

分类号 TP3

密级 _____

UDC _____

编号 _____

中国科学院研究生院

博士学位论文

人脸识别中的光照问题研究

卿来云

指导教师 高文 教授

中国科学院研究生院

申请学位级别 工学博士 学科专业名称 计算机软件与理论

论文提交日期 2005年03月 论文答辩日期 2005年06月

培养单位 中国科学院研究生院

学位授予单位 中国科学院研究生院

答辩委员会主席 _____

声 明

我声明本论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，本论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：

日期：

论文版权使用授权书

本人授权中国科学院研究生院可以保留并向国家有关部门或机构送交本论文的复印件和电子文档，允许本论文被查阅和借阅，可以将本论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编本论文。

（保密论文在解密后适用本授权书。）

作者签名：

导师签名：

日期：

摘 要

由于人脸识别研究有着重要的理论价值和应用价值，经过近四十年的发展，人脸识别技术已经取得了长足进步。当今的人脸识别系统虽然在用户配合的理想情况下取得了令人满意的结果，但如在更大规模人脸数据库、摄像环境不可控、用户不配合的情况下使用，其识别性能急剧下降。因此，现有的人脸识别系统还尚未成熟，尤其需要解决非理想摄像条件（光照变化、摄像设备差异）和用户不配合（视角变化、表情变化、饰物乃至化妆）等问题。

本文主要考虑人脸识别中的光照变化问题，同时兼顾姿态变化。本文从人脸图像的成像模型出发，在估计输入人脸图像的成像参数后，得到与光照无关的特征用于人脸识别，以提高人脸识别系统在光照变化下的性能。简而言之，本文的主要研究成果包括以下几个部分：

(1) 任意光照下人脸图像的光照估计。基于球面谐波理论，朗伯反射对光照而言是一个低通滤波器。因此本文在频域中对成像模型进行分析，对输入人脸图像进行光照估计，即估计光照的低频分量的系数。将人脸表面近似为朗伯反射，融合人脸类的先验知识，本文可对任意输入人脸图像进行光照估计。这种光照表示可表示任何类型的光照环境，包括室外光照环境。同时，由于表示光照的参数很少，本文的光照估计对输入人脸图像的分辨率及配准精度不敏感。

(2) 基于绘制的人脸图像光照补偿。基于绘制的光照补偿是在光照估计的基础上，用光照商图像技术将其绘制到预先定义的标准光照，从而使所有的人脸图像在相同的光照条件下进行比较。本文讨论了两种标准光照：虚拟均匀光照及环境光+正面点光源。其中均匀光照下的图像中只包含人脸的纹理信息而不包含形状信息；环境光+点光源下的图像中既包含人脸的纹理信息，又包含了形状信息。基于绘制的光照补偿依赖于输入人脸的形状，即依赖于特征点的定位，因此对配准比较敏感。在精确配准的情况下，基于绘制的光照补偿对光照变化下的人脸识别能带来很大提高。

(3) 基于差图像的人脸图像光照补偿。差图像为原图像与平均人脸在输入光照下的图像的差。由于二者光照相同，这样差图像中光照变化的影响得到了减弱，同时又保留了用于识别的输入人脸与平均人脸的差别。基于差图像的光照补偿不需要输入人脸的形状，因此其对配准精度不敏感，更适合于实用系统。

(4) 基于图像绘制的方法用于扩充人脸检测中的训练样本。基于统计学习的人脸检测方法的性能依赖于训练样本的分布。通常我们很难收集到足够多的符合要求的训练样本，而基于绘制的方法可以合成任意光照下的虚拟样本。通过合理地配置目标光照，基于扩充后的训练集的人脸检测器的性能得到了较大提高。

(5) 基于特征谐波图像模型的光照和姿态不变的人脸识别。特征谐波图像模型利用

人脸类的谐波基图像的先验知识，从一幅人脸图像重构其球面谐波基图像。谐波基图像包含了人脸的本质特征：形状和反射率，因此我们可以用人脸的谐波基图像用来表示人脸的不变特征，用于不同光照和姿态下的人脸识别。假设人脸的谐波基图像的分布为高斯分布，本部分工作对所有人脸的谐波基图像的分布空间用 PCA 模型表示。结合该先验限制，在光照估计的基础上，可从一幅图像恢复该人脸的谐波基图像。同时由于 PCA 系数的数目相对图像像素数目很少，根据姿态变化的人脸图像中可见部分仍可恢复出该人脸的谐波基图像的 PCA 系数，因此该模型可兼顾姿态变化。

关键词：人脸识别；成像模型；球面谐波模型；光照变化；光照估计、光照补偿、姿态变化、PCA

Study on the Illumination Variations in Face Recognition

Laiyun Qing (Computer Science)

Directed by Prof. Wen Gao

For its great value in theory and application, face recognition has made great progress after two decades of development. The state-of-the-art of the face recognition systems achieves good results under the well-controlled environments. However, if tested on the larger database, or the non-controlled imaging environments, or the incooperative users, the performances of the face recognition systems decrease sharply. Therefore face recognition is far away from mature. It needs of more research on the problems of non-ideal imaging environments (illumination variations and the different camera parameters) and the incooperation of the users (pose variations, expression variations, accessory variations and dressing).

The dissertation discusses the illumination variations as well as the pose variations in face recognition. Based on the imaging formation, we get the illumination insensitive features after estimating the imaging parameters. These features are then used in face recognition to improve the performances of the face recognition systems when the images are collected under varying lighting. The contributions of the dissertation are as follows:

(1) Illumination estimation. Based on the spherical harmonics model, Lambertian surface acts as a low-pass filter for illumination. Analyzing the imaging formation in space-frequency domain, we estimate the illumination affecting the appearance of the Lambertian surface, i.e., the spherical harmonic coefficients of the low frequencies of the illumination. Approximating the human face as a Lambertian surface, we estimate the illumination of the face image exploiting the human face class prior knowledge. This representation of illumination can represent any types of illumination environments, the outdoor illumination environment or the indoor environments. And as the parameters of the illumination are few, the proposed illumination estimation is insensitive to the alignment and the resolution of the input image.

(2) Lighting normalization based on re-lighting. An illumination ratio technique is proposed to re-render the input face image to a pre-defined canonical illumination, based on the illumination estimation. Therefore all the images are compared under the same illumination in face recognition. Two canonical illuminations are discussed in the dissertation: the virtual uniform illumination and a front flash with ambient light, in which the former encodes only the texture information, while the latter encodes both the texture and the shape information of the input face. The re-lighting based lighting normalization depends on the shape of the face. Therefore it depends on the alignment of the feature points. If the face image is perfect aligned, the re-lighting strategy can improve the performance of the face recognition greatly under varying lighting.

(3) Lighting normalization based on the difference image model. The difference image is the difference between the input face image and the average face under the same illumination. In the difference image, the effect of the illumination is weakened, while the differences between faces are still encoded. The shape of the input face is not used in lighting normalization so that the difference image is insensitive to the alignment and it suits real applications.

(4) Re-lighting virtual samples for face detection. The performances of the prevalent face detections based on data-driven learning-based techniques depend on the distribution of the training samples. In general, it is difficult to collect enough samples covering the total distribution. A re-lighting technique is proposed to re-render more samples under varying illuminations to enrich the training set. By configuring the target illuminations properly, the performance of the face detector based on the enriched training set has improved.

(5) EigenHarmonics model for face recognition under varying illuminations and poses. The EigenHarmonics exploiting the statistics of the harmonics of the human face is proposed to reconstruct the harmonic images from one input face image. Both the intrinsic shape and texture are contained in the harmonic images so that the harmonic images are used as the invariant feature for face recognition. The statistics of the harmonic images of the human faces are approximated as a Gaussian distribution and modeled with a PCA (EigenHarmonics) model. Based on illumination estimation, the harmonic images of the input face are reconstructed with this prior restriction. As the parameters of the EigenHarmonics are much fewer than the pixels in the face image, the reconstruction is steady using only the available parts in the face images under varying poses.

Keywords: face recognition, imaging formation, spherical harmonic model, illumination variation, illumination estimation, lighting normalization, pose variation, PCA.

目 录

图目录.....	IX
表目录.....	XI
第一章 绪论	1
1.1 人脸识别的研究意义	2
1.2 人脸识别的研究成果	3
1.3 人脸识别的主要挑战	4
1.4 光照问题的提出及本文的主要贡献.....	6
1.5 论文的组织.....	7
第二章 人脸识别中光照问题综述	9
2.1 光照与人脸图像.....	9
2.1.1 光照变化下的人脸图像的分布.....	9
2.1.2 图像成像模型	10
2.2 人脸识别中光照问题研究概述.....	13
2.2.1 基于图像处理技术的方法	14
2.2.2 基于子空间的方法.....	16
2.2.3 基于成像模型的方法	18
2.3 光照处理方法总结	21
第三章 基于球面谐波理论的人脸图像光照估计	25
3.1 光照估计综述	25
3.2 球面谐波理论	27
3.2.1 球面谐波函数	27
3.2.2 入射光照的球面谐波展开.....	29
3.2.3 朗伯反射的 BRDF 的球面谐波展开.....	29
3.2.4 反射方程的频域表示	31
3.3 光照估计原理	32
3.3.1 球面谐波基图像	32
3.3.2 光照估计	33
3.4 输入人脸的球面谐波基图像	33
3.4.1 输入人脸的形状	33

3.4.2 输入人脸的纹理	35
3.5 光照估计的实验结果分析	35
3.5.1 光照估计视觉结果	37
3.5.2 光照估计对配准精度和图像分辨率的敏感性	38
3.5.3 光照估计结果数值分析	40
3.6 光照估计小结	43
第四章 基于球面谐波理论的人脸图像光照补偿	45
4.1 基于绘制的光照补偿	45
4.1.1 基于绘制的光照补偿的原理	45
4.1.2 标准光照	46
4.1.3 后处理	49
4.2 差图像光照补偿	50
4.3 光照补偿对配准精度的敏感性	52
4.3.1 数据的选择	53
4.3.2 人脸识别实验结果	54
4.4 与其他一些经典的光照处理方法的比较	56
4.4.1 待比较的光照处理算法	57
4.4.2 对比实验结果	59
4.4.3 对比实验结果分析	61
4.5 光照绘制在人脸检测中的应用	62
4.5.1 目标光照的配置	62
4.5.2 人脸检测结果	64
4.6 本章小结	64
第五章 特征谐波图像模型：不同光照与姿态下的人脸识别	67
5.1 动机	67
5.1.1 为什么用谐波基图像 (Harmonics)	67
5.1.2 为什么用 PCA 模型 (Eigen)	68
5.2 基于球面谐波基图像的人脸识别系统框架	68
5.3 特征谐波基图像模型的构造	69
5.4 根据正面人脸图像注册人脸	72
5.4.1 人脸图像与模型配准	72
5.4.2 光照估计	75
5.4.3 谐波基图像重构	78
5.5 从任意姿态人脸图像重构谐波基图像	79

5.5.1 姿态估计	80
5.5.2 谐波基图像重构	82
5.6 人脸识别	84
5.6.1 只有光照变化的实验结果	84
5.6.2 姿态和光照同时变化的实验结果	85
5.7 总结和讨论.....	87
第六章 结束语	89
6.1 本文的主要工作.....	89
6.2 将来的工作.....	91
参考文献	93
致 谢.....	i
作者简历	iii

图目录

图 1-1: FRVT2002 对光照变化部分的测试结果 (摘自[Phillips03])。	6
图 2-1: 不同人脸在不同光照下的图像的分布。	10
图 2-2: 光照和人脸图像。	11
图 2-3: 反射方程的表示。	12
图 2-4: 局部坐标和全局坐标。	13
图 3-1: 朗伯反射函数的频域表示。	31
图 3-2: 人脸球面谐波基图像示例。	32
图 3-3: 平均人脸形状及输入人脸的形状。	34
图 3-4: CMU-PIE 人脸数据库中在 3D 图中显示人脸、13 个相机和 21 个闪光灯的 xyz 坐标。	36
图 3-5: CMU-PIE 人脸数据库中正面姿态不同光照下图像示例。	36
图 3-6: CMU-PIE 人脸数据库上光照估计的结果。	37
图 3-7: 光照的可加性。	38
图 3-8: 光照估计对配准精度的敏感性。	39
图 3-9: 光照估计对图像分辨率的敏感性。	39
图 3-10: CMU-PIE 人脸数据库中各种光照下的光照系数的均值。	41
图 3-11: CMU-PIE 人脸数据库中各种光照下的光照系数的方差。	42
图 3-12: CMU-PIE 人脸数据库中各种光照下光照系数的分类结果。	43
图 4-1: 基于绘制的光照补偿结果。	48
图 4-2: 更多的重新绘制的例子, 将光照条件相对较好的图像绘制到任意光照。	49
图 4-3: 坏点检测。	50
图 4-4: 差图像光照补偿实验结果。	52
图 4-5: Yale B 人脸数据库中光照变化的情况。	53

图 4-6: 人脸图像区域裁减。	54
图 4-7: BVC2004 人脸数据库中一个人脸图像示例。	57
图 4-8: Gamma 灰度校正算法中用的标准图像。	58
图 4-9: 各种光照预处理算法的直观效果图。	58
图 4-10: 人脸检测中目标光照的配置。	62
图 4-11: 人脸检测中的目标光照。	63
图 4-12: 人脸检测中基于图像绘制生成不同光照下的虚拟样本。	63
图 4-13: 不同训练集上基于 SVM 的人脸检测器的 ROC 曲线图。	64
图 5-1: 基于 EigenHarmonics 模型的人脸识别系统框架。	69
图 5-2: 平均人脸在参考空间 (u, v) 中的表示。	70
图 5-3: 特征谐波基图像。	72
图 5-4: 输入人脸及其与模型对齐后的图像。	75
图 5-5: CMU-PIE 人脸数据库上光照估计的结果。	76
图 5-6: CMU-PIE 人脸数据库中各种光照下光照系数的均值。	77
图 5-7: CMU-PIE 人脸数据库中各种光照下光照系数的方差。	78
图 5-8: 根据不同输入人脸在不同光照下的图像重构的球面谐波基图像。	79
图 5-9: 不同姿态下输入人脸图像及其与模型对齐后的图像。	82
图 5-10: 同一人脸在相同光照、不同姿态下图像的光照估计的结果。	83
图 5-11: 根据相同人脸在不同姿态下的图像重构的球面谐波基图像。	83

表目录

表 1-1: 一些主要商业人脸识别系统。	4
表 2-1: 几种典型的基于模型的方法所用的先验知识。	21
表 3-1: 光照系数的物理意义。	40
表 3-2: 几个典型的光照估计结果及其说明。	40
表 4-1: 两种典型的标准光照。	48
表 4-2: 不同光照补偿方法及不同配准精度在 CMU-PIE 人脸数据库上的识别率比较。 .	55
表 4-3: 不同光照补偿方法和不同配准精度在 Yale B 人脸数据库上的识别率比较。	56
表 4-4: 待比较的光照处理算法。	58
表 4-5: 各种方法在 Yale B 人脸数据库上的实验结果。	59
表 4-6: 各种方法在 CMU-PIE 人脸数据库上的实验结果。	60
表 4-7: 各种方法在 BVC2004 人脸数据库上的实验结果。	60
表 5-1: CMU-PIE 人脸数据库上不同光照下的人脸识别结果。	85
表 5-2: CMU-PIE 人脸数据库上不同光照和不同姿态下的人脸识别结果。	86

第一章 绪论

很难想象，没有光，我们这个世界将会怎样？正是因为有了光，才有了这个美妙多彩的世界。通过图像，光和计算机结合，产生了两大研究方向：计算机视觉和计算机图形学。

计算机视觉（Computer Vision）旨在通过电子感知和理解图像，重现人类视觉。我们生活在 3D 世界中，而计算机能得到的大多是 2D 图像。这种低维投影减少了信息量，使得计算机视觉很困难。计算机视觉现在的一个发展趋势是由 2D 视觉向 3D 视觉转变，首先从 2D 图像或图像序列中恢复出 3D 信息，再对 3D 场景进行理解。光使得这个世界美妙多彩，但这也使得计算机视觉问题变得更困难。因为不同光照下相同物体的图像差别很大，从这些变化很大的图像中恢复物体不变的 3D 信息非常困难。

计算机图形学（Computer Graphics）的主要任务是绘制照片级真实感（Photorealism）的图形。计算机图形学利用输入物体的几何形状、反射特性和光线，通过对物理上的光线传播进行仿真，从而生成以假乱真的图形，以模拟这个多彩的世界。这方面的进展主要为物理上正确的光线传播仿真和正确的输入，包括物体的几何形状、反射特性，以及人射光场。计算机图形学的这些输入也被称为视觉模型（Visual Model）。虽然计算机图形学在过去几十年中取得了很大进展，但至今离照片级真实感图形还很远。这主要是因为：(1)大多数光线传播仿真算法都很慢。对于交互式的虚拟绘制，很少能对自然光照或物理正确的反射函数进行绘制。(2)获得正确的输入很困难。一些娱乐应用经常要求繁重的手工劳动以获得这些参数。一个较好的方法是通过真实照片获得场景的高质量的光照输入和物体材质输入。

从上述分析可以看出，计算机图形学和计算机视觉都是研究光线和物体之间的相互作用的计算。**视觉模型**是计算机图形学和计算机 3D 视觉的共同基础。以此为基础，现在计算机视觉和计算机图形学方法相结合的研究越来越多，表现为计算机视觉中很多问题通过基于合成分析（analysis and synthesis）的方法来解决，而计算机图形的很多问题也是通过反绘制（inverse rendering）来实现，即从很多输入照片中估计物体的形状、反射属性和光照等，以获得对建模和识别都适用的物体的本质特征和光照参数。

本文主要考虑一个经典的计算机视觉问题：人脸识别。更确切的说，我们主要考虑光照变化下的人脸图像识别。光照变化是影响现有人脸识别系统性能的关键因素之一。针对这一问题，本文采用基于成像模型的方法，即从人脸图像的成像模型出发，估计输入人脸图像的光照条件和人脸的本质特征。在人脸识别时，我们使用人脸与光照无关或对光照变化不敏感的特征（人脸的本质特征或根据其本质特征绘制的标准光照下的图像），以提高人脸识别系统在光照变化下的性能。

本章是如下组织的。首先介绍人脸识别的研究意义和一些已有的成果，接着我们分

析人脸识别问题中的主要挑战。光照变化是实用人脸识别系统的主要挑战之一。最后给出本文在解决光照问题方面的主要贡献及论文的组织结构。

1.1 人脸识别的研究意义

人脸识别是利用计算机分析人脸图像，进而从中提取出有效的识别信息，自动鉴别用户身份的一门技术。作为一种最直接、最自然、最容易被人接受的生物特征识别技术，人脸识别研究迄今已有 40 多年的历史，是模式识别和计算机视觉中一个非常活跃的研究热点。尤其是从二十世纪九十年代以来，人脸识别取得了飞跃式的发展，国内外出现了很多人脸识别的实验室系统和商用系统。这些系统在用户配合的理想情况下取得了令人满意的结果。而且国内外各知名大学、研究所、IT 公司等都得到了大量的项目支持，以继续推动人脸识别技术的发展。

人脸识别问题之所以得到重视，是因为其具有重要的研究意义，突出表现在其对学科发展的贡献和广阔的潜在应用前景两个方面。

1. 人脸识别研究可促进多门相关学科的发展

人脸识别研究涉及模式识别、计算机视觉、计算机图形学、人工智能、图像处理、心理学、生理学和认知科学等，与人机交互领域和基于其它生物特征的身份识别方法都有着密切的联系。人脸识别作为一个图像模式识别的经典问题，为图像分析与理解、计算机视觉、人工智能、人机交互、认知科学、神经计算、心理学、生理学等多个学科提供了一个很好的具体问题，有利于构建这些学科领域的基础实验平台，用于尝试新方法、验证新理论、解释新现象。人脸识别问题的深入研究和最终解决，可以极大地促进这些学科的成熟和发展。如作为一个模式识别的问题，人脸识别被认为是最具挑战性的问题之一：模式种类数目庞大，不同类别模式差别微妙，这一点由于数据采集过程的噪声、成像设备的精度、外界条件的变化以及数据缺损而显得更加突出。再如，作为一个计算机视觉问题，如何融合人脸的一般性先验形状信息来准确地恢复特定人脸的 3D 结构也是一个非常有价值的研究问题。从计算机图形学的角度，研究人脸在光源照射下的成像过程，可以指导人脸图像 3D 形状分析、光照模式分析及生成虚拟光照和虚拟视图下的图形。同时人脸识别也是智能人机接口领域的重要研究内容之一。人脸识别本质上是要赋予计算机区分不同人类个体的能力，也就是“看”的能力，这恰是智能人机接口研究需要解决的问题之一，也是机器智能的重要表现。

