

学校编码: 10384
学号: 23020141153184

分类号_____密级_____
UDC_____

厦 门 大 学

硕 士 学 位 论 文

流形保持的人脸图像分析方法研究

Automatic Facial Image Analysis based on Locality
Preserving

刘志凌

指导教师姓名: 金泰松 助理教授

专业名称: 计算机技术

论文提交日期: 2017年 月

论文答辩时间: 2017年 月

学位授予日期: 2017年 月

答辩委员会主席: _____

评 阅 人: _____

2017年 月

流形保持的人脸图像分析方法研究

刘志凌

指导老师

金泰松 助理教授

厦门大学

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘要

人脸图像分析是一个具有重要理论意义和应用价值的研究方向。如何提取人脸图像特征、实现算法模型并在图像分析过程中获得更加高效、鲁棒的结果，吸引了模式识别、图像处理、计算机视觉、人工智能和神经网络等多个领域的众多学者对其进行研究。本文以人脸为研究对象，重点研究基于流形保持方法的图像特征提取与识别问题。主要工作如下：

(1) 基于流形保持的人脸图像聚类方法的研究。使用含有干扰信息的人脸图像集作为实验数据集，包含人脸的光照条件变化、遮挡情况变化、面部表情变化、样本数目变化以及随机像素点噪声等干扰因素。选取十种使用邻接图保持局部性的人脸图像分析方法进行聚类分析实验，根据实验结果，探究人脸图像干扰因素对流形保持方法分析结果的影响，分析这些流形保持方法在各种干扰情况下的性能优劣。实验结果表明，使用邻接图保持流形结构的方法中，流形集成思想、 ℓ_1 图结构、弹性网络图结构等对图像干扰具有良好的适应能力。

(2) 提出了一种基于流形保持的协同表示人脸识别方法。概述了协同表示分类的研究进展和流形保持方法在人脸识别领域的现状，在此基础上提出了流形保持和协同表示方法结合的可能性。通过在协同表示方法的优化模型中加入局部结构信息正则项，实现协同表示分类和流形保持方法整合的人脸识别策略，利用增广拉格朗日乘子法将约束优化问题转换为增广拉氏方程的迭代求解问题，得到流形保持协同表示方法的算法实现。在多个人脸图像数据集上进行识别实验，实验结果表明本文提出的方法优于其他参考方法，对图像噪声和异常值具有良好的鲁棒性。

关键字：人脸图像分析、流形学习、局部保持

厦门大学博硕士学位论文摘要库

Abstract

Automatic facial image analysis has very large theoretic and practical values. How to extract facial features, implement the algorithm and obtain more efficient results through facial image analysis and recognition, is attracting a large number of researchers to study from multiple fields such as pattern recognition, computer vision, artificial intelligence and neural network. This thesis focuses on locality preserving manifold learning algorithms of facial image analysis and recognition technology. The main research contents of this thesis can be summarized as follows:

(1) Research of facial image analysis on the basis of locality preserving manifold learning in spectral clustering experiments. The cropped and normalized face images are used as the experimental dataset, which contains the interference factors such as varying illumination conditions, differing mask conditions, differing facial expressions, varying number of samples and random pixel noise. Based on the spectral clustering experimental results, the effect of interference factors on the locality preserving manifold learning algorithms is discussed. The performance of these manifolds preserving methods in a variety of interference situations is analyzed. Experiments show that some locality preserving manifold learning algorithms have superior robust, such as multiple graph regularized method, elastic net hypergraph learning, ℓ_1 -graph learning.

(2) Propose a novel collaborative representation based face recognition method. Many face recognition methods have been proposed, among them, the recently proposed collaborative representation based face recognition has attracted the attention of researchers. Many variants and extensions of collaborative representation based classification (CRC) have been presented. However, most of CRC methods do not consider data locality, which is crucial for classification task. In this thesis, a novel collaborative representation based face recognition method, LP-CRC, is proposed, which balances data locality and collaborative representation. The proposed method incorporates locality adaptor term into the robust collaborative representation based classification framework, leading to a novel unified objective function. The

Augmented Lagrange Multiplier is used to optimize the objective function. Tests on standard benchmarks demonstrate that the proposed face recognition method is superior to existing methods and robust to noise and outliers.

