

Journal of Image  
and Graphics

# 中国图象图形学报



ISSN1006-8961  
CN11-3758/TB

2012 **8**  
Vol.17 No.

中国科学院遥感应用研究所  
中国图象图形学学会主办  
北京应用物理与计算数学研究所

# 中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年8月 第17卷 第8期(总第196期)

## 目次

### 综述

可计算图像美学研究进展 ..... 王伟凝, 蚁静斌, 贺前华(893)

### 图像处理和编码

利用混沌 PSO 或分解的 2 维 Tsallis 灰度熵阈值分割 ..... 吴一全, 吴诗嫻, 张晓杰(902)

H. 264 并行编码中负载平衡方法 ..... 侯兴松, 刘大齐, 盛凯, 赖玉洁(911)

提升细节捕捉能力的非下采样轮廓波变换 ..... 郭洪, 李雪军(919)

SVD 域的图像高斯噪声强度估计 ..... 柳薇(923)

图像扩散去噪模型的分析与改进 ..... 傅绪加, 刘峰, 王信松(934)

小波域最小嵌入失真函数设计及其在隐写中的应用 ..... 李军, 潘峰, 杨晓元, 刘镇(946)

### 图像分析和识别

抗压缩的照片图像与真实感计算机图形识别 ..... 张荣, 王让定(953)

云的自适应阈值检测方法 ..... 黄茜, 王少龙, 徐道磊, 廖婷婷(960)

煤矸区分中的间隔灰度压缩扩阶共生矩阵 ..... 于国防(966)

离群样本划分的半监督模糊学习策略 ..... 宋晓宁, 杨静宇, 杨习贝(971)

多相 Chan-Vese 模型的直接对偶方法 ..... 刘存良, 潘振宽, 郑永果, 王国栋, 丁洁玉(979)

利用梯度信息快速提取直线边缘特征 ..... 夏军营, 徐小泉, 熊九龙(987)

2 维双树复小波不确定度加权融合的人脸识别 ..... 王仕民,叶继华,邓涛,王明文(995)

图像理解和计算机视觉

流形学习与非线性回归结合的头面部姿态估计 ..... 范进富,陈锻生(1002)

计算机图形学

角点距离矩阵和同心圆划分的曲线描述与匹配 ..... 曾接贤,刘秀朋,符祥(1011)

弹性二次曲面模型下的 3 维形状表现与平滑 ..... 郝世杰,蒋建国,郭艳蓉,詹曙(1021)

医学图像处理

主动学习的白细胞图像自动分割 ..... 崔凤,潘晨,吴向平,徐军(1029)

遥感图像处理

Bregman 交替迭代遥感图像复原方法 ..... 徐焕宇,孙权森,罗楠,夏德深(1035)

基于各向异性 Markov 随机场的遥感影像亚像元尺度建筑物提取 ..... 李晓冬,凌峰,杜耘(1042)

中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 8 期

2012 年 8 月 16 日出版

主管单位 中国科学院  
主 办 中国科学院遥感应用研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所  
主 编 李小文  
编辑出版 《中国图象图形学报》编辑出版委员会  
北京 9718 信箱 邮编 100101  
电子信箱:jig@irsa.ac.cn  
电话:010-64807995 010-82614429  
网 址:www.cjig.cn  
印刷装订 北京北林印刷厂  
广告经营许可证 京朝工商广字第 0346 号  
总 发 行 北京报刊发行局  
订 购 全国各地邮局  
国外发行 中国国际图书贸易总公司  
(中国国际书店)  
(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences  
Sponsored by Institute of Remote Sensing Application,  
CAS China Society of Image and Graphics  
Institute of Applied Physics and Computational  
Mathematics  
Chief editor LI Xiaowen  
Editor, Publisher Editorial and Publishing Board  
of Journal of Image and Graphics  
(P. O. Box 9718, Beijing 100101, China)  
E-mail:jig@irsa.ac.cn  
Distributed by Beijing Bureau for Distribution of Newspapers  
and Journals  
Domestic All Local Post Offices in China  
Foreign China International Book Trading Corporation  
(P. O. Box 399, Beijing 100044, China)  
Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

# Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 8 August 2012

## Contents

### Review

Review for computational image aesthetics ..... Wang Weining, Yi Jingjian, He Qianhua (893)

### Image Processing and Coding

Two-dimensional Tsallis gray entropy image thresholding using chaotic particle swarm optimization or decomposition  
..... Wu Yiquan, Wu Shihua, Zhang Xiaojie (902)

Load balancing for H. 264 parallel encoding ..... Hou Xingsong, Liu Daqi, Sheng Kai, Dun Yujie (911)

Non-subsampled contourlet transform algorithm to promote detail information capturing ability ..... Guo Hong, Li Xuejun (919)

Gaussian noise level estimation in SVD domain for images ..... Liu Wei (923)

Analysis and improvement of image diffusion denoising models ..... Fu Xujia, Liu Feng, Wang Xinsong (934)

Design of minimizing embedding impact function in DWT domain and its applications in steganography  
..... Li Jun, Pan Feng, Yang Xiaoyuan, Liu Zhen (946)

### Image Analysis and Recognition

Anti-compression approach to distinguishing photographic images and photorealistic computer graphics  
..... Zhang Rong, Wang Rangding (953)

Adaptive threshold approach for night-sky cloud inspection ..... Huang Qian, Wang Shaolong, Xu Daolei, Liao Tingting (960)

Expanded order co-occurrence matrix to differentiate between coal and gangue based on interval grayscale compression  
..... Yu Guofang (966)

Semi-supervised fuzzy learning strategy by using a way of partitioning the outlier instances  
..... Song Xiaoning, Yang Jingyu, Yang Xibei (971)