2. 人脸识别技术作为一种生物特征识别技术有着广阔的潜在市场前景

身份识别是人类社会日常生活中的基本活动之一。目前我们大多数情况下仍然依赖于传统的身份验证手段(如钥匙、身份证、密码和口令等)来完成身份识别过程。这些传统方式不方便、不安全、不可靠的缺点不言而喻：证件、钥匙携带不便、可能会丢失；证件可以被伪造；密码可能会遗忘或被蓄意窃取。这些缺点使得它们已经越来越不能适应社会发展的需求，我们亟需更方便、更可靠、更安全的身验证方式。生物特征识别

技术 (Biometrics) 则是利用人类本身所拥有的、能够唯一标识其身份的生理特征或者行为特征进行身份验证的技术。与传统的身份识别方法相比, 生物特征识别技术拥有更方便、更高级别的安全性: 如无需额外携带信物、很难在活体上伪造生物特征等。可以预计, 在不远的将来, 包括人脸识别技术在内的各种生物特征识别技术, 作为安全级别更高的新兴技术, 将革命性地在社会应用的各个层面逐步取代传统的基于信物或口令的身份鉴别方法, 在自动身份鉴别的安全领域中占据主导地位。

作为一种典型生物特征识别技术, 与其它生物特征识别技术 (如指纹、虹膜、声音) 相比, 人脸识别在可用性方面具有以下独到的技术优势, 可应用于公共安全、信息安全、访问控制、视觉监控、自然人机交互和动画娱乐等方面:

- (1) 可以隐蔽操作, 不需要用户的主动配合, 尤其适用于安全监控;
- (2) 非接触式采集, 侵犯性较小, 容易被接受;
- (3) 采集方便, 采集设备成本低, 潜在的数据资源丰富;
- (4) 符合人类的识别习惯, 可机读亦可人读。

因此, 人脸识别研究不仅对推动模式识别、计算机视觉等各个学科领域的发展具有重要的理论意义, 而且还具有重大的市场价值和社会意义。

1.2 人脸识别的研究成果

正是人脸识别的重要性及其优势, 人脸识别技术越来越多地引起了人们的关注并得到了长足的发展。值得一提的是, 最近由 Anil K. Jain 和 Stan Z. Li 合编了 <Handbook of Face Recognition> [LiSZ04a] 一书, 书中对人脸分析牵涉的方方面面作了详细介绍, 也可以看作是人脸识别的最新、最全的综述。简而言之, 人脸识别的研究成果主要表现在:

(1) 人脸识别的基础资源如人脸数据库已经非常丰富, 出现了很多专业的大规模人脸资源库。这些人脸数据库一般都包含光照、姿态、表情、年龄等变化。一些典型的人脸图像数据库请参见 <Handbook of Face Recognition> [LiSZ04a] 的第 13 章中的相关介绍。

(2) 很多研究机构在从事人脸识别相关的研究。国际上比较著名的研究机构有美国的卡内基梅隆大学 (CMU) 的机器人研究所、美国麻省理工大学 (MIT) 的媒体实验室和人工智能实验室, 法国的 INRIA 研究院、日本的 ATR 研究所等。我国的人脸分析研究现也进入了蓬勃发展的阶段, 北京工业大学、东南大学、哈尔滨工业大学、复旦大学、南京大学、清华大学、西安电子科技大学、西北大学、浙江大学、中山大学、微软亚洲研究院、中科院自动化所、中科院计算所等高校和研究机构都有一些研究团队在进行人脸分析方面的研究。

(3) 很多公司也开始开发人脸识别方面的产品, 并出现了很多商用系统。表 1-1 列出了一些知名公司及其商用系统。

(4) 已有一些第三方独立组织的人脸识别方面的评测和竞赛, 进一步推动了人脸识别的

发展：如三次 FERET 测试 [Phillips96][Phillips98a][Phillips00]、两次 FRVT 测试 [Blackburn01][Phillips03]、ICPR2004 会议的 BANCA 竞赛 [Messer04] 等。国内从 2004 年开始也有一些单位组织人脸识别评测：公安部组织的人脸识别评测、863 组织的人脸识别评测和由生物特征认证与测评中心联合中国信息安全产品测评认证中心举办的第一届中国生物特征识别竞赛 BVC2004。通过这些评测，国内的人脸识别也将会有更好的发展。

(5) 涌现了很多涵盖人脸分析的学术会议和期刊，如国际计算机视觉会议 (ICCV)、欧洲计算机视觉会议 (ECCV)、计算机视觉与模式识别会议 (CVPR)、计算机图形学大会 (Siggraph) 等都设有人脸分析专题。还有一些专门的国际会议，如：国际人脸与姿态识别会议 (FGR)、基于音视频的身份认证会议 (AVBPA) 等。国内的生物特征识别会议迄今也已举办五届，成为国内同行交流包括人脸识别在内的生物特征识别方面的经验和成果的平台。在国际权威杂志上如：国际计算机视觉杂志 (IJCV)、IEEE 模式分析与机器智能 (PAMI)、模式识别 (PR) 等，每年也有许多关于人脸分析的最新研究成果出现。

表 1-1：一些主要商业人脸识别系统。

Table 1-1. Some commercial face recognition systems.

系统名称	提供商	简要说明
FaceEXPLORER	Viisage http://www.viisage.com	基于 MIT 的特征脸技术，后又融合了 ICA 技术，和弹性图匹配技术
FaceIt	Identix(前身为 Visionics) http://www.identix.com	基于 Rockfeller 大学开发的 LFA 算法
FaceSnap RECORDER	C-VIS http://www.c-vis.com	基于弹性图匹配技术
FaceVACS-SDK	Cognitec http://www.cognitec-systems.com	采用多阶高斯导数滤波为特征，并对这些特征进行判别分析进行匹配识别
GodEye	IS'Vision	基于中科院计算所的自适应多特征融合判别分析
ID-2000	Imagis http://www.imagistechnologies.com	基于变形模板技术
PersonSpotter	Neven Vision (前身为 Eyematic) http://www.eyematic.com	基于 Gabor 特征的人脸识别方法
UnMask Plus	VisionSphere http://www.visionspheretech.com	基于 McGill 大学的基于几何结构特征的识别方法

1.3 人脸识别的主要挑战

上述众多成就表明人脸识别是一个研究热点，但同时也说明人脸识别还是一个远未完善的课题，其中很多问题还有待进一步研究。

关于人脸识别技术的综述文献已有很多,较早的有[Chellappa95][Zhao00a][周00][张00]等。而<Handbook of Face Recognition> [LiSZ04a]可以看成是最新、最全的人脸分析方面的综述。[山04]中对人脸识别的发展也做了精辟的总结。因此本文不再赘述人脸识别的综述,只是简单介绍 FERET 测试和 FRVT 测试得到的一些经典人脸识别系统的性能,以观察人脸识别的现状,然后从技术层面分析产生这种现状的原因。

由美国国防部资助的 FERET 项目共举行了三次人脸识别算法的评测 [Phillips96][Phillips98a][Phillips00],一些主要的人脸识别算法都参与了评测。三次 FERET 测试分别在 1994、1995 和 1996 年进行,1997 年还有一次补充测试。其测试结果被公认反映了当时人脸识别技术的最高学术水平。2000 年发表在 IEEE Trans. on PAMI 上的 FERET'96 的测试报告 [Phillips00]中指出,南加州大学(USC)的弹性图匹配技术 [Lades93][Wiskott97]、马里兰大学(UMD)的子空间判别分析方法 [Zhao98]、麻省理工大学(MIT)的贝叶斯方法 [Moghaddam97][Moghaddam98][Moghaddam00][Moghaddam02]等人脸识别技术具有最好的识别性能。FERET 评测的结果表明人脸识别算法在较理想图像采集条件、用户配合、中小规模的正面人脸数据库上达到了非常好的性能,并指出主流的人脸识别技术对光照、姿态、摄像机等变化和老化等由于非理想采集条件或者用户不配合造成的变化鲁棒性较差。因此,光照问题和姿态问题逐渐成为下一个研究热点。

继 FERET 项目之后,美国国防部有关部门进一步组织了两次针对商用系统的评测——Face Recognition Vendor Test (FRVT): FRVT2000[Blackburn01]、FRVT2002 [Phillips03]。FRVT 测试的数据库的规模更大,达到了 37,437 人的 121,589 幅图像,大多数知名的人脸识别商业系统公司都参加了此次评测。

最近的 FRVT2002 的结果 [Phillips03]得到了下列主要结论:(1)当前的商用人脸识别系统随着人脸测试库的规模扩大,性能大约呈 log 下降;(2)当前的商用人脸识别系统虽然对室内光照的合理变化不再敏感,但对室外人脸图像的识别率还很低,并指出室外光照条件下的人脸识别问题是一个亟待解决的研究课题;(3)人脸识别的性能几乎随图像采集时间的推移呈线性下降;(4)3D 变形模型对非正面的人脸图像的识别率提高很多。

综上所述,现有人脸识别系统在用户配合的理想情况下取得了比较令人满意的结果,但如在更大规模的人脸数据库、摄像环境不可控、用户不配合的情况下使用,目前最好的系统的识别性能急剧下降。因此,现有人脸识别系统尚未成熟,尤其需要针对性地解决在非理想摄像条件(光照变化、背景变化、摄像设备差异)和用户不配合(视角变化、表情变化、饰物乃至化妆)等问题。

从技术层面上来说,人脸识别的主要挑战有 [LiSZ04a]:

(1)人脸图像变化太大:人脸图像受到图像拍摄时的光照、视角、遮挡等影响。人脸还会有表情、年龄等变化。另外还有摄像机的成像参数,如相机的焦距、曝光等。这些变化使得我们很难从图像中提取人脸的本质特征。本文主要讨论光照变化对人脸图像的影响。

(2) 人脸图像分布在高复杂的非线性流形中：上述诸多变化因素导致人脸图像分布非线性流形中，因此线性子空间的方法表现不是太好。这可能有两种方法来解决：一是对图像作归一化，使得人脸图像的流形变得更简单，本文我们对不同光照下的人脸图像做光照补偿也是起到这种作用；另一个办法是设计更强大的学习和分类算法，使得系统可以有效地对非线性流形分类。

(3) 人脸图像的高维性与人脸样本的矛盾：诸多变化使得人脸图像分布在一个高维空间内，但我们可得的每个人脸的样本通常很少。因此很多算法很难泛化到未见过的样本。本文我们也将讨论合成新光照下的虚拟样本，以扩充样本集。

1.4 光照问题的提出及本文的主要贡献

在上述人脸识别问题的三大技术挑战中，光照问题都占据了重要份量。FERET 测试 [Phillips00] 表明即使是最好的人脸识别系统，其识别性能也随人脸图像的光照变化而急剧下降。最近的 FRVT 测试 [Phillips03] 对商用人脸识别系统的测试（测试结果如图 1-1 所示）也表明，虽然当今顶级的人脸识别系统对室内光照的合理变化不再敏感，但其对室外人脸图像的识别率还很低。室外人脸图像的特点之一是光照变化复杂。虽然在过去几年中对光照变化已经有很多相关研究，但对室外人脸图像识别尚无突破性进展。

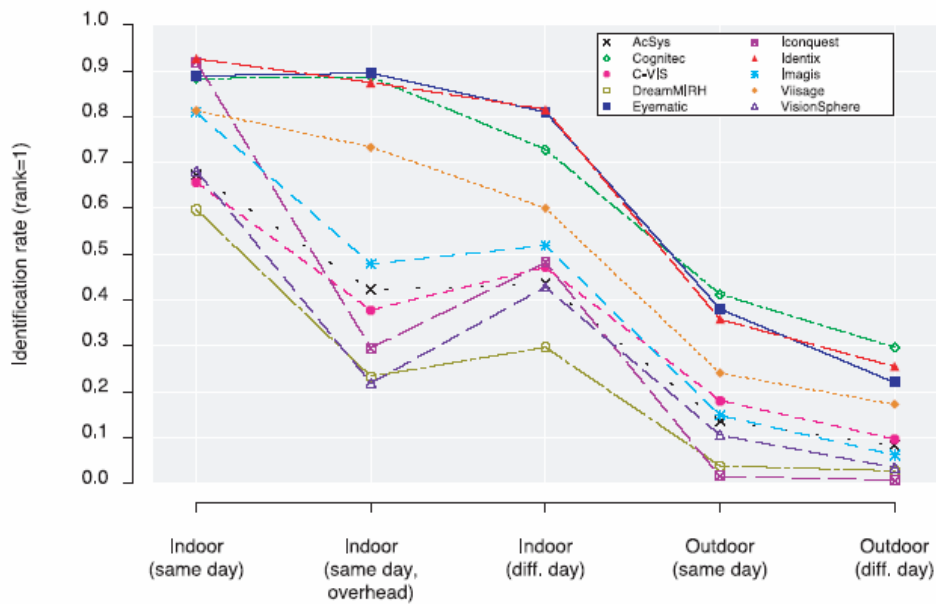


图 1-1: FRVT2002 对光照变化部分的测试结果（摘自 [Phillips03]）。

Figure 1-1. The results of face recognition under varying lighting in FRVT2002.

本文主要研究不同光照下的人脸图像识别，包括室外人脸图像，同时兼顾姿态变化，以找到与光照无关或对光照变化不敏感的特征，实现光照不变的人脸识别。本文的主要贡献有：

1. 任意光照下人脸图像的光照估计。光照问题可分解成两个问题：光照估计和光

照补偿。光照估计是解决光照问题的第一步。基于球面谐波理论，朗伯反射对光照而言是一个低通滤波器。因此本文在频域中对成像模型进行分析，对输入人脸图像进行光照估计，即估计光照的低频分量的系数。将人脸表面近似为朗伯反射，融合人脸类的先验知识，本文可对任意输入人脸图像进行光照估计。在频域中用少量频率系数的光照表示可以表示任何类型的光照，包括室外光照环境。且由于未知数数目很少，光照估计对输入人脸图像的分辨率及配准精度不敏感。

2. **基于绘制的人脸图像光照补偿。**为了实现光照不变的人脸识别，本文对输入不同光照下的人脸图像进行光照补偿。基于绘制的光照补偿是在光照估计的基础上，用光照商图像方法将其绘制到标准光照，从而使得所有的人脸图像在相同的光照条件下进行比较。本文讨论了两种标准光照：虚拟均匀光照及环境光+点光源。其中均匀光照下的图像中只包含人脸的纹理（Texture）信息而不包含影调（Shading）信息，与形状无关；环境光+点光源下的图像中既包含人脸的纹理信息，又包含了其形状信息。在精确配准的情况下，基于绘制的光照补偿对光照变化下的人脸识别能带来很大提高。

3. **基于差图像的人脸图像光照补偿。**本文提出的另一种光照补偿方式为基于差图像的光照补偿。差图像为输入图像与平均人脸在输入光照下的图像差。由于二者光照相同，差图像中光照变化的影响得到了减弱，同时又保留了用于识别的输入人脸与平均人脸的差别。基于差图像的光照补偿不需要输入人脸的形状，因此其对配准精度不敏感，更适合实用系统。

4. **基于 EigenHarmonics 模型的光照和姿态变化下的人脸识别。**谐波基图像包含了人脸的本质特征：形状和反射率，且谐波基图像适合处理人脸识别中的光照变化问题。同时谐波基图像中包含了人脸全部区域的信息，各种姿态下的人脸可以看作人脸区域可见的某部分。因此我们可以用人脸的谐波基图像用来表示人脸的不变特征，用于不同光照和姿态下的人脸识别。从一幅图像恢复人脸的谐波基图像是一个病态问题，因此本部分工作对所有人脸的谐波基图像的分布空间用 PCA 模型表示（EigenHarmonics 模型），结合该先验限制，在光照估计的基础上，可从一幅图像恢复该人脸的谐波基图像。同时由于 PCA 系数数目相对图像像素数目很少，根据姿态变化的人脸图像中可见部分仍可恢复出该人脸的谐波基图像的 PCA 系数，因此该模型可兼顾姿态变化。

1.5 论文的组织

论文第一章讨论了光照问题和计算机视觉及计算机图形学的关系，并引出一个经典的计算机视觉问题——人脸识别。在讨论了人脸识别的研究意义和人脸识别的发展现状后，分析了人脸识别中的主要挑战。然后引出了作为人脸识别中的主要挑战之一的光照问题。最后给出了本文的主要贡献。

第二章中首先从宏观和微观两个角度分析光照变化对人脸图像的影响，并介绍了人脸图像的成像模型，然后回顾了人脸识别中光照问题的主要研究成果并对各种方法作简

单总结。

第三章讨论光照问题的子问题之一——光照估计。本章在讨论朗伯表面对光照的低通特性的基础上，给出了光照估计原理及其在人脸图像上的应用。

在光照估计的基础上，第四章介绍了论文中关于光照补偿部分的工作。光照补偿是将不同光照下的输入人脸图像归一成标准形式。本章给出了两种光照补偿方式：基于绘制的光照补偿和基于差图像的光照补偿。并讨论了不同光照补偿方式对光照变化模式和配准精度的依赖及其对人脸识别的影响。最后还将本文提出的方法与一些经典的光照处理算法进行了详细比较。

第五章介绍了 **EigenHarmonics** 模型，该方法可以兼顾光照和姿态变化。本章按顺序介绍了模型的构造、重构输入人脸的谐波基图像和该模型用在人脸识别的应用。

最后，我们对整个论文进行了总结，并讨论了下一步研究方向。

第二章 人脸识别中光照问题综述

如第一章所述，光照变化问题是当前人脸识别系统的关键问题之一。本章首先从宏观和微观两个方面考察光照变化对人脸图像的影响，然后对解决人脸识别中的光照问题的方法作较详细的回顾，最后对各类方法做简单总结。

2.1 光照与人脸图像

我们首先分析不同光照下的人脸图像的分布以宏观地了解光照变化对人脸识别的影响，然后从图像的成像过程微观地探讨光照变化对人脸图像的影响。

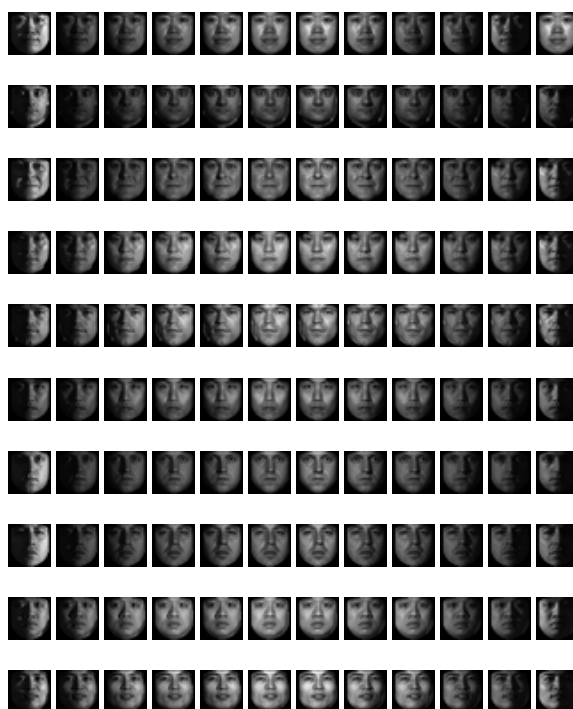
2.1.1 光照变化下的人脸图像的分布

同一人脸在不同光照下的图像之间的差别通常比不同人脸在相同光照下的图像之间的差别更大[Moses92][Moses94]。下面我们用实验进一步验证这一结论。

CMU-PIE 人脸数据库[Sim02]中包含了多个人脸在不同光照下的图像。图 2-1(a)给出了其中 10 个人脸在 12 不同光照下的一些示例图像。从中我们可以直观地感受同一人脸不同光照下的图像之间的差别（同一行）以及相同光照下不同人脸图像之间的差别（同一列）。