Key words: Facial image analysis, Locality preserving, Manifold learning

厦门大学博硕士学位论文摘要库

目录

第一章 绪论	11
1.1 研究背景及意义	11
1.2 国内外研究现状	12
1.2.1 人脸图像分析方法国内外研究现状	12
1.2.2 流形学习研究现状	16
1.3 主要研究内容	19
1.4 章节安排	20
第二章 基于流形保持的人脸图像聚类方法的研究	22
2.1 流形保持方法与聚类分析	22
2.1.1 谱聚类方法	22
2.1.2 基于 K 近邻图的流形保持方法	23
2.1.3 基于图正则化的矩阵分解方法	25
2.1.4 基于 ℓ_1 图的图像分析方法和基于弹性网络的超图学习方法	28
2.1.5 低秩表示方法	30
2.2 实验与讨论	31
2.2.1 实验环境	31
2.2.2 实验评价标准	32
2.2.3 实验设置	32
2.2.4 ORL 数据集实验结果	33
2.2.5 Extended Yale B 数据集实验结果	37
2.3 本章小结	39
第三章 基于流形保持的协同表示人脸识别方法	40
3.1 鲁棒性协同表示分类方法	40
3.2 保持流形结构的线性编码分类介绍	42
3.3 保持局部性的组稀疏分类介绍	43
3.4 LP-CRC 方法的优化模型	43

3.5 LP-CRC 方法优化问题的求解	44
3.5.1 增广拉格朗日乘子法	44
3.5.2 目标函数优化过程	45
3.6 LP-CRC 方法的分类规则	47
3.7 实验过程与讨论	48
3.7.1 数据集处理与实验准备	48
3.7.2 实验参照方法	48
3.7.3 LP-CRC 方法参数设置调整	49
3.7.4 PIE 数据集的实验结果	52
3.7.5 AR 数据集的实验结果	53
3.7.6 LFW 数据集的实验结果	55
3.8 本章小结	57
第四章 总结与展望	58
4.1 总结	58
4.2 展望	59
参考文献	60
研究生期间的科研活动和研究成果	64
致谢	65

Content

Chapter 1 Introduction.....	11
1.1 Research Background and Significance.....	11
1.2 Research Status in the Domestic and Overseas.....	12
1.2.1 Research Status in Automatic Facial Image Analysis.....	12
1.2.2 Research Status in Manifold Learning.....	16
1.3 Main Work of This Thesis	19
1.4 The Organization of This Thesis.....	20
Chapter 2 Clustering analysis based on Locality Preserving	22
2.1 Locality Preserving and Clustering analysis	22
2.1.1 Spectral Clustering	22
2.1.2 Locality Preserving Based on k-Nearest Neighbor Graphs	23
2.1.3 Graph Regularized Non-negative Matrix Factorization	25
2.1.4 ℓ_1 -Graph Learning and Elastic Net Hypergraph Learning	28
2.1.5 Low-Rank Representation	30
2.2 Experimental Results.....	31
2.2.1 Experiment Environment	31
2.2.2 Evaluation Criterion	32
2.2.3 Experiment Setting.....	32
2.2.4 Results on ORL Dataset	33
2.2.5 Results on Extended Yale B Dataset.....	37
2.3 Summary.....	39
Chapter 3 Locality Preserving Collaborative Representation	40
3.1 Robust Collaborative Representation for Classification	40
3.2 Locality Linear Coding for Classification.....	42
3.3 Locality-Constrained Group Sparse Representation	43
3.4 LP-CRC's Unified Objective Function	43
3.5 Optimization of the LP-CRC's Unified Objective Function	44

3.5.1 Augmented Lagrange Method.....	44
3.5.2 Optimization of the Unified Objective Function.....	45
3.6 LP-CRC’s Classification Rules	47
3.7 Experimental Results and Discussion	48
3.7.1 Data Sets and Experiment Environment	48
3.7.2 Comparison Baselines and Configurations.....	48
3.7.3 LP-CRC’s Parameter Setting.....	49
3.7.4 Results on PIE Dataset.....	52
3.7.5 Results on AR Dataset	53
3.7.6 Results on LFW Dataset.....	55
3.8 Summary.....	57
Chapter 4 Conclusion and Prospect.....	58
4.1 Conclusion	58
4.2 Prospect.....	59
References.....	60
Scientific Research Activities and Achievements.....	64
Acknowledgements.....	65

第一章 绪论

1.1 研究背景及意义

随着科技水平的不断发展和生活节奏的逐渐加快,人们开始需求快捷准确的身份识别方式。通过计算机识别技术鉴别身份的生物识别技术(Biometric)通常选用人脸、声音、指纹、笔迹等作为特征信息,其中人脸作为一种简单准确的特征信息受到了众多研究者的青睐。自生物识别技术的思想产生以来,对于人脸图像处理、分析、识别的需求在不断增加。