Direct dual method of Chan-Vese model for multiphase segmentation  
..... Liu Cunliang, Pan Zhenkuan, Zheng Yongguo, Wang Guodong, Ding Jieyu (979)

Fast line edge feature extraction using image gradients ..... Xia Junying, Xu Xiaoquan, Xiong Jiulong (987)

Two-dimensional dual-tree complex wavelet transform uncertainty weighted fusion in face recognition  
..... Wang Shimin, Ye Jihua, Deng Tao, Wang Mingwen (995)

### Image Understanding and Computer Vision

Combining manifold learning and nonlinear regression for head pose estimation ..... Fan Jinfu, Chen Duansheng (1002)

### Computer Graphics

Representation and matching for planar curve based on corner distance matrix and concentric circles  
..... Zeng Jiexian, Liu Xiupeng, Fu Xiang (1011)

Three-dimensional shape representation and smoothing with elastic quadratic patch model  
..... Hao Shijie, Jiang Jianguo, Guo Yanrong, Zhan Shu (1021)

### Medical Image Processing

White blood cell image segmentation based on active learning ..... Cui Feng, Pan Chen, Wu Xiangping, Xu Jun (1029)

### Remote Sensing Image Processing

Bregman alternating iterative method for remote sensing image restoration ..... Xu Huanyu, Sun Quansen, Luo Nan, Xia Deshen (1035)

Building extraction at the sub-pixel scale from remotely sensed images based on anisotropic Markov random field  
..... Li Xiaodong, Ling Feng, Du Yun (1042)

中图分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2012)08-1002-09

论文引用格式: 范进富, 陈锻生. 流形学习与非线性回归结合的头部姿态估计[J]. 中国图象图形学报, 2012, 17(8): 1002-1010.

# 流形学习与非线性回归结合的头部姿态估计

范进富, 陈锻生

华侨大学计算机科学与技术学院, 厦门 361021

**摘要:** 流形学习的目的是发现非线性数据的内在结构, 可用于非线性降维。广义回归网络是人工神经网络的一种, 可用于非线性回归。基于流形学习和非线性回归, 提出了用于解决头部姿态估计的 ManiNLR 方法。该方法首先用流形学习对图像数据进行降维, 然后用非线性回归的方法将数据映射到线性可分空间, 利用非线性回归的结果对人脸的头部姿态进行估计。实验结果表明, ManiNLR 算法能够较好地估计图像中的头部姿态, 并具有较快的速度和较高的鲁棒性。

**关键词:** 流形学习; 头部姿态估计; 非线性回归; 人工神经网络

## Combining manifold learning and nonlinear regression for head pose estimation

Fan Jinfu, Chen Duansheng

College of Computer Science and Technology, HuaQiao University, Xiamen 361021, China

**Abstract:** Manifold learning attempts can be used to obtain the intrinsic structure of the non-linear data, which can be used in non-linear dimensionality reduction. The general regression neural network (GRNN) is a kind of artificial neural network, which can be used in non-linear regression. In this paper, the ManiNLR method, which is based on manifold learning and nonlinear regression, is proposed for head pose estimation. ManiNLR performs manifold learning on the digital image, and then uses GRNN to map the data into the linear separable space, finally using the result to estimate the head pose. Experiments show that ManiNLR can better estimate the head pose in digital images, and has the advantages of high speed and high robustness.

**Key words:** manifold learning; head pose estimation; nonlinear regression; artificial neural network

## 0 引言

获取头部姿态能够得到人脸的朝向和眼睛的注视方向, 是人类动作估计和理解的重要组成部分。在计算机视觉中, 头部姿态估计是指对数字图像进行处理从而获得头部朝向的过程。由于摄像头的物理因素, 拍摄环境的多样化, 以及人类本身外观的多样化(如姿态、眼镜、胡子, 甚至是肤色的多样化)等因素, 使得头部姿态估计成为一个极富挑战性的工作。

对于人脸识别, 头部姿态估计能够对图像进行校正, 提供额外的信息, 是多姿态人脸检测的关键<sup>[1-3]</sup>。因此头部姿态估计在人脸识别领域具有重要地位。

同时, 头部姿态估计在增强现实、嵌入式、人机交互、人工智能等领域都有广泛的应用, 是计算机视觉领域的重要研究课题, 具有广泛的应用价值。

目前, 国内外在头部姿态估计领域中已进行了大量的研究工作, 研究的方法有以下几种<sup>[3-4]</sup>:

1) 外观模板方法: 构建模板, 对已经存在的头

收稿日期: 2011-12-05; 修回日期: 2012-03-13

基金项目: 国家自然科学基金项目(10901062); 福建省自然科学基金项目(2010J01337); 福建省高校产学研合作科技重大项目(2011H6010); 福建省科技计划重点项目(2011H0028); 福建省仿脑智能系统重点实验室开放课题(BLISSOS2010101)

第一作者简介: 范进富(1986—), 男, 华侨大学计算机科学与技术学院 2009 级模式识别与智能系统专业在读硕士研究生, 主要研究方向为图像处理及计算机视觉。E-mail: fanshill@hqu.edu.cn