为了更清楚地观察光照变化对人脸图像分布的影响，我们用 PCA（Principal Component Analysis）方法[Duba01][边 01]将图像降至低维空间后得到人脸图像在不同光照下的流形。将图 2-1(a)中的图像构成一个 PCA 子空间，这个子空间可以看成是人脸图像分布的子空间。然后将这些图像投影到该子空间，得到各人脸的流形（每个人脸在光照变化下的轨迹图）如图 2-1(b)所示。对人脸识别来说，类内差为相同人脸不同光照下的图像之间的距离，类间差是不同人脸之间的距离。从图中可以看出，很多时候都是类内差比类间差大（如图中黄线比红线长）。所以不同光照下的人脸图像识别很困难。

另外图中各个人脸的轨迹形状类似且各个人脸轨迹之间的距离很小，表明光照对不同人脸的影响类似，不同人脸对人脸流形的影响相对较小。人脸识别是希望寻找到的特征只随人脸变化而不随光照变化而变化，而光照估计是希望从图像中获得的光照参数不随人脸变化而变化。从这个意义上来说，这两个问题互为反问题。从人脸识别的角度看，上述结论非常不利。而这对其反问题——光照估计来说却是个好消息。第三章中很多光照估计算法都显式或隐式地利用了上述结论。



(a)

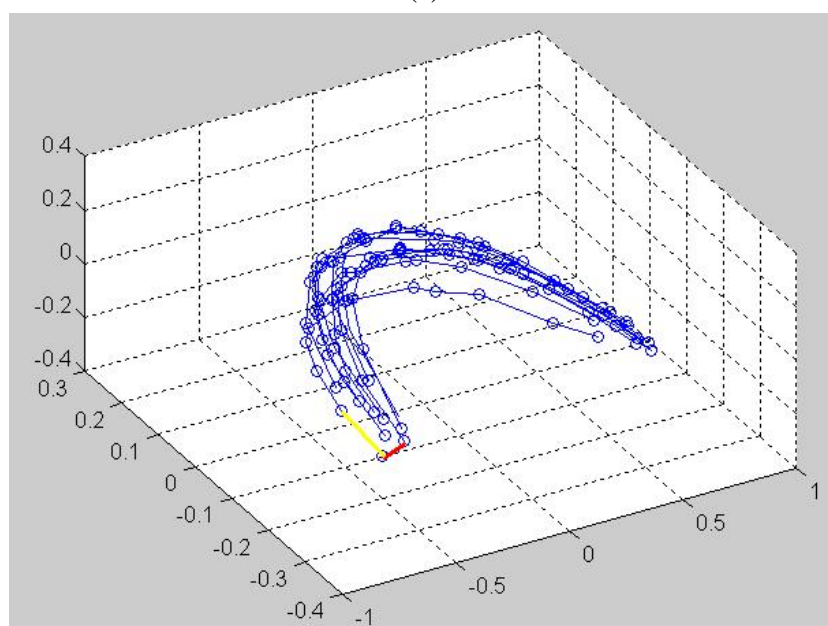


图 2-1: 不同人脸在不同光照下的图像的分布。

(a) 不同人在不同光照下的图像，其中每一行为相同人脸，每一列为相同光照。(b) (a)中的图像在人脸图像子空间的流形（每个人在不同光照下的轨迹）。

Figure 2-1. The manifolds of the persons under different illuminations. (a) The face images of the ten persons under different illuminations. (b) The face manifolds (consisting of “trajectory” of each person under the varying lighting) in the 3-D Eigenfaces.

2.1.2 图像成像模型

上面我们从宏观上了解了光照对人脸识别的影响，下面我们接着从微观上探究光照

对人脸图像的影响。图 2-2 为一幅很常见的人脸图像，可我们看到图中既有阴影区域，也有因镜面反射而饱和的区域。本小节我们从图像形成的过程来简单阐述上述现象产生的原因。关于成像模型更详细的介绍请参见计算机视觉[马 97]或计算机图形学[石 02]教材中辐射学的相关部分。

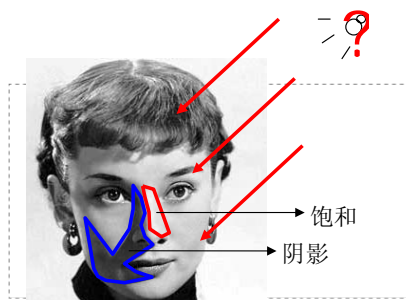


图 2-2: 光照和人脸图像。

由于人脸形状和反射属性的复杂性，光照对人脸图像产生各种影响，如阴影区域和因镜面反射而饱和的区域等。

Figure 2-2. The relation between the illumination and the facial image. For the complex shape and BRDF of the human face, the illumination will cause many effects on the facial image such as the attached shadows, cast shadows and the saturation by specular reflectance.

图像亮度 I 是入射光照照射到物体表面，再由物体表面反射到相机中的光的能量。假设成像系统是理想的，则图像亮度与物体的反射光成正比。因此下面我们并不严格区分图像亮度 I 和物体的反射辐射度 L 。

双向反射分布函数 BRDF (Bi-directional Reflectance Distribution Function) 告诉了我们当表面被从一个方向照明，而从另外一个方向观察时，此表面将显示多亮的信息，记

为 $\rho(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{dL_o(P, \theta_i, \phi_i, \theta_o, \phi_o)}{dE_i(\theta_i, \phi_i)} = \frac{dL_o(P, \theta_i, \phi_i, \theta_o, \phi_o)}{L_i(P, \theta_i, \phi_i) \cos \theta_i d\omega_i}$ ，其中 (θ_i, ϕ_i) 为入射方向，

(θ_o, ϕ_o) 为观察方向， dL_o 为反射的辐射度， dE_i 为入射的辐照度， L_i 为入射辐射度， $d\omega_i$ 微分立体角。

光照满足线性可加性，即由 N 个光源产生的反射光等于分别计算各个光源的反射光相加。对物体表面上的一点 P 来说，其上半球的光线都对该点的亮度有贡献，即所有光源对表面的辐照度等于 $E_o = \int_{\Omega_i} L_i(P, \theta_i, \phi_i) \cos \theta_i d\Omega_i$ ，其中 $\cos \theta_i$ 是考虑从 (θ_i, ϕ_i) 方向观察时表面的投影而引入的，积分区间 Ω_i 为可见的上半球。为求得表面的辐射度，我们在

球面上所有光线射入方向上对 BRDF 与辐照度的乘积作积分。因此

$$L_o(P, \theta_o, \phi_o) = \int_{\Omega_i} \rho(\theta_i, \phi_i, \theta_o, \phi_o) dE_i(\theta_i, \phi_i), \quad (2-1)$$

$$= \int_{\Omega_i} \rho(\theta_i, \phi_i, \theta_o, \phi_o) L_i(P, \theta_i, \phi_i) \cos \theta_i d\Omega_i。$$

式(2-1)被称为反射方程（Reflection Equation）或绘制方程（Rendering Equation），是许多图形绘制工具如 OpenGL 的绘制依据。式(2-1)用图示表示为图 2-3。

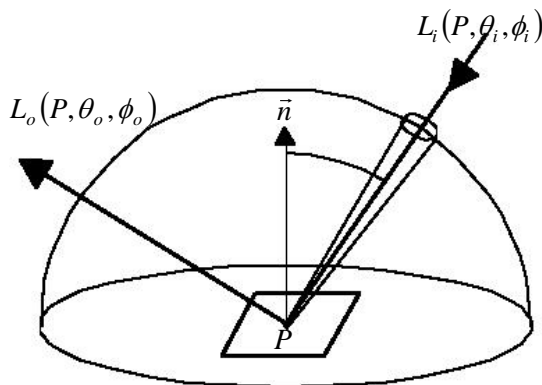


图 2-3: 反射方程的表示。

Figure 2-3. The reflection equation.

在实际应用中，BRDF 一般定义在局部坐标系（Z 轴与局部表面的法线重合）中，而入射光定义在全局坐标系（以观测者为中心）。为了区分局部坐标和全局坐标，下面我们用带上标的 X'Y'Z' 表示局部坐标，用无上标的 XYZ 系统表示全局坐标。为了简单起见，我们还假设光源为远光源（distant illumination），这样光照可以用入射角索引，与表面位置无关，即 $L_i(P, \theta_i, \phi_i)$ 可记为 $L_i(\theta_i, \phi_i)$ 。为了描述方便，我们再定义一个传播函数

$\hat{\rho} = \rho \cos \theta_i'$ ，这样式(2-1)变成

$$L_o(P, \theta_o', \phi_o') = \int_{\Omega_i} \hat{\rho}(\theta_i', \phi_i', \theta_o', \phi_o') L_i(\theta_i, \phi_i) d\Omega_i'。 \quad (2-2)$$

局部坐标和全局坐标的关系如图2-4所示。为了计算，我们需要将光线转换到局部坐标系。令物体表面上点 P 的球坐标为 (α, β) ，我们定义 $R_{\alpha, \beta, \gamma}$ 为一个旋转操作，将局部坐标 (θ_i', ϕ_i') 转换成全局坐标 (θ_i, ϕ_i) 。其中参数 γ 控制各向异性的表面局部切平面绕法线的旋转。

对各向同性的表面，该参数没有意义。 $R_{\alpha, \beta, \gamma}$ 可用标准欧拉角（Euler Angles）表示为

$R_{\alpha, \beta, \gamma} = R_z(\beta)R_y(\alpha)R_z(\gamma)$ ，其中 R_z 表示绕 Z 轴旋转， R_y 表示绕 Y 轴旋转。这样，局部坐标和全局坐标之间的关系为：

$$\begin{aligned} (\theta_i, \phi_i) &= R_{\alpha, \beta, \gamma}(\theta_i', \phi_i') = R_z(\beta)R_y(\alpha)R_z(\gamma)\{\theta_i', \phi_i'\}, \\ (\theta_i', \phi_i') &= R_{\alpha, \beta, \gamma}^{-1}(\theta_i, \phi_i) = R_z(-\gamma)R_y(-\alpha)R_z(-\beta)\{\theta_i, \phi_i\}。 \end{aligned} \quad (2-3)$$

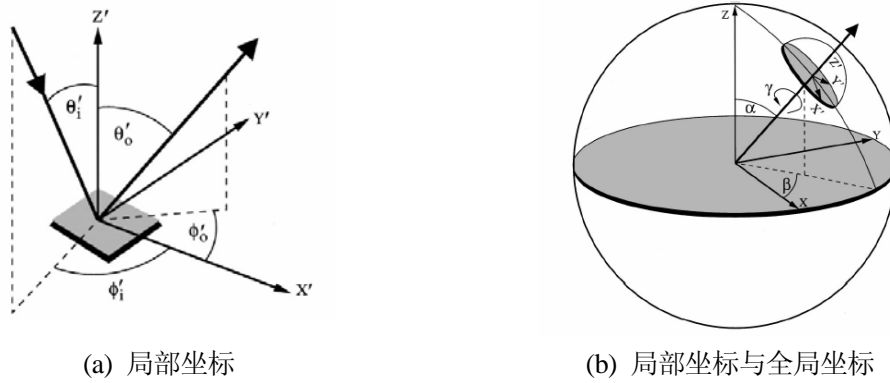


图 2-4: 局部坐标和全局坐标。

Figure 2-4. The local coordination and the global coordination.

因此反射方程式(2-2)可写成

$$L_o(\alpha, \beta, \gamma, \theta'_o, \phi'_o) = \int_{\theta'_i=0}^{\pi/2} \int_{\phi'_i=0}^{2\pi} \hat{\rho}(\theta'_i, \phi'_i, \theta'_o, \phi'_o) L_i(\mathbf{R}_{\alpha, \beta, \gamma}(\theta'_i, \phi'_i)) \sin \theta'_i d\theta'_i d\phi'_i. \quad (2-4)$$

式(2-4)是旋转卷积。式(2-4)在局部坐标系中描述反射方程，将入射光旋转到局部坐标，然后对上半球积分。该式中入射光为输入信号，而BRDF可视为滤波器。

再考虑到投射阴影 (Cast Shadow)，在每个光源前乘一个遮挡函数 $C(i)$ 。 $C(i)$ 为二值函数，其取值为0或1，取决于物体表面是否被遮挡了来自光源 $L_i(\theta_i, \phi_i)$ 的光线，这样

$$L_o(\alpha, \beta, \gamma, \theta'_o, \phi'_o) = \int_{\theta'_i=0}^{\pi/2} \int_{\phi'_i=0}^{2\pi} C(i) \hat{\rho}(\theta'_i, \phi'_i, \theta'_o, \phi'_o) L_i(\mathbf{R}_{\alpha, \beta, \gamma}(\theta'_i, \phi'_i)) \sin \theta'_i d\theta'_i d\phi'_i, \quad (2-5)$$

一般物体表面都包含漫反射和镜面反射成分，所以上式中的 BRDF 可写成这两部分相加的形式，其中 ρ_d 表示漫反射系数， ρ_s 表示镜面反射系数，

$$\hat{\rho}(\theta'_i, \phi'_i, \theta'_o, \phi'_o) = \rho_d + \rho_s(\theta'_i, \phi'_i, \theta'_o, \phi'_o). \quad (2-6)$$

在下面将要介绍的人脸识别中的光照问题解决方法中，除了一部分是完全从图像处理的角度出发外，其余都是基于上述成像模型，只是有些算法对上述模型进行了裁减或变形。如人脸识别中一般不考虑镜面反射成分、不考虑投射阴影，有些算法甚至不考虑attached 阴影。另外一些算法简单地将图像 I 看成是纹理 R 与光照影响 L 相乘的结果， $I = RL$ 。

2.2 人脸识别中光照问题研究概述

自从 FERET 测试 [Phillips00] 指出光照变化对人脸识别算法的影响很大以后，研究者们对人脸识别中的光照问题作了大量工作，提出了很多光照处理算法。最近的 FRVT 测试 [Phillips03] 对商用人脸识别系统的测试也表明，当今顶级的人脸识别系统对室内光照的合理变化已不再敏感，但对室外人脸图像的识别率还较低，并指出室外光照环境下

的人脸图像识别是下一个主要的研究方向。虽然在过去几年中对光照变化已经有很多相关研究，但对室外人脸图像识别尚无突破性进展。光照变化问题仍是一个还在研究发展中的问题。下面我们对人脸识别中光照问题研究的一些代表性工作做简单回顾。

一些人脸识别算法将光照变化对人脸图像的影响看成是影响人脸图像的因素之一，如同表情和姿态等变化一样。由不同人脸不同引起的图像之间的差别为类间差，由光照等引起的相同人脸的图像之间的差别为类内差。这些算法学习类内差的分布，从而使得人脸识别系统能克服这些变化的影响。不同的方法表现在不同的类内差/类间差处理策略。Fisherfaces 方法[Belhumeur97]将光照变化引起的人脸图像看成是类内变化，学习类间差和类内差的分布，再用线性判别分析 (Linear Discriminant Analysis, LDA) [边 01]的方法进行人脸识别。LDA 判别准则是使类间散度最大的同时使类内散度最小。贝叶斯方法[Moghaddam97][Moghaddam98][Moghaddam00][Moghaddam01][Moghaddam02]则对类内差和类间差分别用两个特征子空间表示，相似度用两幅图像差为类内差的贝叶斯概率 (MAP 估计或 ML 估计) 来度量。统一的子空间方法[WangXG03][WangXG04]用一个统一的子空间分析人脸之间的本质差别、各种成像参数 (包括光照) 引起的差别和噪声，并揭示了三种主要的子空间方法: PCA (Eigenfaces)、LDA (Fisherfaces) 和贝叶斯方法之间的内在联系。贝叶斯方法的另一大贡献是使得人脸识别问题由多类问题变成了两类问题。这样很多两类分类方法，如 SVM[Phillips98b][Guo00][Jonsson00][Jonsson02]和 AdaBoost [山 04]策略都可以溶入到人脸识别中来。Q. Li 等对一些线性投影方法在人脸识别中的光照变化问题上作了一些经验比较[LiQ04]。

由于光照对图像的影响实在很大，针对人脸图像中的光照变化问题，研究者们提出了很多专门的方法，FRGC (Face Recognition Grand Challenge)[FRGC05]的任务之一是对各种光照预处理方法做评测。[Sim01]将这些方法分成三大类: 不变特征、标准形式和变化建模。其中不变特征是指从图像中得到的对光照不变或至少对光照变化不敏感的特征，然后用于人脸识别，如商图像[Raviv99]和边缘图等。标准形式是指将任意光照下的图像都重新绘制到标准光照，从而得到其标准图像，如对称 SFS[Zhao00b]中用正面点光源下的标准图像用来做人脸识别。变化建模则是研究光照变化对人脸图像的影响，并对变化光照下的人脸图像建立一个模型表示，如子空间方法[Belhumeur97]。[Sim01]的分类主要是根据最后得到的特征的性质。本文我们主要根据光照处理算法中算法的处理过程，将这些方法分成另外的三大类: 基于图像处理技术的方法、基于子空间的方法和基于成像模型的方法。当然这个分类也不是非常严格，三者有时也会有交叉。

2.2.1 基于图像处理技术的方法

基于图像处理技术的方法主要是从图像处理出发，对输入图像做各种图像处理操作或信号变换，从而寻求光照不变或至少对光照变化不敏感的特征。这类方法包括基本的图像处理技术，如边缘图、梯度、图像与 2DGabor 卷积、log 变换等。另外直方图均衡化(Histogram Equalization, HE)、直方图规定化(Histogram Specify, HS)和 Gamma 灰度校

正(Gamma Intensity Correction, GIC) [Shan03]等灰度变换操作也被很多人脸识别算法作为光照预处理操作。最近一些更复杂的操作, 如比率图(Ratio Image)[Belhumeur98]、图像梯度方向的概率分布 [Chen00]、自商图像(Self Quotient Image)[WangHT04a][WangHT04b]、相对梯度方法(Relative Gradient Edge)[Wei04]和相位图(Phase Spectrum) [Savvides04]在有限的数据集上也表现了很好的效果。

■ 光照方向补偿

最初处理光照变化的方法都是从 2D 图像中直接提取对光照变化不敏感的特征, 如边缘图[Kanade77]等, 然后用于人脸识别。但对各种特征的性能都没有定性或定量的分析。Adini 等[Adini97] 通过实验来评价这些特征对光照变化的敏感性。文中比较了四种特征: 边缘图、Gabor 卷积、灰度的导数及 \log 变换。Adini 的实验结果表明, 任何表示本身都不能克服由光照方向改变引起的变化。这可能是因为在人脸识别中的光照补偿不仅涉及低级视觉层, 还需要高级视觉活动的参与。而文中的表示发生在低级视觉层, 如果包含特定物体或特定类的高层处理可能会更好。这为后续的研究工作提供了很好的参考。

值得一提的是, 虽然 Adini 提出 Gabor 卷积并不是对光照变化不敏感, 多个尺度、多个方向的 Gabor 变换特征还是作为人脸识别中一种光照鲁棒的特征被很多算法所采用 [Tefas01][LiuCJ02][LiuCJ03][山 04]。

■ 比率图 (Ratio Image)

Belhumeur 等[Belhumeur98]通过比较两幅图像的比率图像的复杂度, 来判断两幅图像是来自于不同光照下的同一物体, 还是来自于不同的两个物体。该方法基于这样一个事实: 相同物体在不同方向点光源下的两张图像之间的比率图像要比不同的两个物体在不同的点光源下的图像之间的比率图像更简单。比率图的复杂性用比率图的梯度的模

的平方和表示, 即 $\frac{\partial I}{\partial x J} = \frac{(I_x J - J_x I)^2}{J^4}$ 。由于光照对图像的影响在低频范围: 对平滑后的

图像取 \log 后, 其梯度的欧氏距离可近似为比率图的复杂性度量, 因为

$$\left(\frac{\partial(\log I)}{\partial x} - \frac{\partial(\log J)}{\partial x} \right)^2 = \left(\frac{I_x J - J_x I}{IJ} \right)^2。$$

