进行学习,建立一个能够估计姿态的分类器。这种方法对图像的分辨率要求不高,并且不需要或者需要较少的面部特征点,能够估计角度较大范围姿态。基于模型的方法是利用某

种几何模型来表示人脸的结构和形状,并通过提取某些特征,在模型和图像之间建立起对应关系。这种方法严重地依赖于特征点的定位结果,当图像旋转较大角度时,部分面部特征将丢失,无法进行估计,所以估计的姿态范围较小。

1 基于流形学习的头部姿态估计

基于流形学习的头部姿态估计属于表现类方法,它的基本思想是考虑每个高维头姿态图像都处于一个有姿态变化的连续流形中。目前已经吸引了一些学者对它进行研究,例如 HU N 等人^[4]提出了通过对特定人的姿态流形的学习,在假定姿态流形不变的情况下,利用预测网络来估计其他人的图像的姿态的方法;FU Y^[5]等使用了图嵌入 GE(Graph Embedding)^[6]结合流形学习算法进行人脸的姿态估计研究。

流形学习 LLE^[7](Local Linear Embedding)算法是通

* 基金项目:国家自然科学基金(10901062);福建省自然科学基金(2009J01289, 2010J01337);福建省科技计划重点项目(2008J0021)

过建立局部邻域权重图将数据由高维降到低维,但其邻域均采集自同一流形,对于姿态估计,这将导致姿态估计与人有关。FU Y^[5]等人提出了一种改进的LLE即LEA(Locall Embed Analysis):利用数据集的已知类别信息选择局部邻域时,只考虑属于同一类(即同一姿态)的数据点,并结合图嵌入理论,使改进后的LLE近似线性。这样对于姿态估计将大大提高姿态估计的身份无关性。但这又带来一个新的问题:属于同一类的数据集映射到低维空间中后,退化成为一点,失去了几何拓扑信息,并且所有邻域均为同类样本(即不同人的相同姿态),这使得降维后的流形失去了其相邻姿态间的平滑性。

本文对FU Y等人提出的LEA方法做了进一步的改进:由于邻域的选择是流形学习算法至关重要的第一步,关系到邻域样本权值的计算及最后的降维结果。因此,本文在构造邻域时通过改进邻域距离表示方法,更好地选择邻域,使样本的邻域更好地重构样本本身,以解决LEA降维后的流形不能很好地保持高维时所具有的几何拓扑结构的不足,并使训练流形和测试流形更加靠近,减少姿态估计误差。

2 LLE算法^[7]

给定数据集 $X, x_i \in R^M, i=1, \dots, N, N$ 为样本总数, M 为原始空间维数。

(1)邻域选择:定义点 x_i 的 k 邻域为 $\chi_i = \{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}, x_{i_j} \in X, 1 \leq j \leq k$ 。LLE需要计算点 x_i 和邻域点之间的重构权 $w_{i,j}$,重构权 $w_{i,j}$ 的选取通过极小化重构误差来实现:

$$\varepsilon_M = \sum_i \left\| x_i - \sum_{j=1}^k w_{i,j} x_{i,j} \right\|^2 \quad (1)$$

其中,权 $w_{i,j}$ 表示样本 $x_{i,j}$ 对样本 x_i 的重构的贡献。为了计算权值,引入约束条件:对于所有的样本点 $x_i, \sum_{j=1}^k w_{i,j} = 1$ 且当 $x_{i,j} \notin \chi_i$ 时, $w_{i,j} = 0$,此时通过 $\arg \min(\varepsilon_M)$ 求解权值。

(2)低维嵌入:将低维数据集表示为: $Y = \{y_1, y_2, \dots, y_N\}, y_i$ 是 x_i 的低维表示,且 $y_i \in R_m (m < M)$,应用计算出的权值并最小化嵌入代价函数计算出低维数据集 Y 。嵌入代价函数为:

$$\varepsilon_m = \sum_i \left\| y_i - \sum_{j=1}^k w_{i,j} y_{i,j} \right\|^2 \quad (2)$$

利用式(1)计算出的权值组成一个 $k \times N$ 的矩阵 W ,并将其代入式(2),则 ε_m 可表示为: $\varepsilon_m = \text{Tr}(Y(I-W)^T(I-W))$,令 $\Phi = (I-W)^T(I-W)$,则低维数据集 Y 是 Φ 最小的第2个到第 $m+1$ 个特征向量。

3 LEA算法对LLE的改进

为计算出由高维到低维的线性映射矩阵 P ,需要重新计算权重矩阵 W 。点 x_i 定义 $k \times k$ 的局部图矩阵 G_i 为:

$G_i[j, l] = (x_i - x_{i,j})^T(x_i - x_{i,l}), G_i = (x_i I^T - \chi_i)^T(x_i I^T - \chi_i), I^T$ 是由1组成的 $k \times 1$ 的列向量。 X_i 是 $M \times k$ 的矩阵,包括 x_i 的 k 最近邻样本点。列向量 $w_i = (w_{i,1}, w_{i,2}, \dots, w_{i,k})^T$,则式(1)可表示为

$\varepsilon_M = \sum_{i=1}^N w_i^T G_i w_i$ 。解最小二值问题并取得:

$$w_i = \frac{G_i^{-1} 1}{1^T G_i^{-1} 1} \quad (3)$$

定义 $n \times n$ 稀疏权值矩阵 $W, W[i, i_j] = w_i(j) = w_{i,j}$,矩阵 W 的其他值为0。

LEA将图嵌入的思想用于LLE中,计算出样本的局部邻域后,在用LLE算法计算数据集 X 的低维投影时,考虑低维数据集 $Y = P^T X$,计算出投影矩阵 P ,继而后用公式 $Y = P^T X$ 计算 Y ,而不是直接用LLE算法计算出 Y ,这样就将LLE算法改进为近似线性的LEA方法。

4 改进的算法

(1)邻域选择

LEA在对点 x_i 选取邻域点时,将不同人的相同姿态点作为其邻域点,是为了满足监督图嵌入的要求。这样在理论上满足不同对象的相同姿态靠近,但是会使得降维后的同类间样本点基本重合在一起,退化成为一点,失去了原有的几何拓扑信息,在计算时没有考虑相邻姿态之间的影响,失去了相邻姿态间的平滑性。本文考虑在计算邻域时加入姿态信息,使得在计算投影矩阵时,既考虑相同姿态的影响,也考虑相邻姿态的影响。样本点距离计算如下:

$$D = d \times A(i, j) / A_{\max} \quad (4)$$

其中, (i, j) 为两姿态间的姿态差值, A_{\max} 为姿态之间的最大差值, d 为样本点之间的欧式距离,这样使同类姿态之间的距离变小,不同类姿态间距离随着姿态差值的增大而增大,符合高维空间中样本之间的几何结构,且通过此距离公式计算样本的邻域,即能全部包括同类姿态样本,并且可以包括部分相邻姿态样本。LEA算法中只将同类姿态作为样本的邻域,这样仅能保证同类样本靠近,而不能保证不同类样本之间距离因姿态差值的变化而变化,并且忽视了相邻姿态对样本重构的贡献。

利用式(4)选择邻域,并用式(3)计算权值矩阵 W 后,可考虑将样本集 X 嵌入到低维。

(2) M 到 m 的嵌入

嵌入的问题是如何将 M 维空间映射到 m 维子空间。为了得到线性映射矩阵 P ,应用LEA^[5]算法对LLE^[7]的近似线性化思想,将 $Y = P^T X$ 代入式(2)得:

$$\varepsilon_m = \text{Tr}[P^T X(I-W)^T(I-W)X^T P] \quad (5)$$

约束为 $\sum_{j=1}^k w_{i,j} = 1$ 。图嵌入的约束 $P^T(XX^T)P = nI$,通过拉格朗日优化后的特征值为:

$$X(I-W)^T(I-W)X^T P = \Lambda X X^T P \quad (6)$$

《微型机与应用》2011年第30卷第11期

其中 Λ 是拉格朗日乘数对角矩阵。列 P 是矩阵 $(XX^T)^{-1}X(I-W)^T(I-W)X^T$ 去掉底部特征向量 $P^T X$ 之后的最小的 m 个特征向量。

对于头部姿态估计问题,本文提出算法的流程为:

(1) 训练姿态流形

① 裁剪图片,使图片仅包含头部姿态部分,并对图片预处理、归一化,使所有图片有相同大小。

② 提取特征作为训练特征(也可以不提取,直接用图片像素作为特征),并将特征用一个列向量来表示。

③ 根据式(6)计算样本点之间的距离,求出邻域矩阵,接着求解式(3),计算权值矩阵,然后求解式(5)计算投影矩阵 P 。

④ 应用 $Y=P^T X$, 计算低维映射 Y 。

(2) 测试样本姿态估计

① 同训练步骤①,对测试图片进行裁剪、预处理、归一化等操作。

② 应用投影矩阵 P 计算出测试样本的低维表示。

③ 应用 KNN 分类器估计测试样本姿态。

5 实验

5.1 人脸库

为了验证算法的有效性,本文在 FacePix 人脸姿态数据库上进行了实验。FacePix 人脸库是 2005 年由 CUBiC (the Center for Cognitive Ubiquitous Computing)、Arizona State University 提供,该人脸库包含了不同姿态、不同光照的人脸,本文只介绍不同姿态的图片:具体为 30 人,每人 181 张不同姿态的人脸图像,姿态范围为水平方向上从 $-90^\circ \sim 90^\circ$ (负表示向左旋转),间隔为 1° ,共计 5 430 张分辨率为 128×128 的彩色人脸图像。本文将图片裁剪为 32×32 (人脸库中的第 16、21、27 三人由于图像采集不好,未被纳入实验中)大小的图片。人脸库样例及低维可视化流形如图 1 所示。