部姿态进行标记。新的图像通过与模板进行比较,找出最相似的姿态。

2) 检测器阵列方法:训练一系列的检测器,并进行投票选择,可以结合机器学习的方法。

3) 非线性回归方法:建立图像与头部姿态的非线性函数。

4) 流形学习方法:用低维流形表征高维图像数据,然后用于分类或者回归。

5) 柔性模型:在每个人脸平面中对脸部特征进行非刚性建模,通过对特征或者模型参数的比较进行估计。

6) 几何方法:定位特征点,如眼睛、鼻子、嘴巴等,根据特征点的几何关系确定头部姿态。对特征点的依赖比较大,算法的优劣很大部分取决于特征点的定位。

7) 跟踪方法:根据视频中头部的运动确定头部姿态。运动中存在可提取的信息,可以简化模型。

8) 混合的方法:结合以上若干方法来克服单一方法的缺陷。

其中混合的方法能够结合不同方法的优势,以达到提高准确率的目的,典型的方法有几何方法与跟踪方法相结合,主成分分析与光流法、隐马尔可夫模型、关键帧匹配等跟踪方法的结合<sup>[4]</sup>。

但是,这些方法必然涉及大量的运算,复杂度较高。考虑到人脸数据的高维特性,采用流形学习和非线性回归相结合的方法进行头部姿态估计。通过流形学习的方法,对原始人脸图像数据进行降维,找到高维数据的内在维数,然后用人工神经网络进行非线性回归,将数据映射到线性可分的空间,最后运用线性回归进行姿态估计。实验结果表明,该方法能够在保证较好估计结果的同时,大量减少处理的时间,并具有较高的鲁棒性。

## 1 研究背景

### 1.1 流形学习

数字图像处理常常需要面对高维数据,加上庞大的数据量,就形成维度灾难。目前比较常用的两种线性降维方法是主成分分析(PCA)和线性判别分析(LDA)。二者分别通过获取数据的方差和均值的二阶统计性质来发现数据的线性结构,在分析过程中都没有考虑数据在观测空间的全局和局部性质,导致其在高度非线性分布的数据集上不能找到

完整的分布结构。

传统的线性降维方法实现简单,计算容易。但事物的内部结构通常不是线性的,内在的特征通常也不是特征的线性组合。为了揭示事物复杂的非线性结构,需要用到非线性的降维。

流形学习可用于非线性降维,是一种非监督学习方法,近年来被广泛地应用于模式识别和机器学习当中。与传统的线性降维方法相比,流形学习能够更有效地发现非线性高维数据的本质结构<sup>[5]</sup>。

设  $Y \subset \mathbf{R}^d$  是一个低维流形,  $f: Y \rightarrow \mathbf{R}^D$  是一个光滑嵌入,其中  $D > d$ 。数据集  $\{y_i\}$  是随机生成的,且经过  $f$  映射为观察空间的数据  $\{x_i = f(y_i)\}$ 。流形学习就是在给定观察样本集  $\{x_i\}$  的条件下重构  $f$  和  $\{y_i\}$ 。

如图1所示,流形学习一般包含以下3个步骤:1) 寻找每个样本点的邻域;2) 对样本点进行重构;3) 根据步骤2)的信息将数据映射到低维空间。流形学习的目的是在高维空间中根据数据分布的内在规律性,找到能够有效表达高维空间的少数独立的关键变量。从模式识别的角度看,流形学习就是特征提取的过程,关键在于高维到低维的非线性映射。

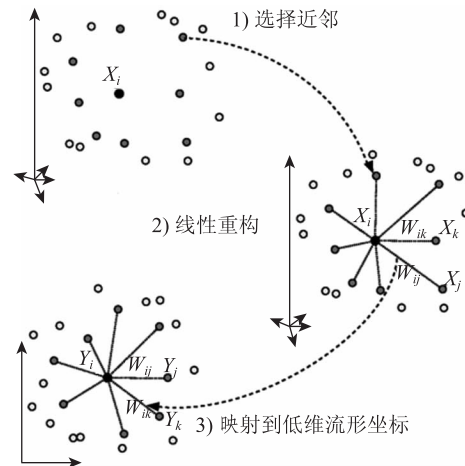


图1 流形学习的过程

Fig. 1 Process of manifold learning

局部线性嵌入(LLE)和拉普拉斯特征映射(LE)是具有代表性的非线性降维方法,局部嵌入分析(LEA)和局部保持投影(LPP)分别是LLE和LE的线性近似。流形学习优势在于:1)流形学习是非线性的,基于数据固有的几何结构,更能体现数据的本质特征;2)流形学习是非参数的方法,不需要特别的参数假设;3)流形学习最终是转化为特征值求解的问题,计算比较简单,计算效率比较高,且不会出现局部极值的问题。但流形学习面临以下问题:

1) 样本点密度变化对结果有影响; 2) 在利用邻域构造低维映射时可能存在偏差; 3) 得到的低维结构通常非线性可分, 存在分类计算复杂的问题。

### 1.2 非线性回归

在数据处理分析过程中, 通过大量的数据构建模型来表示变量之间的关系, 当模型的参数是非线性的时候, 这类模型就叫做非线性模型。非线性回归就是通过分析所观察的数据找到表示这种关系的模型, 具体就是确定函数

$$x \rightarrow y = f(x, \mathbf{c}) \quad (1)$$

式中,  $\mathbf{c} = (c_1, \dots, c_n)$  是模型的待定参数。本文 ManiNLR 方法正是利用了非线性回归将高维空间映射到线性可分的低维空间, 以提高对结果进行分类的速度。同时, 在训练模型时利用强容错性和高鲁棒性的方法, 以弥补流形学习存在的问题。

人工神经网络有很好的非线性映射能力, 常用于非线性回归。人工神经网络一般不要求建模对象的先验知识(如结构、参数、动态特征等), 根据输入、输出数据, 通过网络本身的学习能力就可以达到输入与输出的完全符合。广义回归网络(GRNN)是由 Specht 于 1991 年提出的, 在逼近能力、学习能力和学习速度上与其他神经网络相比有较大的优势, 同时具有很强的容错性和鲁棒性<sup>[6]</sup>。