■ 图像梯度方向的概率分布

H. F. Chen 等[Chen00]证明, 对朗伯表面, 没有任何区分函数是光照无光的, 但图像梯度方向对光照变化相对不敏感。所以文中对图像的梯度的角度的分布采用概率的方法建模。令 $(r_i, \varphi_i), (r_j, \varphi_j)$ 分别为图像 I 和 J 的梯度的模和角度, 两幅图像 I 和 J 的梯度的

角度的联合概率密度函数定义为 $\rho_\varphi(\varphi) = \iint \rho(r_1, \varphi, r_2) dr_1 dr_2, \varphi = \varphi_i - \varphi_j$ 。概率密度函数

$\rho_\varphi(\varphi)$ 通过从训练集中相同物体在不同光照下的图像的梯度角度的分布学习得到。两幅图像 I 和 J 的梯度（角度）的联合概率为 $P(I, J) = \prod_{i \in M} \rho_\varphi(\varphi(i))$ 。最后进行人脸识别时，若 $P(I, J) > \tau$ ，则两幅图像 I 和 J 是由同一对象产生的。

■ 自商图像 (Self Quotient Image)

自商图像[WangHT04a][WangHT04b]是基于简化的成像模型: $I = RL$ ，其中 I 为图像， R 为表面反射， L 为光照。由于光照 L 可以被认为是图像的低频分量，可以用一个低通滤波器 F 来估计，即 $L \approx F * I$ 。这样将得到自商图像 $R = \frac{I}{F * I}$ 。文献中选择多个高斯平滑核 G_1, G_2, \dots, G_n 作为低通滤波器，并且每个高斯窗中各个像素的权值根据其局部信息进行调整。自商图像在 CMU-PIE 人脸数据库上报告了较好的结果。

■ 相对梯度 (Relative Edge)

相对梯度[Wei04]对将梯度进行改进，融入一些局部的信息，从这一点上来说相对梯度与自商图像有些类似。相对梯度定义为 $\bar{G}(x, y) = \frac{|\nabla I(x, y)|}{\max_{(u, v) \in W(x, y)} |\nabla I(u, v)| + c}$ ，其中 $W(x, y)$ 为 (x, y) 附近的一个邻域窗口， $\nabla I(x, y)$ 为图像在点 (x, y) 处的梯度。相对梯度在 CMU-PIE 人脸数据库上报告了很好的实验结果。

■ 相位图 (Phase Spectrum)

Savvides 提出将 EigenPhases[Savvides04]方法用于对光照和遮挡不变的人脸识别。其基本思想是将图像从空域经 FFT 变换到频域后，相位图 (Phase Spectrum) 比能量图 (Magnitude Spectrum) 含有更多的对识别有用的信息，且相位图对光照变化和遮挡不敏感。其报告的在 CMU-PIE 人脸数据库上的识别率很高。我们分析原因主要是光照的影响在低频范围，只用相位谱信息相当于将各个频带的能量都设为相同。由于实际图像中的能量大多集中在低频部分，这样只用相位信息相当于减小了低频信息的能量，起到高通滤波的作用，从而将光照的影响减弱。

2.2.2 基于子空间的方法

基于子空间的方法是同一物体在不同光照下的图像分布在一个低维线性子空间内，识别则是选择与输入图像最相近的子空间。这个低维线性子空间可以通过统计学习构造，也可以是从图像的成像模型分析得到，所以这部分工作也可以归到基于模型的方法。

■ 3D 线性子空间

3D 线性子空间[Shashua97][Shashua94]的基本思想是：忽略成像模型中的镜面反射和阴影的影响，朗伯表面的成像公式为 $I(p) = \lambda(p) \cdot \bar{n}(p) \cdot \bar{s}$ ，其中 $\lambda(p)$ 为点 p 处的反射率， $\bar{n}(p)$ 为其法线， \bar{s} 为光源。这样朗伯表面的成像公式是一个 3 维的线性反射模型。物体在 $k = 3$ 个线性无关的光照下的图像 I_i 就构成了该物体所有光照下的图像的一组基图像，

$$I(p) = \sum_{i=1}^k \alpha_i I_i(p)。$$

■ 4D 线性子空间

4D 线性子空间[Marschner97]最初是用于图像合成的。其基本思想是朗伯表面对光照而言是一个低频滤波器，所以光照的影响可以用一个低阶多项式表示。文中用常数项和 3 个一阶多项式表示。有意思的是，下面将要介绍的 9D 线性子空间的前 4 维刚好和该 4D 线性子空间相同。

■ 5D 线性子空间

受 Eigenfaces[Turk91]的启发，Hallinan [Hallinan94]对各种光照下的人脸图像用 PCA 模型来描述。首先对光照球上的各个光源方向 (θ_i, ϕ_j) 稠密采样，得到相应光照下的图像 $\tilde{I}(\theta_i, \phi_j)$ ，然后通过这些图像进行 PCA 分析找到一组基图像 S_k ， $\tilde{I}(\theta_i, \phi_j) \approx \sum_k \alpha_{ijk} S_k$ 。通过实验，Hallinan 得出结论，5 个特征图像就足够描述人脸在不同光照条件下的人脸图像分布空间。Epstein 将这种方法扩展到其他物体，指出使用 5 ± 2 个特征图像就可以模拟物体在任意光照条件下的图像[Epstein95]。

[Ramamoorthi02]在谐波基图像理论（参见下面的 9D 线性子空间方法）的基础上，解析地证明了人脸在任意光照下的图像在 5D 线性子空间内，这也从理论上印证了 [Hallinan94]的结论。

■ 9D 线性子空间

最近，通过在空间-频域中分析朗伯凸表面的反射方程，[Basri01a][Basri03]和 [Ramamoorthi01a]分别从计算机视觉和计算机图形学的角度独立地证明了朗伯凸表面在任意光照变化下的图像可以由该对象的球面谐波基图像张成的 9D 线性子空间很好近似。该发现极大地推动了对一般光照的建模并使得任意光照下，尤其是室外环境下的人脸识别变得可能。这个 9D 线性子空间的基由低阶多项式（低维球面谐波函数）构成，因此 [Marschner97]的 4D 子空间可以看成是只取 9D 子空间的 4 个低频部分。我们将这个工作称为球面谐波理论。

由于该理论给出了解析的结果，推动了光照变化下的很多视觉[Lee01][Simakov03]

和图形学[Ramamoorthi01b][Ramamoorthi02b][Sloan02]相关理论和应用的发展。其中 9 个点光源方法[Lee01]从光照锥和 9D 球面谐波基图像的关系出发，找到每个人脸 9 个最佳点光源下的图像，这 9 幅图像构成该人脸光照变化空间的非正交基。另外最近还有一些工作试图寻找统计低维线性子空间和 9D 解析线性子空间之间关系[Basri01b][Ramamoorthi02][Nilus03a][Nilus03b][Sato03b]，将统计线性子空间和解析线性子空间统一起来。

2.2.3 基于成像模型的方法

基于成像模型的方法对引起图像变化的各个因素，如人脸的本质特征（形状和反射率）和外在的成像条件（如光照）分别建模，试图从图像中恢复人脸的本质特征。

尽管从一幅图像恢复物体的形状和纹理是一个病态问题[ZhangR99]，但当我们有多幅输入图像[Yuille97][Georghiades01][Shashua94][Shashua97]，或将研究限制在人脸对象，融合人脸这个大类的先验知识[Atick96][Shashua01][Worthington01][Zhao00b][Blanz99][Blanz03]，就可以在在一定程度上实现该功能。人脸类的先验知识大多是人脸的几何形状和反射率在不同人脸之间变化较小。这些先验知识可以通过训练或增加限制得到。不同基于成像模型的方法表现在先验知识的表示和获取的不同，以及先验知识与后验知识（输入图像）的融合策略不同。下文中提到的几种典型算法用到的先验知识如表 2-1 所示。

■ 商图像 (Quotient Image)

商图像方法[Raviv99][Shashua01]是假设人脸为一个有着相同形状但不同反射率的一个理想类中的对象。商图像定义为同类中两个不同对象的反射率之比，又可转化为两个对象在相同光照下的图像之比。即 $Q_y(x, y) = \frac{\rho_y(x, y)}{\rho_a(x, y)} = \frac{\rho_y n^T s_j}{\rho_a n^T s_j} = \frac{y_j}{a_j}$ 。商图像只包含纹理信息，因此与光照无关。

根据 3D 线性子空间理论[Shashua97]，给出 a_1, a_2, a_3 为物体 a 在 s_1, s_2, s_3 三种线性独立光照条件下的三张图像和物体 y 在某光照条件 s 下的图像，那么存在系数 x_1, x_2, x_3 满足 $y_s = (\sum_j x_j a_j) \otimes Q_y$ ，其中 $s = \sum_j x_j s_j$ ，所以如果求出系数 x_1, x_2, x_3 就可以得到商图像。

为了求得系数 x_1, x_2, x_3 ，需要一个训练集。训练集 (bootstrap set) 中包括 N 个物体在 3 种光照条件下的 $3N$ 张图像 $A_i, i = 1, \dots, N$ ， A_i 为 $m \times 3$ 的矩阵，每一列为一个对象的一幅图像。再假设人脸类的纹理为有理扩展 (ration span)，即

$$\frac{\rho(x, y)}{\rho_y(x, y)} = \alpha_1 \rho_1(x, y) + \dots + \alpha_N \rho_N(x, y), \quad \rho(x, y) = \sum_{i=1}^N \rho_i^2(x, y),$$

则图像 y_s 的光照系数为

$$x = \sum_i \alpha_i v_i, \text{ 其中 } v_i = \left(\sum_{r=1}^N A_r^T A_r \right)^{-1} A_i^T y_s, \alpha_i y_s^T y_s - \left(\sum_{r=1}^N \alpha_r v_r \right)^T A_i^T y_s = 0 \text{ 且满足 } \sum_i \alpha_i = N。$$

商图像方法处理时只要一幅图像即可得到对应物体的商图像，且提供了一种简单有效的图形合成的方法。但该方法要求物体是朗伯表面，没有考虑镜面反射和阴影，且假设同类的所有物体具有相同的形状，这些要求对真实环境下的人脸识别来说太严格，如自然光照。在这种情况下，如果忽略阴影，合成的图像就会不真实。为了学习到阴影的影响，[WangHT03]和[Shan03]通过扩大训练集，用类似商图像的思想也取得了较好的实验结果。

■ 光照锥 (Illumination Cone)

光照锥[Belhumeur96][Georghiades98][Georghiades99][Georghiades01]是基于成像模型方法中另一个很成功的方法。光照锥同时考虑光照变化和姿态变化。该方法基于这样一个事实：在固定姿态下，一个物体在任意光照下的图像（ n 个像素）为 IR^n 中的一个凸多面锥（光照锥）。每个人脸可以表示成不同姿态的光照锥的集合。每个光照锥可以由某些极线光照（extreme rays）下的图像表示， $x_{ij} = \max(Bs_{ij}, 0)$ ，其中 $s_{ij} = b_i \times b_j$ 。特定对象的光照锥可以用其3幅合适的图像，采用光度立体视觉算法来构造。实际应用中，特定对象3幅合适的图像可能得不到（如人脸），这时需用更多的图像来构造光照锥。光照锥方法首先从输入的7幅图像中采用扩展的光度立体视觉的算法恢复物体表面的反射率和形状。根据估计出的反射率和形状，再采用光线跟踪技术绘制该对象的极线光照下的图像（通常对光球均匀采样），然后对光照锥用一个低维线性子空间来近似（PCA降维）。人脸识别则是选择与输入图像最近的子空间，从这个角度来说，光照锥方法也可归入子空间方法一类。但由于其主要贡献在人脸形状和纹理的恢复，我们还是将其在此描述。

光照锥报告的识别率是在Yale B人脸数据库上迄今为止是最高的，但该方法需要输入7幅人脸图像且计算复杂，这在一定程度上限制了该方法的更多应用。

■ 对称 SFS (Symmetric Shape From Shading)

纯粹的SFS是一个病态问题，而对称SFS[Zhao00b][Zhao00c]利用人脸的对称信息将未知数的数目减半。该方法的基本思想是利用人脸对称信息，由给定的任一光照条件下的照片 I ，得到其在正面光源条件下的照片 I_p ，然后用 I_p 进行识别。由 I 得到 I_p 的过程

$$\text{为 } I_p(x, y) = \frac{\sqrt{1 + P_s^2 + Q_s^2}}{2(1 + qQ_s)} [I(x, y) + I(-x, y)], \text{ 其中 } q \text{ 是人脸形状的梯度, 可以通过计算人脸}$$

三维模型的偏导数求得， $P_s = k \sin \alpha \cos \tau$ ， $Q_s = k \sin \alpha \sin \tau$ ， k 是光线的强度， α, τ 表示光线方向的角度值。这种方法不需要恢复纹理信息，但需要通过从影调中恢复光源（source from shading）技术得到光源的角度信息，而且要已知人脸的三维模型。

■ 统计 SFS (Statistical Shape From Shading)

统计 SFS[Sim01]通过学习人脸图像的基图像和残差图像的统计分布来恢复输入人脸的基图像。文中提出了一个扩展的朗伯反射方程 $i(x) = \mathbf{n}(x)^t \mathbf{s} + e(x, \mathbf{s})$ ，其中 $\mathbf{b}(x) = \mathbf{n}(x)^t$ 为基图像，是反射率与表面法向的乘积， $e(x, \mathbf{s})$ 表示残差，主要是由阴影和镜面反射等非朗伯反射引起的。给定一组已知光照条件 \mathbf{S} 下的一组图像 \mathbf{B} ，假设 $\mathbf{n}(x)^t$ 和 $e(x, \mathbf{s})$ 为高斯分布，统计 SFS 学习 $\mathbf{N} = (\mathbf{S}^T)^{-1} \mathbf{S} \mathbf{B}^T$, $\mathbf{E} = \mathbf{B} - \mathbf{N}^T \mathbf{S}$ 的统计分布的参数 $\mu_n(x)$, $C_n(x)$ 和 $\mu_e(x, \mathbf{s})$, $\sigma_e(x, \mathbf{s})$ 。

对输入的一幅人脸图像，该方法首先用核回归的方法估计其光照，然后计算 $\mu_e(x, \mathbf{s})$, $\sigma_e^2(x, \mathbf{s})$ ，最后用 MAP 估计求得 $n_{MAP}(x) = \arg \max_{n(x)} \Pr(n(x) | i(x))$ 。

统计球面谐波基图像方法[ZhangL03]类似 Sim 的统计 SFS，只是将基图像 $\mathbf{n}(x)^t$ 换成人脸的球面谐波基图像，使得光照环境可为任意光照且模型中已经考虑 attached 阴影，不必将 attached 阴影放在误差项 $e(x, \mathbf{s})$ 中考虑。统计球面谐波基图像方法也取得了很好的实验结果。

■ 3D 变形模型 (3D Morphable Model)

3D 变形模型[Blanz99][Blanz02][Blanz03]可对姿态和光照同时变化的情况进行处理，而且其报告的姿态变化和光照变化下的人脸识别率很高。3D 变形模型的基本思想是将人脸类的形状和纹理用 PCA 模型表示，即 $\mathbf{S} = \bar{\mathbf{s}} + \sum_{i=1}^{m-1} \alpha_i s_i$ ， $\mathbf{T} = \bar{\mathbf{t}} + \sum_{i=1}^{m-1} \beta_i t_i$ 。这样估计输入人脸的形状和纹理就变成了估计形状和纹理的 PCA 参数： α 和 β 。3D 变形模型在恢复输入人脸的形状和纹理参数的同时，还估计图像绘制参数 ρ 。参数估计是一个迭代优化的

过程，目标是使得能量函数 $E = \frac{1}{\sigma_I^2} E_I + \frac{1}{\sigma_F^2} E_F + \sum_i \frac{\alpha_i^2}{\sigma_{S,i}^2} + \sum_i \frac{\beta_i^2}{\sigma_{T,i}^2} + \sum_i \frac{(\rho_i - \bar{\rho}_i)^2}{\sigma_{R,i}^2}$ 最小。人

脸识别用人脸的本质特征 $\mathbf{C} = (\frac{\alpha_i}{\sigma_{S,i}}, \frac{\beta_i}{\sigma_{T,i}})$ 实现。该算法已经成功应用于

FRVT'2002[Phillips03]中对姿态的矫正。但由于涉及迭代优化，3D 变形模型方法的计算较复杂。

在 3D 变形模型方法之前，[Atick96] 假设人脸的纹理为常数，也对人脸的 3D 形状用 PCA 模型来建模。

■ 本质图像 (Intrinsic Images)

Weiss[Weiss01]提出的从序列图像中恢复物体的本质图像虽然不是针对人脸识别,但如果能得到人脸的本质图像(光照图像和反射率图),反射率图可作为光照不变的特征用于人脸识别。本质图像的基本思想为:给定同一对象不同光照下的几幅图像,然后对这些图像取 \log 运算后进行滤波(滤波器可为 $(1,0,-1)$ 和 $(1,0,-1)^T$),则反射率图的滤波后的图像为所有图像的中值估计的结果。这是基于假设光照图像滤波后的图像是稀疏的(其统计分布符合拉普拉斯分布,这已由图像统计工作[Olshausen96][Huang00][Dror01]等证明)。这种方法随着图像数目的增加效果很好,但其缺点是需要多幅图像,且这些图像需严格对齐,这对实用人脸识别系统来说很苛刻。所以这种方法没有直接被用到人脸识别系统中。[Qing04]将其中的中值(Median)操作替换成最小化(Min)操作,使得对任意两幅图像就可以进行相似度计算从而用于人脸识别,但由于对图像取 \log 运算后的边缘特征本质为图像中的高频分量,其对图像配准要求严格的问题还是没有解决。

表 2-1: 几种典型的基于模型的方法所用的先验知识。

Table 2-1. The prior knowledge used in some typical model based algorithms.