人脸识别(Face Recognition)是通过计算机技术提取人的脸部特征信息进行身份辨别的生物识别技术。在1990年之前,人脸识别的研究只是作为一般性的模式识别问题,研究大多分布在神经科学领域,主要包括:如何保证脸部在身份鉴定中的唯一性;如何提取和分析脸部剪影曲线等人脸几何结构特征;对人脸面部姿态表情的解析与应用;如何提取婴儿的人脸特征等内容,这一阶段主要探索人是如何进行人脸识别过程的。在1990年至1997年之间,自动人脸识别迎来了研究热潮,这段时间内涌现出许多像特征脸(Eigenfaces)这样极具代表性的研究方法,使得人脸识别的研究进程得以快速发展,对人脸图像的分析开始涉及模式识别、图像处理、计算机视觉、人工智能、机器学习和神经网络等研究领域。1998年至今,为了应对光照、表情变化等干扰因素,更具鲁棒性的人脸识别方法开始成为研究热点,人脸识别应用系统也得到更进一步发展。

在人脸识别应用系统的一系列技术当中,人脸图像分析是最重要的一部分,基于人脸图像的智能分析包括人脸定位,预处理,特征提取,模式分类,判决识别等。人脸图像的特征提取是影响分析结果的一个主要因素,由于人脸作为生物特征具有许多不稳定因素,对其进行分析时面临许多困难:(1)不同个体之间存在类似的人脸特征,所有人脸的大致结构都类似;(2)人脸方向的随机性,由于人脸图像是三维物理的投影,同一个体不同角度的人脸图像差异很大;(3)脸部姿态的不确定性,脸部表情的变化、发型变化和佩戴装饰等对分析结果都会造成干扰;(4)环境等其他因素的影响,图像分析时需要考虑不同的光照条件、存在损坏像素点等因素。由于人脸图像分析过程考虑的影响因素较多,因此特征表述

和特征提取方法仍缺乏统一的标准。

人脸图像数据集的样本规模是影响人脸图像分析结果的另一个主要因素，通常情况下，用于识别的参考样本越多，得到的识别结果就越准确。然而庞大的图像数目意味着需要处理维数巨大的数据，高维数据的处理不仅要消耗大量的时间和设备成本，而且对数据处理方法的要求在不断增高，传统的线性降维方法已经难以适用于分析高维的非线性数据。如何降低非线性数据的维度并保持原始数据的内在结构成为研究人脸图像分析方法研究的一个重要课题。

现实世界的的数据往往是高维非线性结构，较为常用的两类非线性降维方法是基于核函数的方法和基于流形学习的方法。核方法基于 Mercer 定理——任何半正定的函数都可以作为核函数，确定核函数，将高维空间的内积运算转化为低维空间的核函数计算，以此达到降维目的。核方法不受输入数据维度的影响，可以有效解决“维数灾难”问题，但是核函数的存在性判断和如何构造是核方法面临的主要问题，满足 Mercer 定理的函数都可以作为核函数，通常需要根据经验选择。

基于流形学习的方法是近年备受瞩目的降维方法。流形的概念是指局部具有欧几里得空间性质的拓扑空间，在拓扑学中用于描述空间结构。近年来对高维空间的微分几何和曲面整体性质的研究发展迅速，微分几何学在多个学科之间都有着广泛的应用，同黎曼几何、拓扑学等有了密切的关系，为研究新的机器学习算法提供了坚实的理论基础。流形学习理论就是在这背景下提出的，它试图将人对流形的认知规律应用到机器学习当中，强调认知过程的整体性，以微分几何学作为实现手段，使学习过程获得的结果更加有效。流形学习旨在获得高维数据的低维流形结构，并求出对应的嵌入映射，是目前较有效的非线性数据降维方法。

本文侧重研究基于流形保持的人脸图像分析方法，具体探讨这些基于流形保持的人脸图像分析方法的实际运用效果以及方法的改进。

1.2 国内外研究现状

1.2.1 人脸图像分析方法国内外研究现状

在模式识别和计算机视觉领域，稀疏表示是近年得到飞速发展和广泛应用的一类人脸图像分析方法。生命科学家 Bruno 在[1]对哺乳动物细胞稀疏性质假说

的描述中提出了稀疏编码(Sparse Coding, SC)的思想,文献[2,3]的研究指出,自然图像有着和简单细胞结构类似的稀疏性质。文献[4-6]中,使用 SC 方法将输入信号 \mathbf{y} 表示为字典元素 Φ 的线性组合,即 $\mathbf{y} = \Phi\alpha$,稀疏向量 α 是原始信号 \mathbf{y} 在字典 Φ 上的稀疏表示。用 ℓ_1 范数来保证 α 的稀疏性, SC 的优化模型形式为:

$$\min_{\alpha} \|\alpha\|_1 \text{ s.t. } \|\mathbf{y} - \Phi\alpha\|_2 \leq \varepsilon \quad (1.1)$$