## 2 ManiNLR 方法

为了对人脸图像进行分类, 将人脸图像看成基于数字图像的高维数据空间中的样本, 并采用低维的流形特征表达其头部姿态的变化<sup>[5,7]</sup>。首先如图 2 所示, 将训练集中的每个人脸图像投影到低维空间的不同点上, 使其能够有效地区分人脸的不同姿态。当对新的人脸图像进行分类时, 将其投影到低

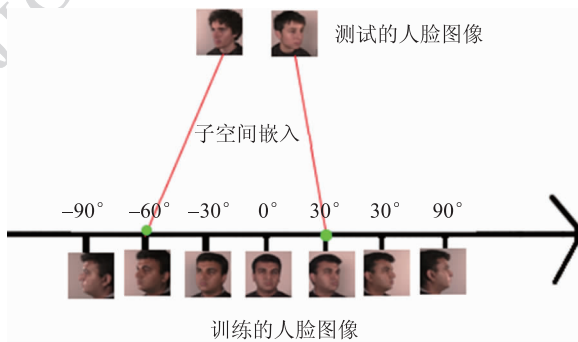


图 2 将人脸图像投影到头部姿态的低维流形中  
Fig. 2 Face image is projected to the low dimensional manifold

维空间中与其姿态相似的点上。这样, 就能找到人脸图像高维空间到低维特征空间的映射关系, 将人脸图像嵌入到低维空间中, 就可以找到表征头部姿态的本质维度, 并得到低维流形结构。

其次, 对于给定的人脸图像, 将高维人脸图像映射到低维的线性可分空间。通过非线性模型对训练数据进行学习, 能够在一定意义上最佳地逼近或拟合已知数据, 即将人脸图像映射到一个或多个头部姿态对应的低维流形, 如图 3 所示。

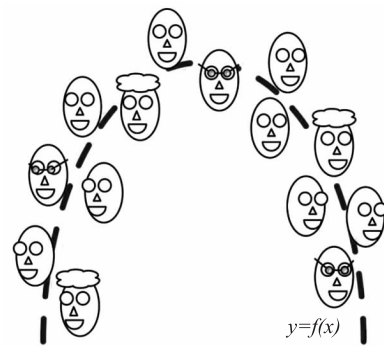


图 3 人脸图像空间到头部姿态的非线性函数  
Fig. 3 Nonlinear function of face image space to head pose

最后, 应用速度具有优势的线性回归, 对非线性映射后的数据进行线性回归, 得到头部姿态。

ManiNLR 方法描述如下:

$$P = \text{ManiNLR}(T, t, l, k, d)$$

输入:

- $T$ : 训练数据;
- $t$ : 测试数据;
- $l$ : 训练数据的标签;
- $k$ : 用来寻找邻域的参数;
- $d$ : 低维流形的维度。

步骤:

- 1) For  $T$  中的每个样本  $x_i$ ;  
找到它的  $k$  个最近邻  $x_j (j = 1, \dots, k)$ ;  
End For
- 2) 获取权值矩阵  $W$ ;
- 3) 嵌入到低维流形  $E$ ;
- 4) 对低维流形进行归一化处理;
- 5) For 初始值到输入向量之间距离的最大值  $\sigma$ ,  
根据最小平方误差准则训练 GRNN;  
End For
- 6) 通过 GRNN 将测试样本映射到低维线性可分空间, 得到低维数据  $D$ ;

7)用线性回归对头部姿态进行估计;  
输出:测试数据的头部姿态  $P$ 。

将 ManiNLR 方法用于头部姿态估计的流程如图 4 所示。



图 4 头部姿态估计的流程图

Fig. 4 Flow chart of head pose estimation

预处理阶段,首先对输入图像归一化大小,并将彩色图像按式(2)降维成灰度图像。

$$g = 0.3R + 0.59G + 0.11B \quad (2)$$

流形学习阶段,分别采用局部线性嵌入(LLE)、局部保持投影(LPP)、局部嵌入分析(LEA)对图像进行降维,同时采用传统的主成分分析(PCA)及线性判别分析(LDA),用以比较分析。

文献[8-10]分别描述了LLE、LPP、LEA算法的实现,这些算法都包含3个步骤,见图1。其中,LLE的待定参数  $k$  和  $d$  需根据具体数据进行选择,同时对样本中的噪声也很敏感,当有新的数据加入时,需要重新计算权重值,因此训练的时间相对于其他方法较长。

LPP 和 LEA 一样,目的是保留数据的局部结构信息。只是在构造权重时,LEA 是通过学习数据空间的内在邻域结构自动构造权重矩阵,而 LPP 是通过热核函数计算权重。因此,LEA 在线性子空间的求解与 LPP 相似,都是将存在低维流形的高维数据嵌入到低维环绕空间时,找到低维流形拉普拉斯算子特征函数的最优线性近似。同时,它们也是定义

在数据的全局空间中而不只是训练数据,因此不需重复计算权重矩阵<sup>[9-10]</sup>。

这些流形学习的方法都是非监督性的方法,没有考虑训练过程中可能利用的姿态信息。在训练过程中,加入已知的姿态信息有利于减少影响人脸图像表示的其他因素,如身份、光照等条件,从而改进预测效果<sup>[7,11]</sup>。训练时利用姿态信息的 LPP 及 LEA 同时拥有监督学习和非监督学习的判别能力,在考虑数据的内在几何结构的同时,可以利用已知的类别信息改进判别结果。