方法	先验知识
商图像[Shashua01]	所有人脸形状相同, 人脸的反射率为有理扩展
统计 SFS[Sim01]	每个像素点的基图像 $b(x)$ 和误差项 $e(x)$ 服从高斯分布
对称 SFS[Zhao00b]	人脸对称, 人脸形状相似(基于稀疏关键点的变形)
3D 变形模型[Vetter03]	反射率和 3D 形状都用 PCA 模型近似

2.3 光照处理方法总结

在人脸识别光照变化问题的方法中,早期的研究大都是集中在基于图像处理技术的不变特征方法,该类方法现在也一直在发展。该类方法主要来源于图像处理技术,计算速度快且无需任何先验知识的学习,原则上可适用于包含人脸在内的任意对象。但该类方法并没有从本质上分离光照,只能从一定程度上减弱光照变化的影响,离光照不变的人脸识别还是有较大差距。最近的相位图特征对光照变化取得了较好的效果,但相位图[Savvides04]对图像配准敏感(类似高通滤波的方法都对配准敏感),所以将其用在实用系统中还有待进一步研究。

基于图像处理技术研究中最近一个明显的趋势是将基本的图像处理技术结合区域进行分析。如相对梯度[Wei04]、自商图像[WangHT04a][WangHT04b]、基于块的直方图均衡化(block-based histogram equalization)[Xie05]和更早一些的基于区域的直方图均衡化和 Gamma 灰度校正[Shan03]。这些方法相对其对应的基本图像处理技术,人脸识别结果都有了很大提高。我们分析其中主要原因是因为光照对人脸图像的影响在低频范围,即在一个局部区域内光照对图像影响的变化是很小的。如在相对梯度方法中,取局部区域内最大的梯度,我们可以认为此最大梯度通常不是由光照变化引起的,而是由人脸本身

的纹理变化引起的。因此最后的特征对光照变化不敏感。我们期望其他图像处理技术与区域处理相结合也会得到较好的结果。

从上世纪九十年代初 **Eigenfaces** 方法[Turk91]出现之后,基于统计子空间的方法引起了人们的关注。基于统计子空间的方法从训练集中学习光照变化下人脸图像的分布, Hallinan [Hallinan94]证明了光照对图像的影响在一个 5D 子空间内。而且实践证明当训练集的光照条件覆盖了测试集的光照条件时,该类方法表现得确实很好。但是基于统计学习的方法性能依赖于训练集,如果训练集的光照条件不能覆盖测试集的光照条件,其性能不佳。而当前系统对训练集中光照条件的采样都是经验性的,还没有理论证明如何对光照采样是最优的。最近[Sato04]基于采样定理和球面谐波基图像模型的工作认为应该均匀采样 $4N^2$ 个采样点,其中 N 为频带个数。当 $N=9$,需 36 个采样点。这对一般系统来说采样个数还是太多。解析线性子空间方法则需要知道物体的 3D 表示(形状和纹理)或者需要不同光照下的多幅图像来构造子空间的基。最近有一些工作试图将统计子空间和解析构造子空间方法之间建立联系[Basri01a][Ramamoorthi02][Nilius03a][Nilius03b][Sato03b],并已经得到了一些很有意思的结论,但这方面工作还有待深入。

因为图像的形成是一个 3D 向 2D 投影的过程,仅仅在 2D 图像上处理不能从根本上解决光照问题。要彻底解决问题,还是从 3D 角度来分析。同时计算机运算速度的大大加快,使得基于成像模型的方法近些年涌现出来。解析线性子空间和基于成像模型的方法从图像的成像模型出发,分析光照变化对人脸图像的影响,可从本质上解决光照变化问题。但是为了计算方便,多数方法都假设人脸为朗伯凸表面,即不考虑人脸图像中的互反射(interreflection)和投射阴影,也不考虑镜面反射。在商图像方法[Shashua01]中,甚至连 attached 阴影也没有考虑。这些方法在人脸识别中虽然取得了较好的结果,但由于人脸的形状和表面反射的复杂性,还是损失了一部分精度。如果已知人脸的 3D 形状,[Ishiyama02]用一个二值函数表示投射阴影。3D 变形模型[Blanz03]考虑了投射阴影和镜面反射,但考虑的光照条件为环境光+一个主光源。另外其运算速度还是很慢,虽然其改进的匹配算法速度已经大大提高[Romdhani03]。基于成像模型的方法中的主要问题是对模型做更进一步的细化。虽然简单的模型已经取得了一定效果,但要得到更完美的效果,还需要在各方面精耕细作,既包括考虑现实的成像模型,也包括从算法的速度上做工作。

现在很多算法假设比较简单的光照环境(如单个点光源或一个点光源+背景光),但现有人脸识别系统的瓶颈之一是室外光照,而室外光照的一个特点是光照变化复杂。最近的 9D 线性子空间的方法可以解析地表示任意光照,从这一点上来说,其意义重大。

基于上述考虑,本文从图像的成像模型出发,基于球面谐波理论,首先估计输入图像的光照,然后将图像做光照补偿或直接用人脸的本质特征(球面谐波基图像)用于人脸识别。9D 线性子空间方法[Basri01a]需要知道输入人脸的 3D 结构,9 个点光源的方法[Lee01]需要输入人脸在特定 9 个点光源下的图像,这些输入对实用的人脸识别系统来说

有些困难。与前面的基于球面谐波理论的方法不同，我们尝试只用一幅人脸图像，融合人脸类的先验知识，达到对任意光照下的人脸图像进行光照估计和光照补偿的目的。从这一点来说，我们的工作和统计谐波基图像方法[ZhangL03]类似，但我们所用的人脸类的先验知识的表示不同，从而最后的融合策略也有所区别。在后续的几个章节，我们将详细介绍我们在不同的阶段的人脸类的先验知识的表示及其最终结果。

第三章 基于球面谐波理论的人脸图像光照估计

人脸识别中的光照问题可以分解成两部分：光照估计和光照补偿。根据视觉心理学实验，人在识别同种光照条件下的图像比识别不同光照下的图像的速度快，因此认为人在识别不同光照条件下的物体时有一个类似光照补偿的过程。而且人们在不知道具体物体（同类别）时，就可以估计出光照条件。

本文我们借用该结论，将光照估计放在光照补偿之前进行。从算法实现的角度分析，光照对图像的影响是全局的，光照条件的参数通常比图像中的像素数少得多，所以光照估计利用图像整体信息，可以在光照补偿之前进行。

本章是如下组织的。首先我们介绍关于光照估计方面前人做的一些代表性工作。从第二章中，我们得到结论：**BRDF** 对光照而言是一个滤波器，通过进一步分析，球面谐波理论得到结论：朗伯表面的 **BRDF** 对光照是一个低通滤波器。然后基于这一结论，给出人脸图像的光照估计的详细过程。在给出多个数据集上人脸图像光照估计的实验结果及其分析之后，最后小结本章。

3.1 光照估计综述

光照估计是给定物体的一幅/多幅图像，物体的形状和反射属性可以是已知或未知，输出物体成像时的光照条件。一根光线可以用 $L(\theta, \phi)$ 表示，其中 (θ, ϕ) 表示光线的方向， L 表示光线的亮度。由于真实自然界中光线可能来自各个方向，因此要描述整个光照变化还是需要无穷多个参数。所以实际系统中都是假定特定类型的光照（如一个或几个点光源）或是对整个光照空间离散采样。光照估计就是估计用来描述光照的这些参数。光照估计在计算机视觉中被称为从影调中恢复光源（**Source From Shading**）。在计算机图形学中被称为反加光（**Inverse Lighting**），可以看作是反绘制（**Inverse Rendering**）的一部分。

最初的光照估计[Hron86]是从影调恢复形状（**Shape From Shading, SFS**）的附属品。这种方法通常假定最简单的光照，即只有一个点光源的情况。为了处理更复杂的光照环境，更多的算法用一个参考对象（如已知形状和反射率的球）来测量场景的光照。1976年 Blinn 和 Newell[Blinn76]首次采用景物的镜面包围球及其经、纬度定义环境映射（**Environment Map**），可以看成是各个方向光照的记录。基于这种思想，Debevec 用对一个用全反射球（**mirror ball**）的多张不同曝光的多张图像获得的环境映射的高动态图（**High Dynamic Range Image, HDRI**）[Debevec97][Debevec98][Debevec00]来表示来自各个方向的光线，以克服图像的低动态范围低（通常为 0~255），真实地表示各个方向光线的强度。由于全反射玻璃球会影响其附近物体的局部光照，Zhang[ZhangYF00][ZhangYF01]等提出用一个漫反射参考球来减小参考球对周围物体的影响。其基本思想是

通过确定球面上的关键点（Critical Points）来估计场景中的多个点光源的方向。其中关键点的法线与光源方向垂直，所以关键点的亮度变化最大。对于球面大圆上的一段弧 τ 和弧上一点 θ ，用 $E(\theta) = b_i \sin \theta + c_i \cos \theta$ 表示其辐照度。如果两个相邻两段子弧 $[\theta - h, \theta]$ 和 $[\theta, \theta + h]$ ，对应的系数 b_-, c_- 和 b_+, c_+ 之间的距离 $d = \sqrt{(b_- - b_+)^2 + (c_- - c_+)^2} > T$ ，则认为这两段弧之间必定有至少一个关键点，其中 h 为搜索步长， T 为阈值。Y. Wang 等[WangY02]将该工作扩展到已知形状的任意物体，并结合投射阴影来估计多个光源的方向。

Sato 等根据投射阴影在光照估计方面做了很多工作 [Sato99a][Sato99b][Sato01][Sato03a]，因为阴影中包含更多的场景光照的信息。这些工作用一个已知形状的物体在另一个已知形状的场景上产生的阴影作为输入，将光照离散采样成一个测地球上的 n 个点光源，这样光照估计就是求这 n 个点光源的亮度，可用一组线性方程组求解。同时该方法还可以迭代求解阴影区的反射参数，可以用上一次的光照估计结果指导下次的光照自适应采样。

Ramamoorthi[Ramamoorthi01c]从信号处理的角度，在频域中分析反射方程，得到反射方程在频域中的表示为 $B_{lmpq} = \Lambda_l L_{lm} \hat{\rho}_{lpq}$ （各向同性反射表面），其中 Λ_l 为归一化因子， B_{lmpq} 为反射光场的频谱系数， L_{lm} 为光照的频谱系数， $\hat{\rho}_{lpq}$ 为反射函数的频谱系数。根据该反射方程表示，可得到一般意义下的光照问题可解的条件。

Marschner[Marschner97][Marschner98]则用一组光照基来表示光照对图像的影响，任意光照用这些光照基的线性组合来表示。这样光照估计就是估计这些光照基的系数，可用一组线性方程组求解。

与人脸识别相关的光照估计中，很多方法将未知光照表示为训练光照的线性组合，尤其是从线性子空间延伸过来的方法，如3D线性子空间[Shashua97]、商图像[Shashua01]、光照锥[Georghiades01]和9D线性子空间方法[Lee01]。这样光照估计就是估计线性组合的系数。这类问题通常可以用最小二乘法求解或SVD（Singular Value Decomposition）分解来求解。另外统计SFS[Sim01]和统计谐波基图像[ZhangL03]采用核回归方法来求解线性组合的系数。3D变形模型方法[Blanz03]则是用迭代优化的方法直接求解环境光和主光源的颜色和位置参数。上述方法估计出来的光照是物理上真实的光照，还有一些方法求出的相对光照，如分块直方图方法[Xie05]。

本文我们基于Baris[Baris01a]和Ramamoorthi[Ramamoorthi01c]等的工作，在频域中分析图像的成像模型。由于3D球面上的频域基函数为球面谐波函数，我们称之为球面谐波理论。根据球面谐波理论，光照对人脸图像的影响在低频范围，则对人脸图像有影响的光照用其低频系数表示即可。光照估计就是估计光照的低频系数。这种光照表示可以表示物理上任意复杂的光照，包括室外光照环境。所以下面我们首先介绍由Baris[Baris01a]和Ramamoorthi [Ramamoorthi01c]提出的球面谐波理论，然后介绍光照估计的

原理，最后介绍我们如何融合输入人脸图像和人脸类的先验知识得到输入人脸的球面谐波基图像，从而最终实现光照估计。

3.2 球面谐波理论

为计算方便，我们假设人脸表面为朗伯凸表面，即不考虑镜面反射成分，也不考虑互反射和投射阴影。在此基础上，我们对成像模型进行分析，进而对输入人脸图像进行光照估计。

因为成像模型（如第 2.1.2 节中介绍的反射方程）为旋转卷积，因此适合在空间—频域中分析。反射方程描述图像亮度/反射光照 $L_o(P, \theta_o, \phi_o)$ 与入射光照 $L_i(\theta_i, \phi_i)$ 、BRDF $\hat{\rho}(\theta_i, \phi_i, \theta_o, \phi_o)$ 在球面坐标系下的关系。因此下面我们分别介绍 3D 球面上的频域基函数——球面谐波函数、朗伯反射的球面谐波展开和入射光照的球面谐波展开，最后得到反射方程在频域中的表示形式。

3.2.1 球面谐波函数

类似于 2D 平面上的傅立叶变换，在 3D 球面上的频域基函数为球面谐波函数 Y_{lm} (Spherical Harmonic)。球面谐波函数来源于许多物理系统，如量子物理等。球面谐波函数的介绍详见参考文献[Tannoudji77][Inui90]。

球面谐波函数 Y_{lm} 定义为：

$$Y_{lm}(\theta, \phi) = N_{lm} P_{lm}(\cos \theta) \exp(\text{Im} \phi),$$

$$N_{lm} = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}, \quad (3-1)$$

其中 N_{lm} 为归一化因子。在上式中，方位角 ϕ 用傅立叶基函数表示，而 θ 的依赖性用勒让德函数 (Legendre functions) P_{lm} 表示。下标索引满足条件 $l \geq 0$ 且 $-l \leq m \leq l$ 。因此第 l 阶上有 $2l+1$ 个基函数，前 l 阶共有 l^2 个基函数。如前 3 阶 ($l=0,1,2$) 共 9 个球面谐波函数。这 9 个球面谐波函数的三角形式 (球面坐标系中用 θ, ϕ 表示) 及多项式形式 (笛卡尔坐标系中用 x, y, z 表示, 其中 $x^2 + y^2 + z^2 = 1$) 为式 (3-2),

$$Y_{00}(\theta, \phi) = \frac{1}{2} \frac{1}{\sqrt{\pi}}, \quad (3-2)$$

$$\begin{aligned}
 Y_{1-1}(\theta, \phi) &= \frac{1}{2} \sqrt{\frac{3}{2\pi}} \sin \theta e^{-i\phi} = \frac{1}{2} \sqrt{\frac{3}{2\pi}} (x - iy), \\
 Y_{10}(\theta, \phi) &= \frac{1}{2} \sqrt{\frac{3}{\pi}} \cos \theta = \frac{1}{2} \sqrt{\frac{3}{\pi}} z, \\
 Y_{11}(\theta, \phi) &= -\frac{1}{2} \sqrt{\frac{3}{2\pi}} \sin \theta e^{i\phi} = -\frac{1}{2} \sqrt{\frac{3}{2\pi}} (x + iy), \\
 Y_{2-2}(\theta, \phi) &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{-2i\phi} = \frac{1}{4} \sqrt{\frac{15}{2\pi}} (x - iy)^2, \\
 Y_{2-1}(\theta, \phi) &= \frac{1}{2} \sqrt{\frac{15}{2\pi}} \sin \theta \cos \theta e^{-i\phi} = \frac{1}{2} \sqrt{\frac{15}{2\pi}} (z(x - iy)), \\
 Y_{20}(\theta, \phi) &= \frac{1}{4} \sqrt{\frac{5}{\pi}} (3 \cos^2 \theta - 1) = \frac{1}{4} \sqrt{\frac{5}{\pi}} (3z^2 - 1), \\
 Y_{21}(\theta, \phi) &= -\frac{1}{2} \sqrt{\frac{15}{2\pi}} \sin \theta \cos \theta e^{i\phi} = -\frac{1}{2} \sqrt{\frac{15}{2\pi}} (z(x + iy)), \\
 Y_{22}(\theta, \phi) &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\phi} = \frac{1}{4} \sqrt{\frac{15}{2\pi}} (x + iy)^2.
 \end{aligned}$$

在计算时，有时用球面谐波函数的实数形式更方便。球面谐波函数的实数形式的标准定义为

$$Y_{lm}(\theta, \phi) = \begin{cases} \sqrt{2} N_{lm} P_{lm}(\cos \theta) \cos(m\phi) & m > 0 \\ N_{lm} P_{lm}(\cos \theta) & m = 0. \\ \sqrt{2} N_{lm} P_{lm}(\cos \theta) \sin(m\phi) & m < 0 \end{cases} \quad (3-3)$$

但是这种定义使得符号有时不一致。对前 3 阶 ($l = 0, 1, 2$) 的球面谐波函数的实数形式，我们写成

$$\begin{aligned}
 (x, y, z) &= (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \\
 Y_{00}(\theta, \phi) &= \frac{1}{2} \frac{1}{\sqrt{\pi}}, \\
 (Y_{1-1}; Y_{10}; Y_{11})(\theta, \phi) &= \frac{1}{2} \sqrt{\frac{3}{2\pi}} (y; z; x), \\
 (Y_{2-2}; Y_{2-1}; Y_{21})(\theta, \phi) &= \frac{1}{2} \sqrt{\frac{15}{2\pi}} (xy; yz; xz), \\
 Y_{20}(\theta, \phi) &= \frac{1}{4} \sqrt{\frac{5}{\pi}} (3z^2 - 1), \\
 Y_{22}(\theta, \phi) &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} (x^2 - y^2).
 \end{aligned} \quad (3-4)$$

类似于 2D 圆上的傅立叶函数，球面谐波函数在 3D 球面上形成一组完全标准正交基。因此任何函数 $f(\theta, \phi) \in L^2$ ，都可展开成球面谐波函数的线性组合，

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\theta, \phi), \quad (3-5)$$

$$f_{lm} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} f(\theta, \phi) Y_{lm}^*(\theta, \phi) \sin \theta d\theta d\phi.$$

其中 $Y_{lm}^*(\theta, \phi)$ 表示球面谐波函数 $Y_{lm}(\theta, \phi)$ 的共轭形式。

球面谐波函数的旋转形式定义为

$$Y_{lm}(R_{\alpha, \beta, \gamma}(\theta, \phi)) = \sum_{m'=-l}^l D_{mm'}^l(\alpha, \beta, \gamma) Y_{lm'}(\theta, \phi). \quad (3-6)$$

也就是说，旋转后的球面谐波函数用其不同 m 索引的球面谐波函数的组合表示，而索引 l 不变。式中 D^l 为一个 $(2l+1) \times (2l+1)$ 的矩阵，它表示旋转后球面谐波函数如何通过同阶的原球面谐波函数的线性组合来重构。值得一提的是由于旋转 $R_{\alpha, \beta, \gamma} = R_Z(\beta)R_Y(\alpha)R_Z(\gamma)$ ， D^l 对 β, γ 的依赖很简单，因为球面谐波函数绕 Z 轴的旋转很直观，所以

$$D_{mm'}^l(\alpha, \beta, \gamma) = d_{mm'}^l(\alpha) e^{lm\beta} e^{lm'\gamma}, \quad (3-7)$$

其中 d^l 表示绕 Y 轴旋转后球面谐波函数该如何变换。

在群 (Group) 理论中， D^l 为旋转群 $SO(3)$ 的 $2l+1$ 维表示。 D^l 的解析形式比较复杂，可参见文献[Inui90]。为了方便后续章节的描述，这里给出矩阵 D^l 的一个重要性质：

$$D_{m0}^l(\alpha, \beta, \gamma) = d_{m0}^l(\alpha) e^{lm\beta} = \sqrt{\frac{4\pi}{2l+1}} Y_{lm}(\alpha, \beta). \quad (3-8)$$

3.2.2 入射光照的球面谐波展开

入射光照 $L_i(\theta_i, \phi_i)$ 在全局坐标中可用球面谐波函数表示为

$$L_i(\theta_i, \phi_i) = \sum_{l=0}^{\infty} \sum_{m=-l}^l L_{lm} Y_{lm}(\theta_i, \phi_i). \quad (3-9)$$

为了得到其在局部坐标系中的表示，我们对入射光做一个旋转：

$$L_i(\theta_i, \phi_i) = L_i(R_{\alpha, \beta, \gamma}(\theta'_i, \phi'_i)) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_{m'=-l}^l L_{lm} D_{mm'}^l(\alpha, \beta, \gamma) Y_{lm'}(\theta'_i, \phi'_i). \quad (3-10)$$

3.2.3 朗伯反射的 BRDF 的球面谐波展开

忽略与入射角和出射角都无关的朗伯反射系数 ρ_d ，朗伯反射的 BRDF 为

$\hat{\rho}(\theta'_i, \phi'_i, \theta'_o, \phi'_o) = \cos \theta'_i$ 。将其展开到频域表示为

$$\hat{\rho}(\theta'_i, \phi'_i, \theta'_o, \phi'_o) = \sum_{l=0}^{\infty} \hat{\rho}_l Y_{l0}(\theta'_i, 0), \quad (3-11)$$

其中系数

$$\hat{\rho}_l = 2\pi \int_0^{\pi} Y_{l0}(\theta'_i) \cos \theta'_i \sin \theta'_i d\theta'_i, \quad (3-12)$$

其中 2π 来自于对方位角的积分，积分上限为 $\frac{\pi}{2}$ 是因为只对可见的上半球积分。

上式可以简单地用Legendre多项式 $P(\cos \theta'_i)$ 表示，用 $u = \cos \theta'_i$ 带入上积分式， $P_l(u) = u$ ，

$$Y_{l0}(\theta'_i, 0) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos \theta'_i), \quad \text{则}$$

$$\hat{\rho}_l = 2\pi \sqrt{\frac{2l+1}{4\pi}} \int_0^1 P_l(u) P_l(u) du, \quad (3-13)$$

P_l 的奇偶性和 l 相同，Legendre多项式在定义域 $[-1, 1]$ 上是正交的，即

$$\int_{-1}^1 P_a(u) P_b(u) du = \frac{2}{2a+1} \delta_{a,b}.$$

当 $l=1$ 时， $P_a(u)P_b(u)$ 为偶函数，

$$\hat{\rho}_l = 2\pi \sqrt{\frac{2l+1}{4\pi}} \times \frac{1}{2} \int_{-1}^1 P_1(u) P_1(u) du = \pi \sqrt{\frac{2l+1}{4\pi}} \times \frac{2}{2+1} = \sqrt{\frac{\pi}{3}}. \quad (3-14)$$