其中 ε 是拟合时作为参考标准的一个非常小的正数。基于 SC 方法的运用广泛涉及数据处理的各个领域,例如图像恢复[7-10]、压缩感知[11]、形态成分分析[12]和超分辨率[13-14]等。

Wright 等人[15]在人脸图像分析方法中结合 SC 理论,提出了名为稀疏表示(Sparse Representation for image Classification, SRC)的人脸识别方法。定义 $\mathbf{X}_i \in \mathbb{R}^{m \times n}$ 为第 i 类的所有训练样本, \mathbf{X}_i 的每一列是一个训练样本, \mathbf{X} 共有 K 类,即 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K]$,对于待分类的人脸图像 $\mathbf{y} \in \mathbb{R}^m$, SRC 通过 \mathbf{X} 对 \mathbf{y} 进行稀疏表示: $\mathbf{y} \approx \mathbf{X}\alpha$ 。分析稀疏向量 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K]$,其中 α_i 是与 \mathbf{X}_i 对应的子向量, SRC 方法提出如果 \mathbf{y} 属于 i 类,则有 $\mathbf{y} \approx \mathbf{X}_i\alpha_i$ 成立,由此可得稀疏向量 α 中大多数 $\alpha_k (k \neq i)$ 的值都比较小,只有 α_i 的值较大,因此能够使用 α 中稀疏的非零值对 \mathbf{y} 进行表示,即在模型中加入 α 的 ℓ_1 范数正则化项, SRC 的算法执行过程如下:

SRC 算法

1. 使用 ℓ_2 范数对 \mathbf{X} 进行归一化处理
2. \mathbf{y} 在 \mathbf{X} 上的稀疏表示优化模型为:

$$\hat{\alpha} = \operatorname{argmin}_{\alpha} \{\|\mathbf{y} - \mathbf{X}\alpha\|_2^2 + \lambda\|\alpha\|_1\} \quad (1.2)$$

其中 $\lambda > 0$ 是平衡稀疏正则化项影响的常量参数

3. 计算残差: $\mathbf{r}_i = \|\mathbf{y} - \mathbf{X}_i\hat{\alpha}_i\|_2$ (1.3)

其中 $\hat{\alpha}_i$ 是稀疏向量中第 i 类对应的部分

4. 输出对 \mathbf{y} 的分类: $\operatorname{identity}(\mathbf{y}) = \operatorname{argmin}_i \{\mathbf{r}_i\}$
-

当测试图像数据 \mathbf{y} 存在遮挡或损坏时, Wright 等人在[15]中提出了使用单位矩阵 \mathbf{I} 作为附加字典来表示离群像素的优化模型:

$$\hat{\alpha} = \operatorname{argmin}_{\alpha} \{\|\mathbf{y} - [\mathbf{X}, \mathbf{I}][\alpha; \beta]\|_2^2 + \lambda\|[\alpha; \beta]\|_1\} \quad (1.4)$$

可以看出它与优化模型 (1.2) 结构相同, 加入单位矩阵 \mathbf{I} 以此增强对损坏数据的鲁棒性。

SRC 方法在人脸识别方面展现了优秀的性能，它激起大量基于稀疏性约束的人脸识别方法研究，例如 Gao 等人[16]提出了稀疏表示的核形式；Yang 等人[17]使用 Gabor 遮挡字典进行带遮挡人脸图像的稀疏表示学习过程；Cheng 等人[18]在优化模型中加入 ℓ_1 图保留局部性质；Qiao 等人[19]提出了基于稀疏表示的无监督学习方法；Yang 等人[20]在图像分类中将稀疏编码与线性金字塔匹配(Linear Pyramid Matching, LPM)相结合。

在 SRC 方法中人脸图像分析默认数据是正脸图像，因而产生了基于 SRC 处理非正脸和姿势发生变化的人脸图像的扩展方法。Huang 等人[21]提出了能够应对人脸图像朝向变化的方法；Wagner 等人[22]提出了应对未校准人脸图像和应对光照条件变化的方法；Peng 等人[23]提出运用低秩分解校准一批包含严重数据损坏的线性相关图像的方法；文献[24-27]提出了增强 SRC 识别效果的字典学习方法。