非线性回归阶段,采用广义回归网络的结构,如图 5 所示。

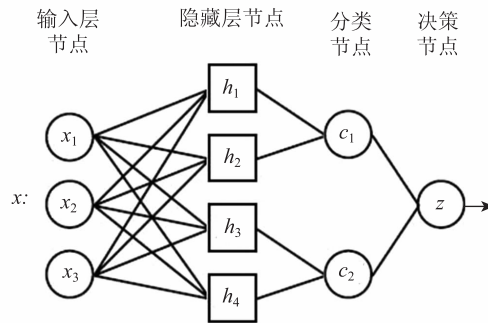


图 5 广义回归网络的结构图

Fig. 5 Structure diagram of GRNN

具体的实现过程如下:

1)输入层,其节点是原始数据,每个数据向量对应一个神经单元。

2)隐藏层也称作模式层,每个神经单元对应一类数据,并通过 C-均值方法构建聚类中心。然后计算输入与每个聚类中心的距离(通常是欧氏距离),并累加,使用预先设定的阈值  $\sigma$  作为每个神经元的径向基计算非线性激活函数,即指数函数,将结果输出给下一层。

3)分类层也称总结层,只有两个神经单元,代表分母、分子两个系数。分类层将隐藏层的输出结果与一个权重向量进行点积运算,该权重值是每个聚类中心所代表样本的个数,得出分子系数。分母系数是每个隐藏层神经单元的权重之和。将结果输出给下一层。

4)最后是决策层,将分类层分母系数相同的神经单元对应的分子系数进行累加,对结果进行判断。

由于广义回归网络只有前向反馈,不存在递归,因此能够快速学习并收敛到最优回归面,特别是对于稀疏数据,在运算速度上具有明显的优势。

对于广义回归网络的径向基函数的半径,即参数  $\sigma$ ,其越大得到的回归函数越平滑,越小得到的结

果拟合的越精确。考虑到广义回归网络训练时,寻找最优参数需要大量的计算,对广义回归网络的径向基函数的参数进行如下处理:计算输入所有向量之间的距离,取最大距离,并设置  $\sigma$  初始值为 0.1,每次迭代增加 1,直到  $\sigma$  大于输入向量之间的最大距离,计算终止。在实验过程中,发现  $\sigma$  参数的偏差在小范围内(小于 1)对神经网络的输出结果影响较小,为了简化计算过程,本实验将迭代步距设为 1。

### 3 实验结果

实验环境: Pentium (R) Dual-Core CPU E5400<sup>®</sup> 2.7 GHz, 2.69 GHz, 4 GB 内存。所采用的数据库是 Facepix 人脸数据库<sup>[12]</sup>,该数据库包含 30 个不同身份的彩色人脸图像,每个人拥有间隔 1°,范围从 -90°到 +90°的连续水平偏转姿态。每个人有 181 张人脸图像,整个数据集共有 5 430 张彩色人脸图像。其中,人脸的特征具有多样性,包含性别、发型、胡须、帽子、眼镜等特征的差异。实验对每个图像进行归一化处理,形成 32 × 32 共 1 024 维的数据。

图 6 是分别用 LLE、LEA、LPP(有监督)、LPP、PCA 及 LDA 将 10 个不同的人脸图像数据降维到 3 维后的效果图。从图 6(c)(f)中可以明显看出加入头部姿态信息的有监督流形学习方法,即有监督 LPP 和 LDA,能够很好地展示人脸头部姿态的分布,基本能够排除身份等其他因素的影响,区分头部姿态。从图 6(a)(b)中可看出 LLE 及 LEA 降维得到的流形也能较好地地区分人脸头部姿态,个别人的身份影响比较大。而 PCA 的低维流形(图 6(e)),对于人的身份区分比较明显,这样对姿态的估计可能造成干扰。无监督 LPP 的低维流形(图 6(d)),相对就比较混乱,无法准确区分头部姿态。

ManiNLR 方法的时间主要在 GRNN 的训练上,即图 4 中的非线性回归阶段,直接根据误差结果最小化设置参数时,对 10 组人的人脸图像进行读取、归一化、降维、训练的时间大于 3 h,但提出的根据输入数据设置 GRNN 参数的方法处理时间在 1.5 h 以内,能够明显地提高 GRNN 的训练速度。

GRNN 训练完成后,即图 4 中的线性判别阶段,进行预测的速度基本可以达到实时,实验中对 10 组人的人脸图像进行预测时,速度稳定在 0.5 s 左右,与直接采用欧氏距离判别<sup>[13]</sup>相比有明显的提高,测试阶段所用的时间见表 1。

表 1 测试时间

Table 1 Test time

判别方法	降维方法	降维后维度					平均时间 /s
		3	10	30	60	90	
线性	LLE	0.5	0.5	0.5	0.5	0.5	0.5
	LEA	0.5	0.5	0.5	0.5	0.5	0.5
	LPP	0.4	0.4	0.4	0.5	0.5	0.4
	LPP(有监督)	0.4	0.4	0.5	0.5	0.4	0.4
欧氏距	LEA	17.4	17.6	18.0	18.7	19.3	18.2
	LPP	17.7	17.4	18.1	18.6	19.3	18.2
	LPP(有监督)	17.4	17.5	18.0	18.2	18.3	17.9

采用 Dietterich<sup>[14]</sup>提出的 5 × 2 交叉确认的方法,将 20 个人随机分成两组,每组 10 个人,分别作为训练集和确认集。经过 5 次对折得到 10 个训练集和确认集,对折次数及实验方法的选择见文献[14]。

分别对生成的 10 组训练集和测试集采用直接线性回归、欧氏距离判定,及 ManiNLR 方法进行预测,表 2 是实验中头部姿态估计值的平均误差和方差,表中的数据是 10 组实验的平均结果。