当 $l > 1$ 的奇数时， $P_a(u)P_b(u)$ 为偶函数， $\int_0^1 P_l(u) P_l(u) du = \frac{1}{2} \int_{-1}^1 P_l(u) P_l(u) du$ ，

$$\hat{\rho}_l = 2\pi \sqrt{\frac{2l+1}{4\pi}} \times \frac{1}{2} \int_{-1}^1 P_l(u) P_l(u) du = \pi \sqrt{\frac{2l+1}{4\pi}} \times 0 = 0. \quad (3-15)$$

当 $l > 1$ 的偶数时， $P_a(u)P_b(u)$ 为奇函数， $\int_0^1 P_l(u) P_l(u) du = -\int_{-1}^0 P_l(u) P_l(u) du$ ，

$$\hat{\rho}_l = 2\pi \sqrt{\frac{2l+1}{4\pi}} \times \frac{(-1)^{l/2-1}}{(l+1)(l-1)} \times \left[\frac{l!}{2^l ((l/2)!)^2} \right]. \quad (3-16)$$

总而言之，当 $l > 1$ 的奇数时， $\hat{\rho}_l = 0$ ；当 $l > 1$ 的偶数时， $\hat{\rho}_l \sim l^{-2}$ ，因此 l 增大时， $\hat{\rho}_l$ 迅速下降。图3-1为 $\hat{\rho}_l \sim l$ 图。所以朗伯反射函数对光照而言是一个低通滤波器，即只有光照的低频部分对朗伯表面有影响。从Heisenberg测不准原理也可以得到类似的定性的结论，因为朗伯表面的反射方程在空域中积分区域很大（上半球），则其在频域中一定局限在低频范围。

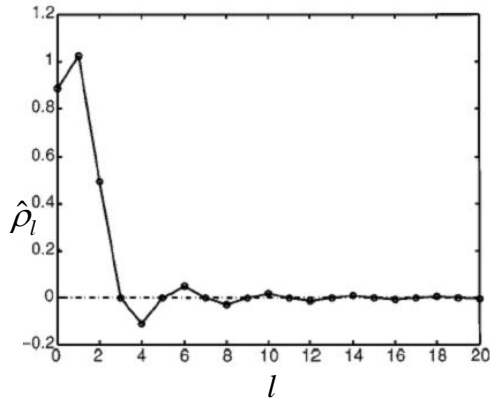


图 3-1: 朗伯反射函数的频域表示。

朗伯反射函数各频带的系数 $\hat{\rho}_l$ ，可以看出奇数阶 $l > 0$ 的系数 $\hat{\rho}_l = 0$ 。偶数阶的系数随着 l 的增大急剧下降。

Figure 3-1. The plot of Lambertian BRDF coefficients $\hat{\rho}_l$ versus l . It can be seen that odd terms with $l > 0$ have $\hat{\rho}_l = 0$. Also, as l increases, the coefficients decay rapidly.

3.2.4 反射方程的频域表示

对朗伯凸表面来说，由于式 (2-4) 中的出射角 (θ'_o, ϕ'_o) 没有意义，用以表示各向异性的参数 γ 也没有意义，因此可以简化为

$$L_o(\alpha, \beta) = \int_{\theta'_i=0}^{\pi} \int_{\phi'_i=0}^{2\pi} \hat{\rho}(\theta'_i, \phi'_i, \theta'_o, \phi'_o) L_1(\mathbf{R}_{\alpha, \beta, 0}(\theta'_i, \phi'_i)) \sin \theta'_i d\theta'_i d\phi'_i. \quad (3-17)$$

将式 (3-10) 和式 (3-11) 代入式 (3-17)，得到

$$L_o(\alpha, \beta) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_{m'=-l}^l \sum_{n=0}^{\infty} L_{lm} D_{mm'}^l(\alpha, \beta, 0) \hat{\rho}_n T_{lm'n} \quad (3-18)$$

$$T_{lm'n} = \int_{\theta'_i=0}^{\pi} \int_{\phi'_i=0}^{2\pi} Y_{lm'}(\theta'_i, \phi'_i) Y_{n0}^*(\theta'_i, \phi'_i) \sin \theta'_i d\theta'_i d\phi'_i = \delta_{ln} \delta_{m'0}.$$

所以 $n = l, m' = 0$ ，等式 (3-18) 变成

$$L_o(\alpha, \beta) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_{m'=-l}^l L_{lm} D_{m0}^l(\alpha, \beta, 0) \hat{\rho}_l. \quad (3-19)$$

将式 (3-8) 代入式 (3-19)，得到

$$L_o(\alpha, \beta) = \sum_{l=0}^{\infty} \sum_{m=-l}^l L_{lm} \hat{\rho}_l \sqrt{\frac{4\pi}{2l+1}} Y_{lm}(\alpha, \beta). \quad (3-20)$$

为了书写方便, 令新变量 $A_l = \hat{\rho}_l \sqrt{\frac{4\pi}{2l+1}}$, 则

$$L_o(\alpha, \beta) = \sum_{l=0}^{\infty} \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha, \beta). \quad (3-21)$$

由于朗伯反射的BRDF对光照是一个低通滤波器, 即上式中取 $l=0,1,2$ 即可, 所以

$$L_o(\alpha, \beta) \approx \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha, \beta). \quad (3-22)$$

其中 $A_0 = \pi, A_1 = \frac{2\pi}{3}, A_2 = \frac{\pi}{4}$, L_{lm} 为光照的球面谐波系数。

3.3 光照估计原理

3.3.1 球面谐波基图像

假设人脸为朗伯凸表面, 考虑到人脸各点的反射率, 用 $\lambda(x, y)$ 表示人脸表面一点 (x, y) 的反射率, $(\alpha(x, y), \beta(x, y))$ 表示该点的法向, 则该点的图像亮度为

$$\begin{aligned} I(x, y) &\approx \lambda(x, y) \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha(x, y), \beta(x, y)), \\ &= \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} b_{lm}(x, y), \end{aligned} \quad (3-23)$$

其中

$$b_{lm}(x, y) = \lambda(x, y) A_l Y_{lm}(\alpha(x, y), \beta(x, y)), \quad (3-24)$$

为人脸的球面谐波基图像。球面谐波基图像可以看作是人脸在谐波光照下的图像。谐波光照为只含有一个谐波成分的光照环境, 由于谐波光照中函数值有时为负, 所以谐波光照只是虚拟光照, 在真实世界中并不存在。一个人脸的球面谐波基图像的如图 3-2 所示。

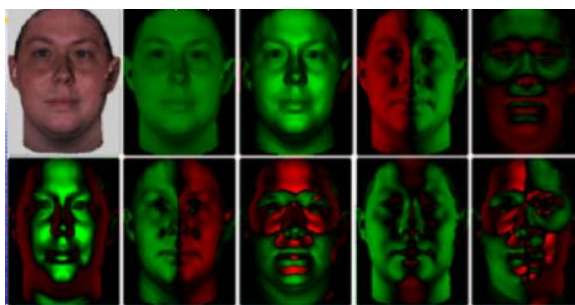


图 3-2: 人脸球面谐波基图像示例。

第一行第一个图像为输入人脸, 其余图像为 9 个球面谐波基图像, 其中绿色表示正值, 红色表示负值。

Figure 3-2. The example of spherical harmonic images of a face.

3.3.2 光照估计

根据式(3-24)，球面谐波基图像与光照无关，而只与输入人脸的本质特征（形状和纹理）有关。因此如果已知输入人脸的球面谐波基图像，根据式(3-23)，光照估计就是找到一种光照，使得在该光照下的图像与输入图像差别最小。这样，给定一幅人脸图像 I (n 维列向量，其中 n 为图像中像素的个数)，光照估计就是估计式(3-23)中的 9 个光照系数 L_{lm} ($0 \leq l \leq 2$)。这可以通过解下述最小二乘问题

$$\hat{\mathbf{L}} = \arg \min_{\mathbf{L}} \|\mathbf{BL} - \mathbf{I}\| \quad (3-25)$$

得到，其中 \mathbf{B} 为谐波基图像的矩阵形式，是一个 $n \times 9$ 矩阵，矩阵的每一列为一个谐波基图像 $b_{lm}(x, y)$ 。

3.4 输入人脸的球面谐波基图像

为了利用式(3-25)进行光照估计，我们需要知道人脸的球面谐波基图像 b_{lm} 。球面谐波基图像 $b_{lm}(x, y) = \lambda(x, y)A_l Y_{lm}(\alpha(x, y), \beta(x, y))$ ，包含了人脸的纹理信息和法线（形状）信息，因此需知道人脸的反射率图 $\lambda(x, y)$ 和法线图 $(\alpha(x, y), \beta(x, y))$ ，其中法线又与 3D 形状相关。9D 线性子空间方法[Basri01a]利用已知的人脸 3D 结构计算球面谐波基图像，L. Zhang[ZhangL03]等假设人脸谐波基图像每点的分布符合高斯分布，再融合输入人脸的信息，用 MAP 估计得到输入人脸的谐波基图像。在本文中，输入人脸的反射率图和法线图也是通过“人脸类”的先验知识得到。但我们所用的人脸类的先验知识的表示不同，从而最后的融合策略也有所区别。

由于球面谐波基图像包含了人脸的形状和纹理信息，下面我们分别就人脸的形状和纹理加以叙述。如果已知输入人脸纹理和形状，即可根据公式(3-24)计算其球面谐波基图像。

3.4.1 输入人脸的形状

人脸的形状大致相同，该假设已用于很多视觉算法中[Shashua01][Zhao00][Wen03]。由于表示光照的参数只有 9 个，相对图像的像素数目很少，因此可以将图像的分辨率降低。而分辨率降低之后，不同人脸的形状差异就更小了。且光照对图像的影响是全局的，微小的形状偏差不会对光照估计造成严重影响。另据视觉心理学实验报告人在不知道具体物体（同类别）时，就可以估计出光照条件，因此光照估计时可以用平均人脸模型替代输入人脸的形状。

为了尽可能地减小输入人脸形状与平均人脸形状的差别，本文采用基于特征点的变形（Morph）技术，将平均 3D 人脸形状变形到输入人脸，以得到尽可能与输入人脸相似的 3D 形状。给定输入人脸图像，我们首先将输入人脸图像与平均 3D 人脸形状的关键特征点对齐，

其余像素通过图像变形技术对齐。本文中我们用一种局部纹理约束的主动表观模型 (Local Texture Constraint Active Appearance Model, LTC-AAM) [WangW02][山04], 定位人脸图像上的关键特征点。在输入人脸图像的特征点与平均3D人脸形状的特征点对齐之后, 其他的顶点我们采用Kriging插值[黄01], 将平均3D人脸形状变形到输入人脸, 以得到尽可能与输入人脸相似的3D形状。Kriging插值是一种最佳线性无偏估计, 它用已知点观察到的数据的线性组合来预测未知量, 利用方差来表述空间变化, 其目标是使得用预测值的空间分布来预测的得到的值的误差的均值为0, 方差最小。本文中我们使用的Kriging插值的半方差模型为指数模型, 即

$$\gamma(h) = \begin{cases} 0 & |h| = 0 \\ C_0 + C_1 \left(1 - \exp\left(\frac{-3|h|}{a}\right) \right) & \text{其他} \end{cases}, \quad (3-26)$$

其中核方差 (Nugget) $C_0 = 0.0$, 梁 (Sill) $C_0 + C_1$ 是最大方差, 与待变形的两个形状特征点位移的方差有关, 变程 (Range) a 为常数, 步长 (Distance) h 为两个特征点之间的方差。

有了输入人脸的形状, 其姿态也可以根据特征点的相对位置确定, 则我们可以很容易地计算输入人脸的法线图。图 3-3 给出了本文采用的平均人脸形状、输入人脸图像的特征点及其经过对齐后的特定人脸形状和法线图。

虽然 LTC-AAM 与基本的 ASM[Cootes95]和 AAM[Cootes98]相比, 特征点定位的精度有了很大提高, 但光照变化对特征点定位仍然是一个巨大的挑战。在光照条件很恶劣的情况下, LTC-AAM 自动定位的特征点不太准确。本文实验部分我们比较了只用两个眼睛位置对齐平均人脸形状模型和输入人脸图像的结果, 以观察系统对配准 (alignment) 精度的敏感性。

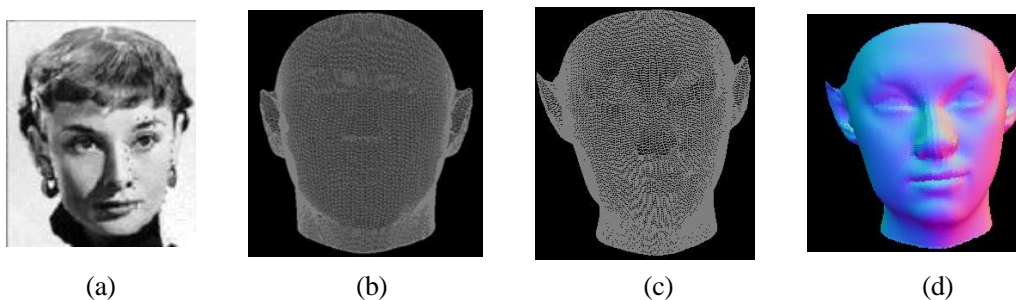


图 3-3: 平均人脸形状及输入人脸的形状。

(a) 输入人脸图像及其特征点位置。(b) 平均人脸形状。(c)变形得到的输入人脸的形状。(d) 输入人脸的法线图。

Figure 3-3. The shape of the average human face and the input face. (a) The input facial image and its feature points. (b) The mesh of the average human face. (c) The morphed mesh of the input face. (d) The normal map of the input face.

3.4.2 输入人脸的纹理

关于纹理，我们假设人脸表面各点的反射率近似相等。更确切的，我们假设人脸的纹理图中除了直流分量 λ_{00} 外不包含其他低频分量($1 \leq l \leq 4$)。因为反射率图也可在频域表示

为 $\lambda(\alpha, \beta) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \lambda_{lm} Y_{lm}(\alpha, \beta)$ ， (α, β) 为人脸表面上某点的法线的球面坐标。这样根据式

(3-23)，图像为

$$\begin{aligned} I(x, y) &\approx \lambda(x, y) \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha(x, y), \beta(x, y)), \\ &= \sum_{l1=0}^{\infty} \sum_{m1=-l1}^{l1} \sum_{l2=0}^2 \sum_{m2=-l2}^{l2} \lambda_{l1m1} L_{l2m2} A_{l2} Y_{l1m1}(\alpha(x, y), \beta(x, y)) Y_{l2m2}(\alpha(x, y), \beta(x, y)). \end{aligned} \quad (3-27)$$

而球面谐波函数乘法由Clebsch-Gordan展开[Tannoudji77]决定，即

$$Y_{l1m1}(\theta, \phi) Y_{l2m2}(\theta, \phi) = \sum_{l=|l1-l2|}^{l1+l2} \sum_{m=-l}^l \{C\langle l1, l2: 0, 0 | l, 0 \rangle \cdot C\langle l1, l2: m1, m2 | l, m \rangle Y_{lm}(\theta, \phi)\}, \quad (3-28)$$

其中 $C\langle l1, l2: m1, m2 | l, m \rangle$ 为Clebsch-Gordan系数。上式右边 $Y_{lm}(\theta, \phi)$ 的系数

$C\langle l1, l2: 0, 0 | l, 0 \rangle \cdot C\langle l1, l2: m1, m2 | l, m \rangle$ 当且仅当 $m = m1 + m2$ 且 $l1 + l2 - l$ 为偶数是才不等于0。

由于光照的阶数满足 $0 \leq l2 \leq 2$ ，当 $l1 > 4$ （反射率图中第1-4阶能量为零）时， $l \geq |l1 - l2| > 2$ ，即只有大于2阶的项。为了保证图像的第0、1和2阶等于对应阶的辐照度乘以反射率的直流分量 λ_{00} ，要求反射率图的 $\lambda_{lm} = 0, 1 \leq l \leq 4$ 。

虽然人脸表面并不严格满足上述假设，我们发现在实际应用中还是取得了很好的效果。人眼观察到的人脸皮肤颜色大致相同也证明了这一点。

由于光照参数的维数很少，在实际应用中，为了使输入更符合这一点，我们将图像的分辨率可以降低得很低，这样各点的反射率更接近，且反射率与皮肤差别较大的区域，如眉毛、眼睛、鼻子和嘴巴部分不参与光照估计运算。

3.5 光照估计的实验结果分析

为了评价光照估计的结果，我们选择光照严格控制的CMU-PIE人脸数据库 [Sim02]作为我们的实验数据。

CMU-PIE人脸数据库中共有68人，包含姿态、光照和表情的变化。本章我们选择其中不同光照下的正面人脸图像作为实验数据。CMU-PIE人脸数据库有21个不同方向的闪光灯和房间的背景光用来照亮人脸，通过控制这些光源的开或关，共产生43种不同的光照条件的人脸图像。针对光照变化，CMU-PIE人脸数据库中给出了两个子集：

“illum”和“lights”，其中前者只用闪光灯用来照亮人脸，每个人脸共有 21 幅图像 (f02~f22)，对应不同的闪光灯方向；后者既有闪光灯，也有背景光照亮人脸，每个人脸共有 22 幅图像(f01, f02~f22)，其中 f01 只有背景光。因此光照条件不仅包含点光源，还包括环境光+点光源的情况。光源、相机和人脸的相对位置如图 3-4 所示。由于我们暂时不考虑姿态变化，我们在实验中选取其中正面姿态下的图像，即由编号为 c27 的相机拍摄的图像。CMU-PIE 人脸数据库中“illum”和“lights”集合中的示例图像及其对应的闪光灯编号如图 3-5 所示。

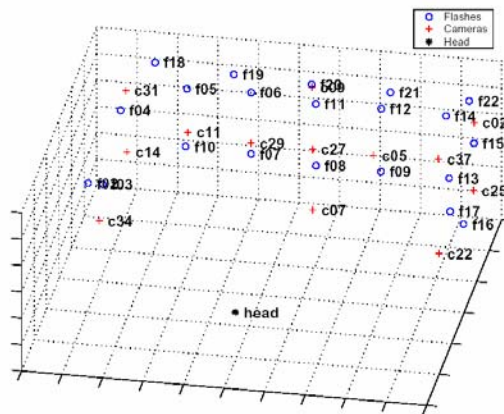


图 3-4: CMU-PIE 人脸数据库中在 3D 图中显示人脸、13 个相机和 21 个闪光灯的 xyz 坐标。

Figure 3-4. The xyz-locations of the head position, the 13 cameras, and the 21 flashes plotted in 3D to illustrate their relative locations in CMU-PIE face database.

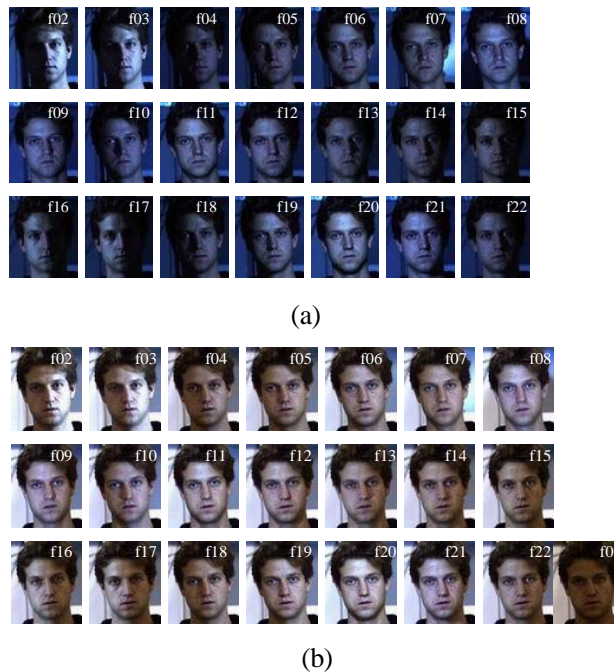


图 3-5: CMU-PIE 人脸数据库中正面姿态不同光照下图像示例。

(a) “illum”集合中的图像。(b) “lights”集合中的图像。

Figure 3-5. The frontal images under different illuminations from CMU-PIE face database. (a) The example images in “illum” subset. (b) The example images in “lights” subset.