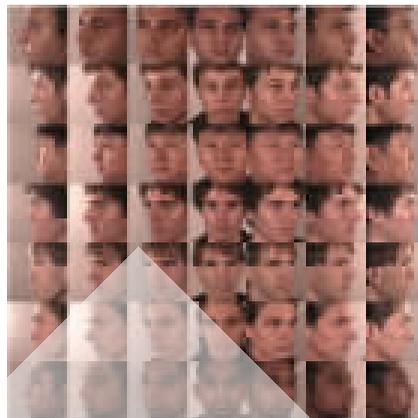
5.2 实验结果及分析

(1) 低维可视化效果

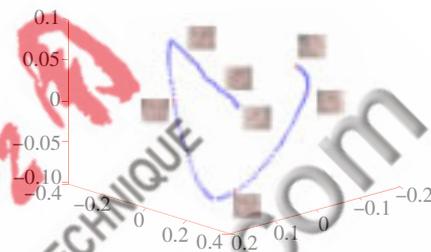
图 1 (a) 是 FacePix 人脸库经裁剪后的部分样例图,按照每行为同一人,每列为同一姿态排列,姿态从左到右分别为 -90° 、 -60° 、 -30° 、 0° 、 30° 、 60° 、 90° ;图 1 (b) 是 FacePix 人脸库中第一个人的 181 张姿态图像经本文改进的 LEA 算法降维后的三维嵌入流形,嵌入流形的图片姿态按照 -90° 、 -60° 、 -30° 、 0° 、 30° 、 60° 、 90° 排列,邻域 $K=80$,特征为裁剪并处理后的灰度图。由图 1 (b) 可以看出,不同姿态处在低维不同位置,且按照姿态顺序呈流形分布。

(2) 头部姿态实验

训练及测试样本三维流形如图 2 所示,图 2 实验选取的特征均为裁剪并处理后的灰度图。图 2 (a) 为 LEA 算法的低维嵌入图,邻域 $k=8$,图中颜色较深的线为人脸库中前 9 个人的流形,为训练流形;颜色较浅的线为中间



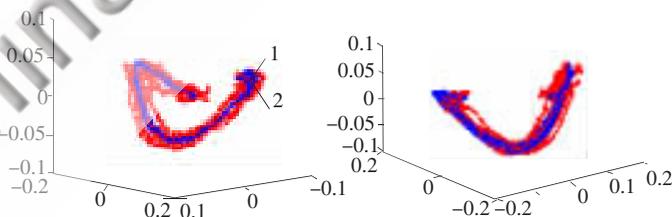
(a) FacePix 人脸库经裁剪后的部分样例图



(b) 经本文改进的 LEA 算法降维后的三维嵌入流形

图 1 人脸库样例及低维可视化流形

9 个人应用训练出的投影矩阵 P 投影后的结果,为测试流形。图 2 (b) 为本文改进后的算法的嵌入图,邻域 $k=13$,图中不同颜色的含义同图 2 (a)。通过图中效果比可以看出,改进后的算法更能使测试样本和训练样本的相同姿态靠近,利于分类误差的降低。



(a) LEA 算法的低维嵌入图

(b) 本文改进后的算法的嵌入图

图 2 训练及测试样本三维流形

对 LEA 算法和改进后的算法做相同条件下的对比试验。分别选取 FacePix 人脸库中前 9 人、前 12 人、前 15 人、前 18 人做训练样本,对应的后 18 人、后 15 人、后 12 人、后 9 人做测试样本,每人 181 张图片。由于改进后的算法仍是基于 LLE 算法的,所以邻域、嵌入维数以及参与训练的图片数对实验效果均有一定影响。实验中的特征均为裁剪并处理后的灰度图。表 1 为实验的姿态估计平均误差表。

图 3 中实验为:低维维度 $m=14$,训练样本为 9 个人 1 629 张图片,测试样本为 18 个人 3 258 张图片,LEA 算法邻域取 $k=8$,改进算法邻域取 $k=10$ 。图中实线为 LEA 算法姿态估计误差,其平均误差为 3.44° ;虚线为改进算法姿态估计误差,其平均误差为 2.99° (如表 1 所示)。