从表 2 可以看出,使用 ManiNLR 方法能够很好地找到数据的非线性映射关系,在降到 10 维的时候,预测误差就达到 8°左右。而当降维到 30 维的时候,对于几种流形学习的方法最终都能够达到 5°的误差,方差也都保持在较小的范围,20 以内。而直接对流形学习方法降维得到的数据进行线性回归则不能较好地进行姿态估计,降到 10 维时误差在 10°以上;降到 30 维时,基本误差都在 9°左右。同时,预测值的方差比较大,都在 40 以上,反映了其预测结果不稳定,鲁棒性差。

图 4 的线性判别阶段,直接比较测试数据与训练数据的欧氏距离的方法,在较低维度的时候能够得到较好结果,但随着维度的增加,预测效果反而降低,同时该方法同样存在预测不稳定的问题,方差一直都在 20 以上,甚至达到 100 以上,容错性能和鲁棒性都较差。图 7 是头部姿态估计的结果趋势图,可以对各类方法进行比较,从中也可以明显的看出上述结果。

图 7 中,左侧是实验结果的均差,右侧是与左侧对应方法的方差。从图 7(a)可以发现,当数据降到 30 维时,预测的平均误差都在 8°下,LLE、LPP、LDA 甚至达到了 4°的误差。对应地从图 7(b)可以发现,

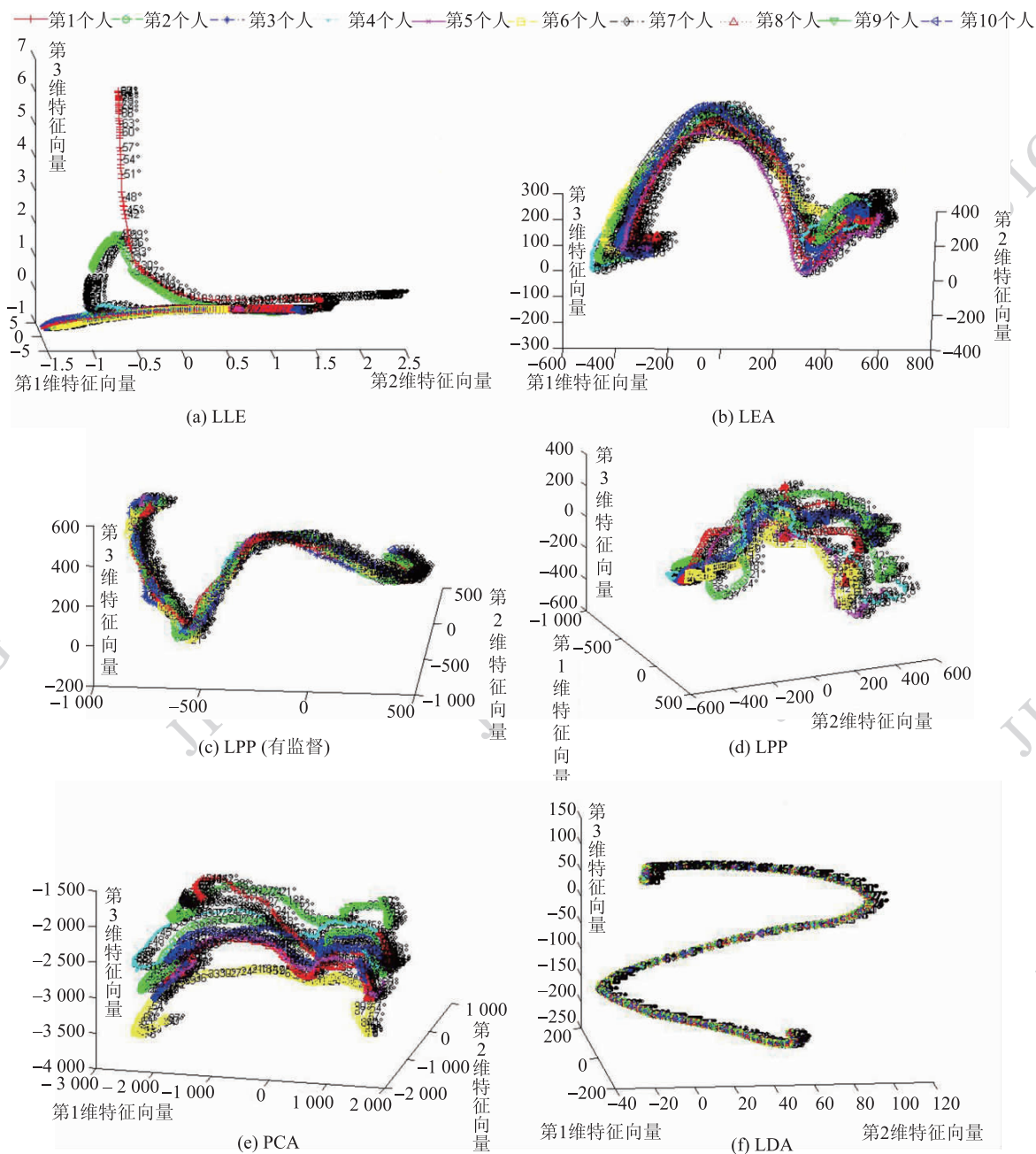


图 6 人脸图像数据降维到 3 维后的低维流形

Fig. 6 Charts of low dimensional manifolds which face image data reduce to three dimensions

在数据降到 10 维时,预测结果的方差就开始稳定,最后基本都低于 20。这表明在流形学习降维,甚至是采用传统的降维方法后,广义回归网络都能够较好的对数据进行映射,而其本身所具有的高鲁棒性也使结果比较稳定。