含有稀疏性约束的优化模型其 ℓ_1 范数最小化过程是比较耗时的，文献[28-35]提出了加速该类优化模型 ℓ_1 范数最小化过程的高速算法，文献[30]中叙述了五个具有代表性的加速算法：梯度投影 (Gradient Projection)、同伦 (Homotopy)、迭代阈值收缩 (Iterative Shrinkage-Thresholding)、近似梯度 (Proximal Gradient) 和增广拉格朗日乘法 (Augmented Lagrange Multiplier, ALM)。文献[30]对各算法进行了比较，在人脸识别的研究应用中，同伦[34]，ALM 和 ℓ_1 最小平方法[35]能够兼顾较好的准确性和运算速度。

虽然 SRC 方法在人脸识别领域已得到广泛的研究应用，但是其获得优秀性能的原理尚未被完全阐明。在人脸图像分析研究中时常涉及稀疏性约束的 ℓ_1 范数，这些方法旨在改进稀疏向量 α 的 ℓ_1 正则化项，例如 Liu 等人[36]在优化模型中增加稀疏向量 α 的非负约束；Gao 等人[37]提出了在 SC 优化模型中加入 α 的拉普拉斯正则化项；Yuan 和 Yan[38]注重保持训练图像中存在的组结构关系，即属于同一分类的样本数据之间的关系，通过在稀疏表示优化模型中使用联合 ℓ_1 和 ℓ_2 范数的组 Lasso 约束来实现；Elhamifar 等人[39]提出了提高算法鲁棒性的结构化稀疏表示方法。

上述方法都只注重改进稀疏向量 α 的稀疏性约束 ℓ_1 正则化项，而忽略了测试样本 y 与所有训练样本 $[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K]$ 之间的联系。文献[40]的研究质疑了稀疏性

约束在优化模型中的主导作用, Shi 等人[41]对 SRC 方法的 ℓ_1 正则化项进行讨论研究, 证明了 α 的稀疏性约束 ℓ_1 正则化项不是 SRC 方法性能优越的关键因素。Rigamonti 等人[42]比较了 SRC 方法和卷积滤波方法(Convolution Filter)的区别, 得出两个优化模型的学习过程能得到相似结果的结论, 进一步证明稀疏性约束的 ℓ_1 正则化项在优化模型中是能被其他正则化项替代的。

Zhang 等人[43]在人脸图像分析实验中对比了使用 ℓ_2 范数正则化项的协同表示和使用 ℓ_1 范数正则化项的 SRC 方法, 实验结果表明协同表示的性能要优于 SRC, 在优化模型中使用稀疏向量 α 的 ℓ_2 正则化项代替 ℓ_1 正则化项还能提高算法的运行速度, 以此为基础, Zhang 等人在后续研究[44]中详细讨论了 SRC 方法的协同表示性质, 进而提出了使用 ℓ_2 范数正则化项的协同表示分类方法 (Collaborative Representatin based Classification, CRC), 该方法的算法复杂度比 SRC 小得多, 从实验结果能看出其出色的计算效率和值得借鉴的性能表现。考虑到增强 CRC 方法处理异常值和噪声的鲁棒性, Zhang 等人在[44]中提出了注重处理遮挡和损坏图像数据的 CRC 方法的鲁棒性形式 (R-CRC)。

CRC 方法的优越性能引起了大量基于协同表示的人脸图像分析方法的改进研究, 考虑到 CRC 方法无法处理高度非线性分布的数据, Wei 等人[45]提出了作为 CRC 方法非线性扩展的数据局部性约束的 CRC 核方法(KLCRC)。Timofte 等人[46-47]提出了加权协同表示分类方法 (WCRC)。

人脸图像分析方法需要研究在无约束的环境下对人脸图像识别的鲁棒性[48-50], 文献[50]中给出了对此类图像识别效果的综合分析。对于无约束的人脸图像, 分析方法面临的主要问题是人脸姿势的不确定性, 人脸的面部朝向及表情变化都会影响分析结果。Ding 等人[48]提出了能够处理面部朝向偏转在 90° 以内的人脸识别方法。Ding 等人[49]在此基础上提出了多方位多级双交叉模式 (MDML-DCPs) 的人脸图像识别方法, 该方法在处理人脸姿势和表情变化的方面有着良好的计算效率和鲁棒性。近年来出现了许多人脸图像分析方法鲁棒性技术的研究成果, 例如 Xu 等人[51]提出的多视角完整空间学习 (MISL), 其在损失函数中使用柯西估计求解残差, 以此增大对异常值的鲁棒性。He 等人[52]提出了一种计算人脸图像鲁棒性稀疏表示的分析方法。

不仅考虑稀疏性, 许多研究者开始倾向于研究保持局部性质的人脸图像分析

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库