表 1 FacePix 人脸库头部姿态估计平均误差(度)

算法	人数	m							
		6	8	10	12	14	16	18	20
LEA	9(k=8)	3.64	3.43	3.31	3.37	3.44	3.49	3.57	3.60
	12(k=11)	3.51	3.39	3.28	3.26	3.33	3.46	3.52	3.59
	15(k=14)	3.51	3.37	3.24	3.20	3.18	3.27	3.42	3.51
	18(k=17)	3.43	3.31	3.23	3.19	3.17	3.17	3.20	3.28
改进算法	9(k=10)	3.38	3.13	3.05	2.98	2.99	2.95	2.91	2.91
	12(k=14)	3.26	3.11	3.04	3.00	2.94	2.84	2.89	2.89
	15(k=16)	3.29	3.09	3.01	2.94	2.85	2.72	2.76	2.90
	18(k=20)	3.13	3.04	2.96	2.85	2.77	2.69	2.64	2.68

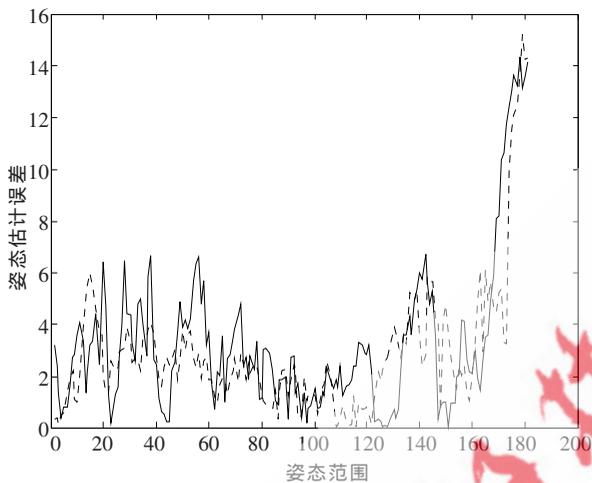


图 3 改进算法与 LEA 算法姿态估计误差对比

从表 1 及图 3 可以看出,改进后的算法与原来的算法相比,其误差降低不少。主要原因:如图 2 所示,由于 LEA 算法的邻域取自同姿态样本,其缺点是降维后同类样本重合在一起,理论上是类间距离越小越好。但是由于人的差异性,同样的姿态不同的人会有差距,所以导致训练出的流形与测试样本的流形有很大差距。改进算法由于适当扩大邻域,既包括同类样本又包括姿态相近的样本,这样训练流形与测试流形的差距就会缩小。

本文提出了一种对局部嵌入分析(LEA)算法改进的头部姿态估计方法(即一种新的邻域选择方法),在邻域

选择时充分利用先验姿态信息,使降维后流形更加符合高维时的姿态间的几何关系,降低姿态估计误差。由实验可知,本文对 LEA 算法改进的有效性。然而由于流形学习算法的实验结果与参数(如邻域 k 、降维维度 m 等)有很大有关,并且数据库由于图像裁剪不同,实验效果也不尽相同,因此算法还有待进一步的研究与探讨。

参考文献

- [1] CHUTORIAN E M, TRIVEDI M M. Head pose estimation in computer vision: a survey [J]. IEEE transactions on Pattern Analysis and Machine Intelligence, 2009,31(4):607-626.
- [2] LI S Z, FU Q D. Kernel machine based learning for multi-view face detection and pose estimation [C]. Proceedings of 8th IEEE International Conference on Computer Vision. Vancouver, Canada; 2001.
- [3] 马丙鹏. 基于表观的人脸姿态估计问题研究[M].北京:中国科学院,2009.
- [4] HU N, HUANG W, RANGANATH S. Head pose estimation by non-linear embedding and mapping [C]. Proceeding. IEEE International Conference on Image Processing. 2005.
- [5] FU Y, HUANG T S. Graph embedded analysis for head pose estimation [C]. Proceeding IEEE International Conference on Automatic Face and Gesture Recognition. 2006.
- [6] YAN S, XU D, ZHANG B, et al. Graph embedding: a general framework for dimensionality reduction [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005.
- [7] ROWEIS S, SAUL L. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000,290:2323-2326.

(收稿日期:2011-01-10)

作者简介:

李维清,男,1986年生,硕士研究生,主要研究方向:图像处理,模式识别。

陈锻生,男,1959年生,博士,教授,主要研究方向:图像模式识别与计算机视觉。