图 7(c)中,有监督的 LPP 方法效果明显优于无监督的 LPP 方法,无监督 LPP 误差比有监督的 LPP 高出 10°左右。而使用本文的 ManiNLR 方法与使用欧氏距离测度的方法效果相当。但从图

7(d)的比较中可以看出,ManiNLR 方法具有更高的鲁棒性。

图 7(e)(f)是直接对降维后的数据进行线性回归,可以看出效果不理想,即使降维到 90 维,平均误差还在 8°以上,而方差则在 50 左右。图 7(g)(h)是传统降维方法 PCA 和 LDA 的实验结果,PCA 降维方法的效果不好,LDA 的平均误差能够达到较为理想的水平,和流形学习的方法相当,但方差比较大,特别是在降维后的维度在 30 以内时。

表2 实验结果(误差的均值和方差)

Table 2 Experimental results(error's average value and variance)

/(°)

方法	降维后维度				
	3	10	30	60	90
LLE	(8.8,43.9)	(8.2,30.9)	(5.1,15.6)	(4.3,12.1)	(4.5,13.0)
LEA(欧氏距)	(6.0,53.9)	(4.2,22.4)	(3.6,19.0)	(3.9,34.0)	(4.6,115.8)
LEA	(20.7,176.9)	(11.7,81.6)	(9.4,62.0)	(7.9,43.4)	(7.9,52.0)
LEA(GRNN)	(10.2,56.8)	(8.1,40.8)	(7.7,30.7)	(6.2,21.3)	(6.1,20.3)
LDA(欧氏距)	(5.2,35.6)	(3.2,11.8)	(4.6,84.6)	(6.0,154.8)	(6.9,222.8)
LDA	(21.0,244.7)	(13.1,138.9)	(9.1,59.0)	(8.8,54.4)	(7.5,49.1)
LDA(GRNN)	(15.9,179.8)	(5.6,22.2)	(4.6,13.2)	(4.3,12.0)	(4.4,13.4)
PCA(欧氏距)	(5.7,37.3)	(4.2,22.5)	(3.4,20.2)	(3.2,24.2)	(3.6,44.9)
PCA	(18.9,220.5)	(12.7,90.9)	(8.4,41.1)	(8.3,41.2)	(7.0,33.6)
PCA(GRNN)	(8.6,51.4)	(7.5,39.3)	(7.1,29.2)	(5.3,17.4)	(4.8,14.4)
LPP(欧氏距,有监督)	(5.4,28.5)	(2.9,8.0)	(2.8,9.6)	(3.0,17.1)	(5.0,18.1)
LPP(有监督)	(19.7,254.1)	(14.3,95.7)	(8.8,52.9)	(10.4,61.5)	(10.8,65.2)
LPP(欧氏距)	(12.0,66.8)	(7.4,83.9)	(6.1,95.1)	(6.1,86.6)	(4.5,66.8)
LPP	(22.8,320.9)	(18.1,189.3)	(13.6,105.2)	(11.3,68.0)	(8.6,43.9)
LPP(GRNN,有监督)	(11.7,104.9)	(5.3,17.0)	(4.9,14.9)	(5.0,15.8)	(4.8,14.1)
LPP(GRNN)	(19.7,298.7)	(9.8,74.3)	(7.8,41.0)	(6.8,28.6)	(5.0,15.5)

从图7(b)与图7(d)(f)(h)的对比可以看出,采用 ManiNLR 方法具有较好的稳定性。同时,实验中是随机抽取样本,数据集中人脸特征的差别对结果也有影响,但 ManiNLR 方法仍然能保持稳定性,持有比较低的误差和波动性。

通过上面的分析比较,表明本文 ManiNLR 方法能够结合流形学习和非线性回归的优点,对人脸头部姿态进行估计,在保证精确度的同时,具有较快的速度和较高的鲁棒性。在训练过程中加入监督信息能够改进对预测结果进行改进。

## 4 结论

利用流形学习能够发现非线性数据内在结构的特性,结合非线性回归方法,将人脸图像映射

到线性可分的低维数据,最后通过速度较快的线性回归对人脸的头部姿态进行估计。实验表明,提出的 ManiNLR 方法在保证预测精度的情况下,能够显著地减少运行时间,并且具有较高的鲁棒性。

值得注意的是,在降维后低维度数提高时,将 ManiNLR 方法中的流形学习降维换成传统的线性降维方法,也能得到较为理想的效果。而试验中加入监督信息,可以改进预测结果。基于以上两点,特别是有监督的线性判别方法在精度上能够得到与 ManiNLR 方法相近的结果。

在实验过程中,可以看出添加标签信息对于人脸图像数据降维后流形的表现有很好的效果,同时对于结果的判断也有改进作用。对所提供的监督信息的使用方法也是一个可研究的问题。