3.5.1 光照估计视觉结果

对于朗伯反射表面，其成像公式也可写成

$$\begin{aligned} I(x, y) &= \int_{\theta'_i=0}^{\frac{\pi}{2}} \int_{\phi'_i=0}^{2\pi} \rho(\theta'_i, \phi'_i, \theta_o, \phi_o) dE_i(\theta'_i, \phi'_i) d\theta'_i d\phi'_i, \\ &= \lambda(x, y) \int_{\theta'_i=0}^{\frac{\pi}{2}} \int_{\phi'_i=0}^{2\pi} dE_i(\theta'_i, \phi'_i) d\theta'_i d\phi'_i, \\ &= \lambda(x, y) E(x, y), \end{aligned} \quad (3-29)$$

即光照对人脸图像的影响通过辐照度 E 实现，因此对光照估计的结果的度量，我们用 E 表示。比较式 (3-23) 和式(3-27)，并将辐照度图用角度索引，得到

$$E(\theta, \phi) \approx \sum_{l=0}^2 \sum_{m=-l}^l L_m A_l Y_{lm}(\theta, \phi). \quad (3-30)$$

图 3-6 给出了 CMU-PIE 人脸数据库中两个不同人脸在几种光照下的光照估计的结果（光照环境分别为“illum”集合中的 f02, f10、f22，和“lights”集合中的 f01, f02, f10 和 f22）。实验中图像的分辨率为图像中人脸区域实际大小，约为 164×200 。虽然两幅图中第一行和第二行中的人脸的形状和反射率差别较大，但从相同光照下的图像估计出的光照很相似，说明光照估计结果只与光照条件有关，不同人脸的影响很小。

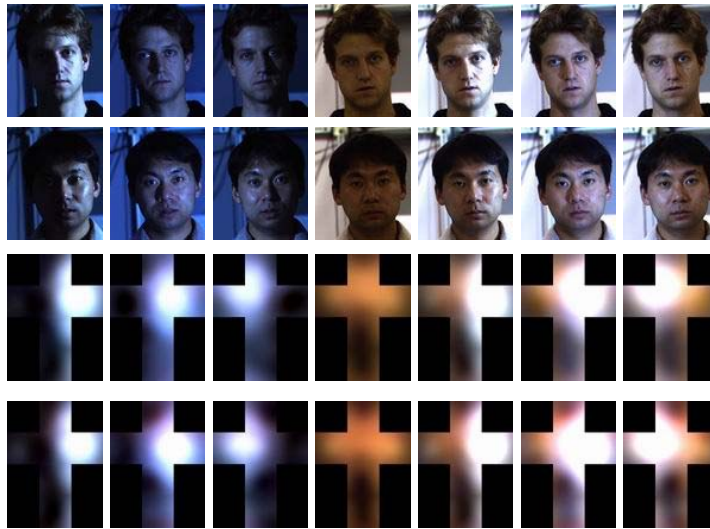


图 3-6: CMU-PIE 人脸数据库上光照估计的结果。

前两行为两个人脸在不同光照下的图像，后两行是其光照估计的结果，其中每一列的光照条件相同。虽然两个人脸差别较大，但相同光照下的光照估计结果非常相似。这里辐照度图以立方体的形式给出。

Figure 3-6. Some examples of the estimated irradiance maps from the CMU-PIE face database. The first two rows are the images of two persons and the last two rows are the estimated irradiance maps. Every column of the first two rows is captured under the same illumination. Evidently, the estimated irradiance maps are fairly similar though the faces in the first two rows are different. The irradiance maps are illustrated as cubic environments.

前面我们说过光照满足可见性，即分别在两个光照条件下拍摄的图像之和等于将这两种光照加在一起的光照下拍摄的图像。CMU-PIE 人脸库数据中包含只用背景光的图像、只用闪光灯的图像和两者都用的图像。图 3-7 第一行中给出了从图像角度验证光照可加性的结果。即只用背景光的图像与只用闪光灯的图像之和与“背景光+闪光灯”下的图像几乎相等。下面我们对同一人脸的上述三幅的图像分别进行光照估计，并计算“背景光+闪光灯”的光照估计结果与只有背景光光照估计结果之差。该差与只有闪光灯的光照估计结果作比较，从图 3-7 中可以看出二者非常相似。这证明了光照估计后的光照仍然满足光照可加性。这也从另一个侧面说明光照估计的结果是可信的。

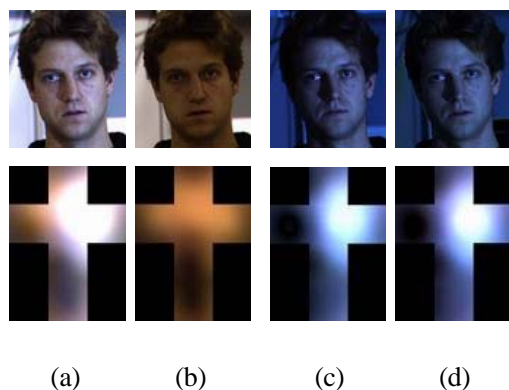


图 3-7：光照的可加性。

(a) 背景光+闪光灯。(b) 背景光。(c) (a)与(b)之差。(d) 闪光灯。(c)与只有闪光灯的结果(d)很相似，证明光照估计的结果满足可加性。

Figure 3-7. An example of the subtracting of the estimated irradiance maps from the CMU-PIE database. (a) Room lights with flash. (b) Room lights. (c) Difference of (a) and (b). (d) Flash only. The difference irradiance map is quite similar to the estimated irradiance maps of the same flash with the room lights off.

3.5.2 光照估计对配准精度和图像分辨率的敏感性

当光照条件不理想时，用 LTC-AAM 方法自动定位的特征点有时不十分准确。因此我们需要知道光照估计结果对特征点定位精度的敏感性。我们比较了精确配准和粗略配准下的光照估计结果。精确配准情况下特征点的位置先用 LTC-AAM 方法自动定位，再用手工调整。粗略配准中只用了两个眼睛的位置。而自动 LTC-AAM 标注的结果可视为介于二者之间。光照估计对配准精度的敏感性定义为

$$S = \frac{\sqrt{\sum_{i=1}^n \|E_{perfect}(i) - E_{loose}(i)\|^2}}{\sum_{i=1}^n E_{perfect}(i)}, \quad (3-31)$$

其中 n 为辐照度图中像素的个数， $E_{perfect}$ 和 E_{loose} 分别表示精确配准情况下和粗略配准情况下的光照估计的结果。

CMU-PIE 人脸库中各种光照条件下的平均敏感度在图 3-8 中给出, 其中精确对齐时图像的分辨率为图像实际大小, 约为 164×200 , 而粗略对齐时图像的分辨率为 64×64 。由于在原始图像分辨率的图像中, 可用像素数很多, 光照估计时也进行了下采样。从图中可以看出, 各种光照下敏感度都很小。这意味着光照估计的结果对配准精度不是很敏感, 也就是说在精确配准不可得的情况下, 我们可以用粗略配准来进行光照估计。

实验中我们还比较了不同分辨率对光照估计的影响。相应的, 我们定义光照估计对分辨率的敏感性定义为

$$S = \frac{\sqrt{\sum_{i=1}^n \|E_1(i) - E_2(i)\|^2}}{\sum_{i=1}^n |E_1(i)|}, \quad (3-32)$$

其中 n 为辐照度图中像素的个数, E_1 和 E_2 分别表示两种分辨率下的光照估计结果。

图 3-9 给出了分辨率为 32×32 相对分辨率为 64×64 时光照估计的敏感性。由于低分辨率时特征点精确定位的意义已不大, 这两种光照估计的结果都是在粗略配准情况下的结果。图中的值都很小, 说明光照估计对图像分辨率不敏感。

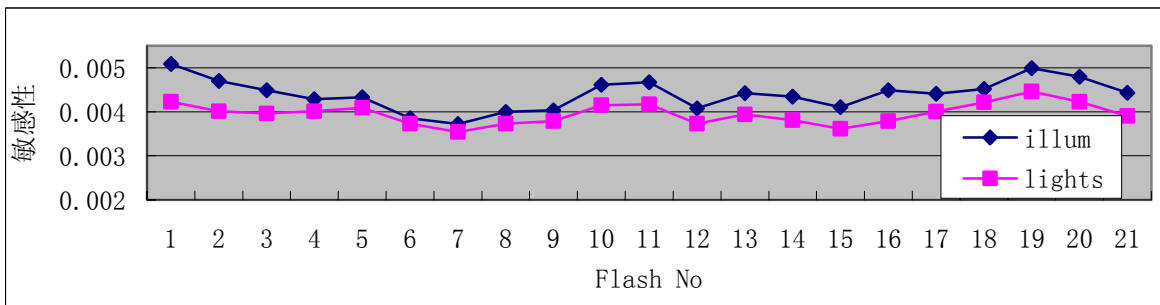


图 3-8: 光照估计对配准精度的敏感性。

Figure 3-8. The average sensitivity of the lighting estimation to alignment.

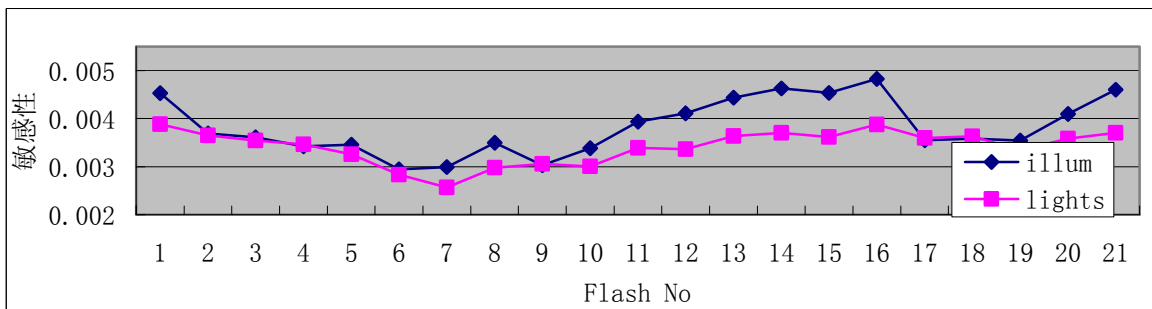


图 3-9: 光照估计对图像分辨率的敏感性。

Figure 3-9. The average sensitivity of the lighting estimation to resolution.

3.5.3 光照估计结果数值分析

光照估计的直接结果是光照的 9 个低频分量的系数。根据式 (3-4) 球面谐波函数 Y_{lm} 的表示以及式 (3-5) 中光照系数 L_{lm} 的计算, 可推断得到各光照系数的物理意义。表 3-1 列出了第 0 阶和第 1 阶各分量的光照系数的物理意义, 二阶以上的系数的物理含义稍微模糊一些, 但亦可按此推理。表 3-2 中的几个实例的实验结果也直观地验证上述结论。

表 3-1: 光照系数的物理意义。

Table 3-1. The physical meaning of the spherical harmonics coefficients the illumination.

光照分量系数	物理意义
直流系数 L_{00}	表示光照的平均能量
一阶系数 L_{1-1}	表示主光源的 Y 轴方向, 其符号为 + 表示主光源的主方向为 Y 轴正方向
一阶系数 L_{10}	表示主光源的 Z 轴方向, 其符号为 + 表示主光源的主方向为 Z 轴正方向
一阶系数 L_{11}	表示主光源的 X 轴方向, 其符号为 + 表示主光源的主方向为 X 轴正方向

表 3-2: 几个典型的光照估计结果及其说明。

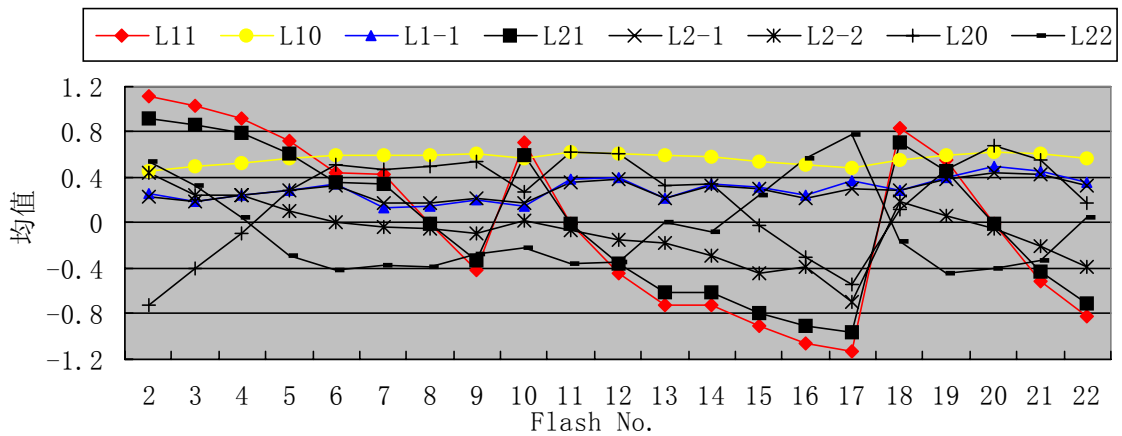
Table 3-2. Some typical results of illumination estimations.

输入图像	光照估计结果						直观解释
	辐照图	颜色通道	L_{00}	L_{1-1}	L_{10}	L_{11}	
		R	57	13	29	67	只有闪光灯, 主光源在 X 轴正方向
		G	66	11	32	82	
		B	86	11	42	101	
		R	45	22	29	2	只有闪光灯, 光源位于正上方
		G	51	21	33	2	
		B	74	27	48	3	
		R	95	46	55	10	背景光+闪光灯, 主光源位于正上方
		G	84	40	50	9	
		B	90	42	56	6	
		R	55	23	29	8	背景光, 主要能量集中在直流分量
		G	35	15	18	6	
		B	20	8	10	3	

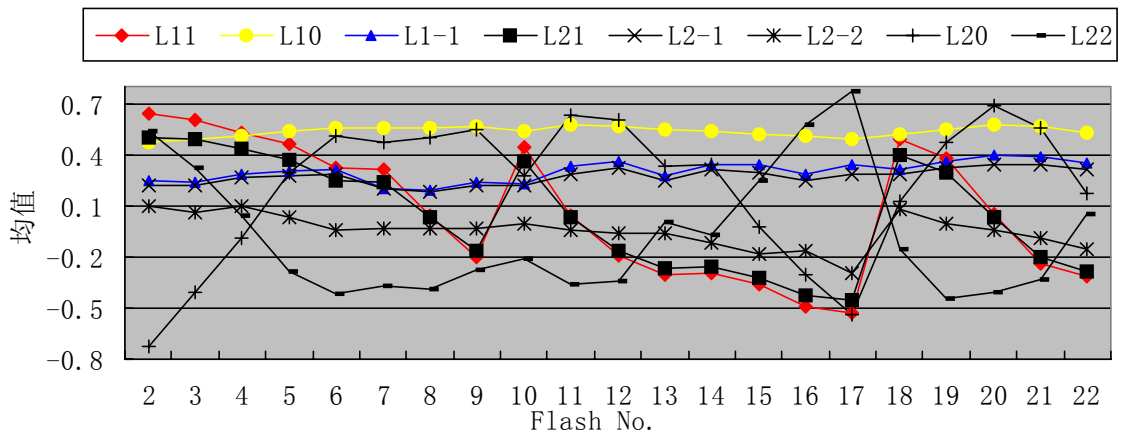
更多的实验结果如图 3-10 所示。图中分别给出了 CMU-PIE 人脸数据库中 “illum” 集合 (无背景光) 中和 “lights” 集合 (有背景光) 中各种光照下的光照系数的均值。为

了消除不同人脸平均反射率的影响，后 8 个系数都除以直流系 L_{00} 数作归一。这些结果同样也说明了上述结论。如在“illum”集合中，闪光灯编号为“f02”与编号为“f17”时，系数 L_{11} 的数值基本相等而符号相反，表示这两种光照的主光源方向相对 YOZ 平面对称，亮度相同。这与真实情况一致。另外有背景光时对应光照的各光照的系数比无背景光的情况小，说明背景光对直流分量贡献较大（各系数除以直流系数，所以反而变小了）。

根据光照系数的物体意义，可进一步得到光照系数与主光源偏光程度之间的关系：一阶光照系数与直流系数的比值（ L_{1-1}/L_{00} 、 L_{10}/L_{00} 和 L_{11}/L_{00} ）越大，表明在主光源在相应轴的方向偏光越严重。如闪光灯编号为“f02”与编号为“f17”时， L_{11}/L_{00} 较大，表明在这两种情况下，主光源在 X 方向偏光严重。



(a) “illum” 集合



(b) “lights” 集合

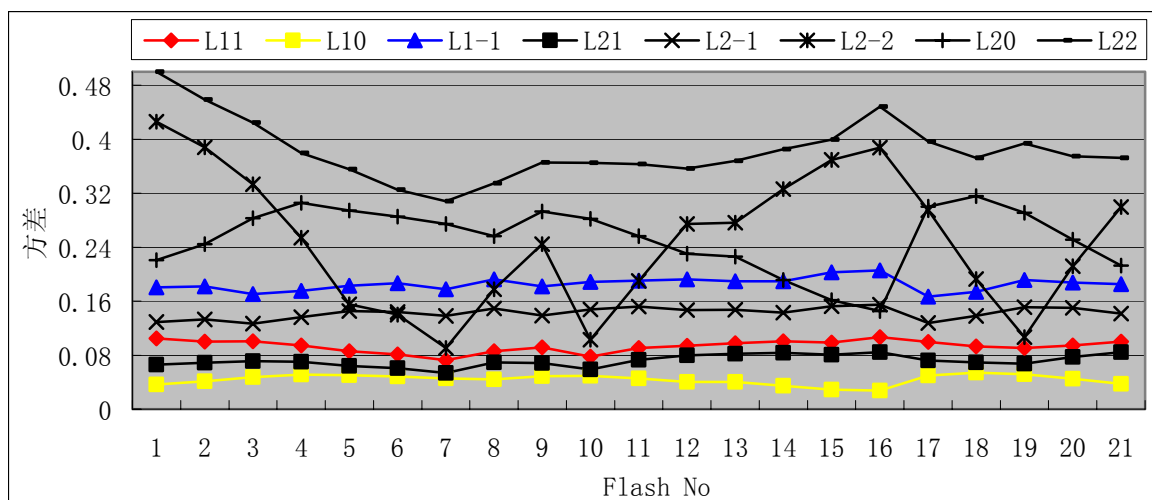
图 3-10: CMU-PIE 人脸数据库中各种光照下的光照系数的均值。

Figure 3-10. The mean of the coefficients of the illuminations in the “illum” set (a) and the “lights” set (b) of the CMU-PIE face database.