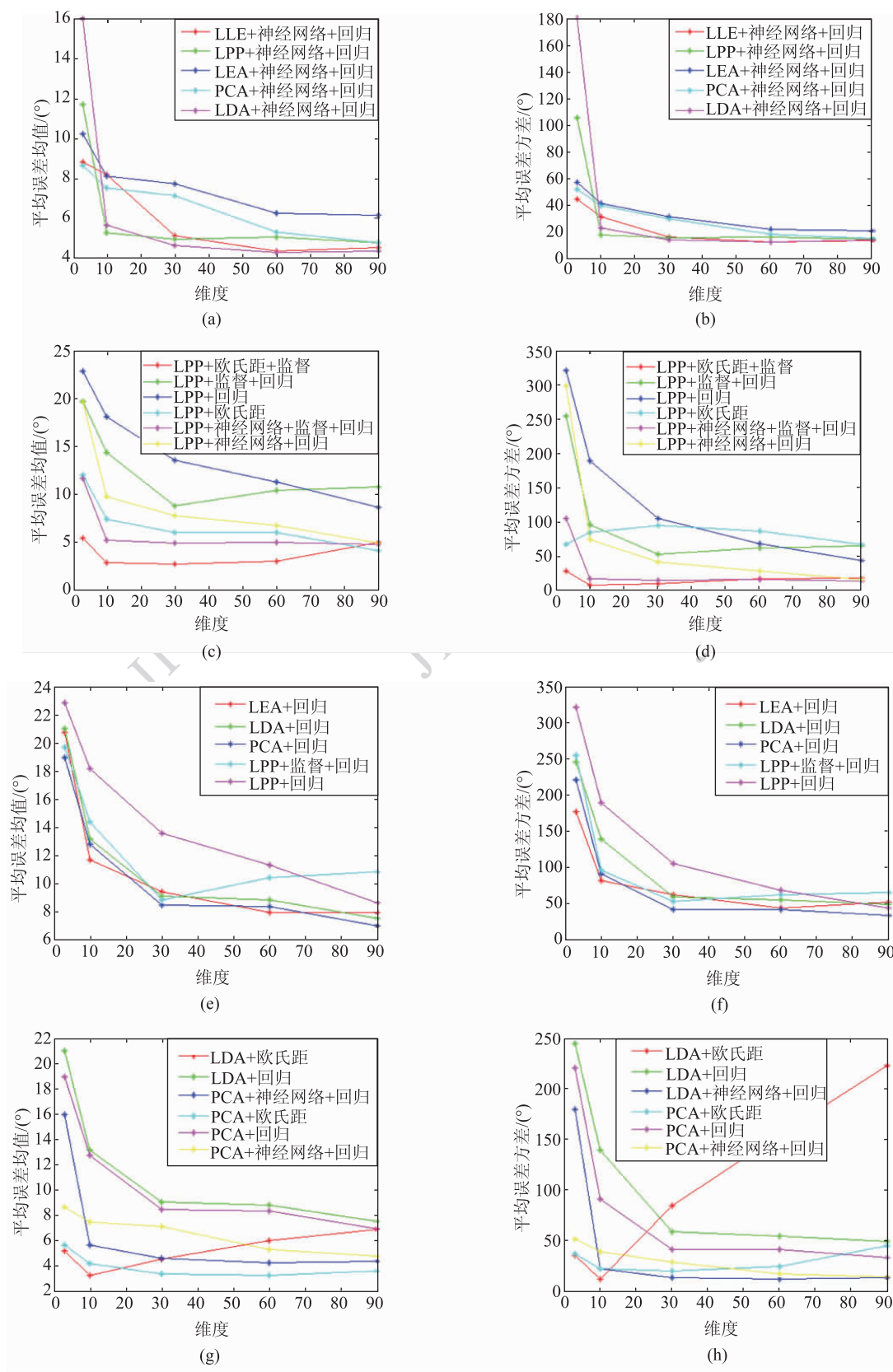


图 7 人脸头部姿态估计结果图  
Fig. 7 Results of head pose estimation

## 参考文献(References)

- [ 1 ] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [ 2 ] Yan Y, Zhang Y J. State-of-the-art on video-based face recognition [J]. Chinese Journal of Computers, 2009, 32(5): 878-886. [ 严严, 章毓晋. 基于视频的人脸识别研究进展 [J]. 计算机学报, 2009, 32(5): 878-886. ]
- [ 3 ] Wang Y H. Face Recognition-Theory, Method and Technology [M]. Beijing: Science Press, 2010: 294-306. [ 王映辉. 人脸识别——原理、方法与技术 [M]. 北京: 科学出版社, 2010: 294-306. ]
- [ 4 ] Murphy-Chutorian E, 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.
- [ 5 ] Wang J. Research on manifold learning: theories and approaches [D]. Hangzhou: ZheJiang University, 2006. [ 王靖. 流形学习的理论与方法研究 [D]. 杭州: 浙江大学, 2006. ]
- [ 6 ] Specht D. A general regression neural network [J]. IEEE Transactions on Neural Networks, 1991, 18(4): 568-576.
- [ 7 ] Balasubramanian V N, Ye J P. Biased manifold embedding: A framework for person-independent head pose estimation [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA: IEEE, 2007: 1-7.
- [ 8 ] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290(5500): 2323-2326.
- [ 9 ] He X F, Niyogi P. Locality preserving projections [J]. Matrix, 2003, 16(12): 153-160.
- [ 10 ] Fu Y, Huang T S. Locally linear embedded eigenspace analysis, IFP-TR [R]. Illinois: Univ. of Illinois at Urbana-Champaign, 2005.
- [ 11 ] Balasubramanian V, Krishna S. Person-independent head pose estimation using biased manifold embedding [J]. EURASIP Journal on Advances in Signal Processing, 2008, 8(2): 1-16.
- [ 12 ] Black J A, Gargsha M, et al. A framework for performance evaluation of face recognition algorithms [C]// Proc. SPIE. Internet Multimedia Management Systems III. Boston, USA: SPIE, 2002: 163-174.
- [ 13 ] Wang Z C, Li X Y. Head pose estimation based on manifold embedding and distance metric learning [J]. Opto-Electronic Engineering, 2008, 35(8): 62-65, 77.
- [ 14 ] Alpaydm E. Introduction to Machine Learning [M]. Beijing: Machine Press, 2009: 215-216. [ 阿培丁. 机器学习导论 [M]. 范明, 咎红英, 牛常勇, 等译; 北京: 机械工业出版社, 2009: 215-216. ]