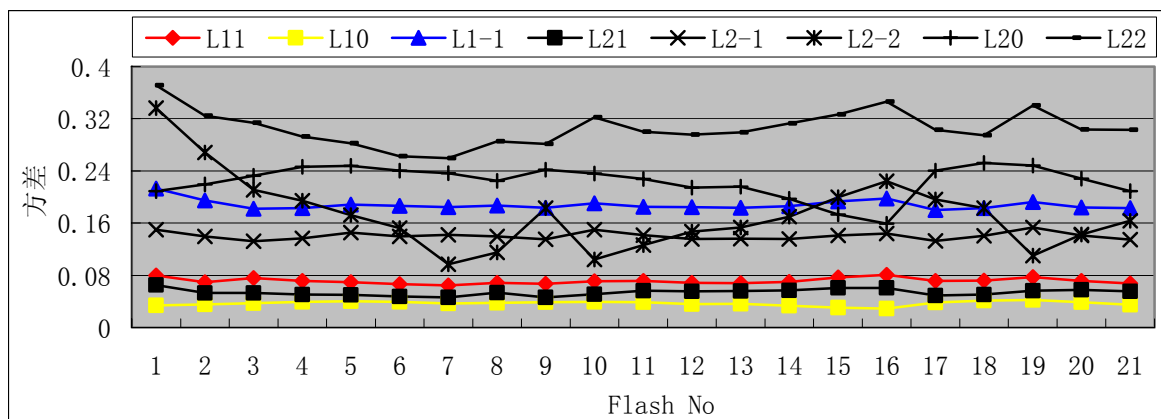
因此最后我们考察光照估计的稳定性。不同人脸在相同光照下的光照估计结果应该

比较稳定。我们通过两组数据来验证光照估计的稳定性：光照系数的方差和光照系数分类的结果。

CMU-PIE 人脸数据库中各种光照下不同人脸的光照估计的系数的方差如图 3-11 所示。图中各光照系数的方差不大，尤其是一阶的 3 个分量。如在各种光照下 L_{11} 的均值变化范围很大（如图 3-9(a)所示，从 1.2~1.2），但其方差一直保持在 0.09 左右，也就是说各种光照下光照估计的结果随人脸变化很小。8 个系数中有 3 个系数的方差较大，因为图像只是人脸正面的可见部分，因此图像中对人脸背后的光照的信息不够，导致光照估计的结果受噪声的影响较大。



(a) “illum” 集合



(b) “lights” 集合

图 3-11: CMU-PIE 人脸数据库中各种光照下的光照系数的方差。

Figure 3-11. The standard deviation of the coefficients of the illuminations in the “illum” set (a) and the “lights” set (b) of the CMU-PIE face database.

对 CMU-PIE 人脸数据库中的光照估计的系数进行分类的结果如图 3-12 所示。分类时，任选一人脸在各个光照下图像，根据这些图像估计得到其光照系数，作为类似人脸识别时的原型图 (Gallery)。其他人脸在各种光照的图像的光照估计结果作为测试数据

(Probe)用以光照分类。分类时相似性度量为两个光照系数串联而成的向量之间的相关系数。我们将离得最近的3个闪光灯组成一个集合，分类时结果为集合中任何一个闪光灯都认为是正确的。分类的结果不是100%正确，因为各个人脸毕竟有些差异，且不同人脸在相同光照下的图像中的姿态也略有不同。另外有些闪光灯离得确实太近。对错误分类的结果分析显示，所有错误分类的结果都是其邻近的光照情况，如“lights”集合中光照为“f06”的几个错分结果中，是将其分别错分为8个“f10”、3个“f20”和1个“f08”。从图3-3中各闪光灯的位置和图3-4中的人脸图像来看，“f06”确实离上述3个闪光灯都很近，对应光照下的人脸图像也很相似。而在其最近集合中只包含了{f05, f07, f19}。“lights”集合中光照分类的结果相对“illum”集合稍差一些，这是因为由于存在环境光，相邻闪光灯下的人脸图像更相似，所以分类时有更多的结果被错分到其邻近的光照。总的来说，光照估计的结果很稳定，结果主要随光照变化而变化，随人脸的变化很小。

光照估计主要是为了下一步的光照补偿。在此讨论光照分类问题是因为有些情况下我们可能只需知道光照的大体种类即可。因为一般的人脸识别算法对光照变化都有一定的容忍，在算法容忍的范围内可以不作光照补偿。而光照补偿本身作了一些近似和假设，会带来一定噪声。在这些情况下作光照补偿可能会得不偿失、矫枉过正。

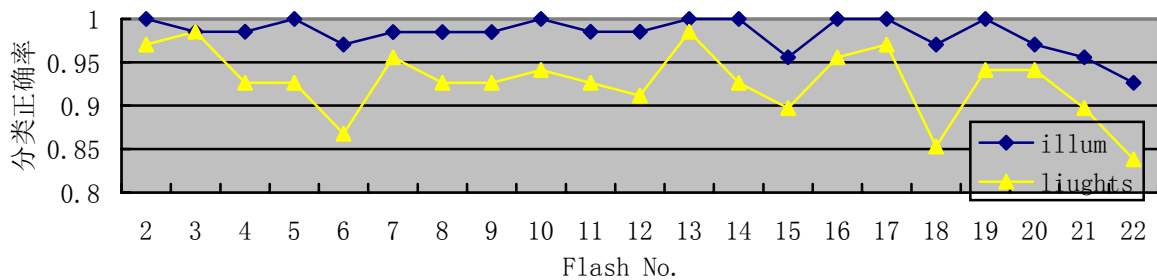


图 3-12: CMU-PIE 人脸数据库中各种光照下光照系数的分类结果。

Figure 3-12. The classification of the lighting estimation on CMU-PIE face database.

3.6 光照估计小结

根据球面谐波基图像理论，朗伯反射对光照而言是一个低通滤波器，即只有光照的低频部分对朗伯反射有影响。基于该理论，对朗伯反射，光照估计只需估计其低频系数即可。这样用少数几个（9个）参数就可以表示影响朗伯表面的任意类型的光照。

假设人脸表面为朗伯凸表面，并根据人脸类的先验知识：（1）不同人脸的形状基本相同和（2）人脸的各点反射率近似相等，我们能够根据一幅输入图像计算其球面谐波基图像。从而光照估计可以通过解一个最小二乘方程得到，即求解光照的9个低频分量的系数。

由于光照的参数很少，光照估计结果对配准精度和图像分辨率变化不敏感。为了能更有效地估计光照，我们还给出了尽量符合上述假设的操作：（1）形状用基于特征点的

变形；(2) 降低分辨率。尤其是降低分辨率后人脸的形状差异更小，使得光照估计不太依赖于特征点定位的结果。这对实用系统很重要，因为实用系统在光照变化剧烈时特征点的定位通常都不太准确。另外降低分辨率后，人脸表面各点的反射率差别也更小了，使得人脸表面纹理近似为常数的假设更合理。

不同光照下的真实人脸图像的实验结果也证实了本文中光照估计的准确性和稳定性。

第四章 基于球面谐波理论的人脸图像光照补偿

在光照估计的基础上,本章我们讨论对输入人脸图像进行光照补偿。(光照估计的另一个作用是告诉我们在合适的光照变化范围内可以不作光照补偿直接进行人脸识别。)与光照估计的全局过程不同,光照补偿是一个局部过程,也就是说光照补偿是对图像中每个像素单独进行,因此需要输入人脸特有的信息。这也是人脸识别的基础。

光照补偿是期望将输入人脸图像经过处理后得到一个与光照无关或至少对光照变化不敏感的表达,从而使得人脸图像的流形分布更简单,这样人脸识别变得更容易。这可以用一个统一的计算模型来表达:

$$F = T(I), \quad (4-1)$$

其中 I 表示原始图像, T 是削弱光照变化影响的变换, F 是变换后的特征表示,可以是图像或者其他特征。

本章是如下组织的。首先介绍本文提出的两种光照补偿方法:基于绘制的光照补偿和基于差图像的光照补偿,接着给出两种光照补偿方法的实验结果及其与其他一些经典的光照处理方法的对比实验结果,然后简单介绍了图像绘制在人脸检测中扩大训练样本中的应用。最后为本章小结。

4.1 基于绘制的光照补偿

因为光照变化对人脸图像影响很大,导致光照变化下人脸图像的内类差比类间差还大,从而人脸图像识别很困难。一个直觉的想法是将所有图像都绘制到相同光照下,然后再做识别。基于绘制的光照补偿就是基于该思想,将不同光照的输入人脸图像,都重新绘制到一种标准光照环境,从而使所有的人脸图像都在相同的光照条件下进行比较,以消除光照对人脸识别的影响。对应式(4-1),变换 T 为图像绘制过程, F 为标准光照下的图像。

4.1.1 基于绘制的光照补偿的原理

假设人脸表面为朗伯凸表面,如式(3-29),

$$I(x, y) = \lambda(x, y)E(x, y), \quad (4-2)$$

其中

$$E(x, y) \approx \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha(x, y), \beta(x, y)), \quad (4-3)$$

$\lambda(x, y)$ 表示图像中点 (x, y) 处的纹理, $(\alpha(x, y), \beta(x, y))$ 为该点的法线,可根据人脸的3D形状计算。这种表示形式的一个好处是可以将人脸图像对纹理 $\lambda(x, y)$ 的依赖与形状的依赖

分开，其中图像对形状和光照的依赖关系用辐照度 $E(x, y)$ 表示。

比率图方法在人脸图像绘制和表情合成中取得了很好的结果 [Shashua01] [LiuZC01][Wen03]。本文我们将比率图扩展到光照比率图。光照比率图定义为同一人脸在不同光照下的图像比。因此输入图像的光照相对标准光照的比率图为

$$R(x, y) = \frac{I(x, y)}{I_0(x, y)} = \frac{\hat{E}(\alpha(x, y), \beta(x, y))}{E_0(\alpha(x, y), \beta(x, y))}, \quad (4-4)$$

其中 I 和 I_0 分别表示和输入图像和在标准光照下 L_0 的图像， \hat{E} 和 E_0 分别表示输入图像的光照下的辐照度和标准光照下的辐照度。

标准光照下的图像可以通过基于光照商图像的绘制

$$I_0(x, y) = \frac{I(x, y)}{R(x, y)} \quad (4-5)$$

得到。

需要指出的是，由于光照比率图 $R(x, y) = \frac{\hat{E}(\alpha(x, y), \beta(x, y))}{E_0(\alpha(x, y), \beta(x, y))}$ 与物体的纹理无关，所

以如果假设人脸的形状相同（商图像方法 [Shashua01] 如是假设），也可以用一个人脸的光照比率图来合成另外一个人脸在标准光照下的图像。不过在我们的系统中利用的是同一个人的形状信息，得到的是特定人的光照比率图。

为了利用式 (4-5) 进行绘制，我们需要知道输入光照条件相对标准光照的光照比率图 R 。而计算 R 又需知道两种光照的辐照度图，即需知道两种光照和输入人脸的形状。其中标准光照可以预先定义，输入人脸图像的光照通过论文第三章中的方法估计得到，输入人脸的形状同第 3.4.1 节，通过将平均 3D 人脸变形到输入人脸来获得。第三章中光照估计对输入人脸特定的 3D 形状不是很依赖，用全局信息就可以估计出光照。而光照补偿则依赖于输入人脸表面每点的纹理和形状。在基于光照比率图的图像绘制中，我们不必显式地求出纹理，而是直接利用图像中隐含的纹理信息。而对形状，计算时需要知道人脸表面上每点的法线。因此下一小节我们将讨论不同的标准光照对输入人脸 3D 形状的依赖性。

4.1.2 标准光照

人脸识别时有用的信息为人脸的本质特征：纹理和形状。而二者的组合又有三种情况：形状、纹理、形状+纹理。由于我们在计算时输入人脸的形状信息是通过将平均人脸变形到输入人脸得到，且光照变化时很难得到精确定位的特征点，我们认为单独用这种形状信息不足以区分不同的人脸。因此本文主要考虑后两种情况。

根据式 (4-2)，人脸图像对形状的依赖用辐照度 $E(x, y)$ 表示，如式 (4-3) 所示，

$E(x, y) \approx \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha(x, y), \beta(x, y))$ 。因此 $E(x, y)$ 对形状的依赖又取决于球面谐波函数 $Y_{lm}(\alpha(x, y), \beta(x, y))$ 。根据球面谐波函数的多项式形式表示 (式 (3-4)), 当 $l=0$, Y_{00} 为常数 $c_0 = \frac{1}{2} \frac{1}{\sqrt{\pi}}$, 与形状无关; 当 $l=1, 2$ 时, Y_{lm} 为分别为人脸法向量的一阶多项式和二阶多项式, 函数值取决于具体的自变量 (法向量)。在 $E(x, y)$ 的表达式中, 球面谐波函数 Y_{lm} 对辐照度 $E(x, y)$ 的贡献还与权重 A_l 有关。而 $A_1 > A_2$ ($A_1 = \frac{2\pi}{3}, A_2 = \frac{\pi}{4}$), 因此辐照度 $E(x, y)$ 对形状的依赖主要为一阶分量, 即标准光照中一阶能量越大, 其辐照度的一阶能量越大, 则物体在该标准光照下的图像更依赖于物体的形状。

因此, 若标准光照只包含直流分量 L_{00} , 即当 $l \neq 0$ 时, 其他分量 $L_{lm} = 0$, 则此标准光照下的人脸图像为

$$I_0(x, y) = \lambda(x, y) \sum_{l=0}^2 \sum_{m=-l}^l A_l L_{lm} Y_{lm}(\alpha(x, y), \beta(x, y)) = c_0 \lambda(x, y) L_{00} A_0 \quad (4-6)$$

此时标准图像与输入人脸的法线 $\alpha(x, y), \beta(x, y)$ 无关, 即与人脸的形状无关, 而只与人脸的纹理 $\lambda(x, y)$ 相关。只包含直流分量的光照在物理上是均匀光照, 各个方向的光照强度相等。均匀光照的辐照度图中前三阶的能量比可视为 1.0: 0.0: 0.0, 由表 4-1(a) 给出。

真实环境中一种比较典型的光照是“环境光+一个主光源”, 这也是很多视觉系统假设的光照环境。CMU-PIE 人脸数据库[Sim02]中的“lights”集合中的图像就是在这种光照条件下拍摄的。根据 CMU-PIE 人脸数据库中的“lights”集合中闪光灯编号为“f11” (正面点光源) 中的一幅图像估计出的光照结果如表 4-1(b) 所示。该光照辐照度图中前三阶的能量比为 0.8: 0.18: 0.02, 其中第 l 阶能量计算为 $\sum_{m=-l}^l \|A_l L_{lm}\|^2$ 。与均匀光照相比, 该

光照环境中一阶分量中包含了更多的能量, 因此其标准形式中包含了更多的形状信息。

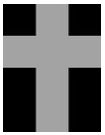

在定义了标准光照之后, 我们就可以根据式 (4-2) 计算光照比率图 R 和式 (4-3) 将输入图像绘制到标准光照, 得到标准图像 I_0 。基于绘制的光照补偿的一些示例图像如图 4-1 所示。从图中可以看出, 输入的相同人脸不同光照下的图像经过光照补偿后变得更相似了, 这表明该处理确实起到了光照补偿的效果。同时光照补偿后不同人脸之间的差异仍然得到了有效保持, 这样人脸识别的依据仍然得到了保留。

由于均匀标准光照绘制时只有纹理信息而不包含输入人脸的形状信息, 我们期望这种光照补偿方式不会对特征点对齐特别敏感; 而环境光+正面点光源绘制包含的形状信

息，因此最终的补偿效果会比较依赖于特征点的精确定位。反之，如果特征点定位准确，我们期望标准光照为环境光+正面点光源时，效果会更好。

表 4-1：两种典型的标准光照。

Tab. 4-1. The typical canonical illuminations.

	辐照度图	9 个光照系数									
			L_{00}	L_{11}	L_{10}	L_{1-1}	L_{21}	L_{2-1}	L_{2-2}	L_{20}	L_{22}
均匀光		R	100	0	0	0	0	0	0	0	0
		G	100	0	0	0	0	0	0	0	0
		B	100	0	0	0	0	0	0	0	0
环境光+正面点光源		R	103	0.4	59	38	0.4	33	12	16	36
		G	89	-0.1	53	34	-0.1	32	7	26	37
		B	100	-2	62	39	-1	37	6	52	15

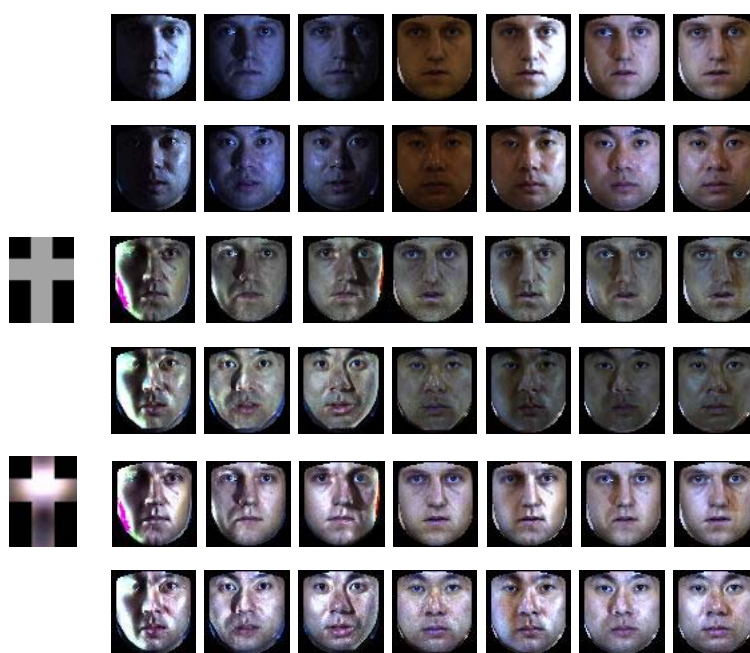


图 4-1：基于绘制的光照补偿结果。

前两行为两个人脸在不同光照下的输入图像，中间两行和最后两行分别为前两行中对应图像绘制到均匀标准光照和环境光+正面点光源标准光照下的图像。

Figure 4-1. The results of re-lighting. The first two rows illustrate the cropped facial images. The next four rows contain the corresponding re-rendered images under the uniform illumination and under the room lights with a frontal flash, respectively.

4.1.3 后处理

由于数字图像的数值范围有限（0~255）且基于光照比率图像的绘制涉及除法，所以当光照比率图 R 中数值很小，即像素对应的 \hat{E} 很小时（如偏光严重时），会引入噪声，对人脸识别也会有干扰。我们称这些像素点为坏点（bad pixels），同统计意义上的孤立点（Outliers）。

一些重新绘制的算法[WangHT03]通过将原型图像集合（Gallery）中的图像打光到测试图像（Probe）的光照以回避这个问题，因为原型图像集合中的图像光照条件通常较好，包含信息较多，所以绘制后得到的图像噪声较少，从而人脸识别性能也会更好一些。图 4-2 给出了用本文方法将原型图像加光到其他图像光照的一些实例，从图中可以看出，重新绘制得到的图像中没有太多噪声，与真实图像很相似。但是这种策略每处理一个测试图像就要对所有原型图像集合中的图像作一次绘制，这对实用人脸识别系统来说是不可接受的。



图 4-2: 更多的重新绘制的例子，将光照条件相对较好的图像绘制到任意光照。

图中的一行为输入光照条件较好的图像。第二行为不同光照下的输入图像。第三行为将第一行中的输入图像绘制到第二行的图像对应的光照得到的图像。可以看出第三行的图像与第二行对应的图像很相似。

Figure 4-2. More results of re-lighting. The first row is the input facial images under the ideal illumination. The second row is the real facial images taken under different lighting conditions and the third is the corresponding re-rendered images (re-rendering the input image to the illumination of the corresponding image in the second row). The corresponding images in the last two rows look similar.

本文中我们将所有图像都绘制到相同的标准光照，原型图像集合中的图像只需在注册时绘制一次，因此不会影响后续的人脸识别部分的速度。但测试图像的光照条件千变万化，有时光照条件很恶劣，被光照遮挡的部分包含信息很少。从信息较少的输入中要得到较多信息，通常只能用先验猜测，会引入较多噪声（如图 4-1 中前 3 列）。这些坏点的纹理可以用全局统计信息（所有人脸纹理的概率分布）或局部统计信息（输入人脸的纹理分布）[Wen03][Hertzmann01]来限制和校正，以得到更好的视觉效果。由于合成的

纹理主要是统计学习得到的信息，没有为识别增加额外的有用信息，因此在识别时，本文只是简单地抛弃这些像素。这和部分人脸识别类似，只是这种遮挡是光照遮挡。但如果识别时用的是全局特征（如 PCA），上述校正过程还是很有必要的。

这些坏点的检测通过计算 \hat{E} 实现。 \hat{E} 中的像素值过低的点一定在阴影区域即为孤立点。一些例子如图 4-3 所示。但在有些情况我们可以不必作坏点检测操作，如在图 4-1 中每行的后 4 幅图像，标准图像中几乎没有噪声点。我们用 9 个光照系数之间的关系来判断。如果直流系数的能量比一阶 3 个分量能量和相对较大，则不必处理。因为黑暗的阴影通常是由于一阶分量引起的。例如，“illum”集合的直流能量与一阶能量之比为 2.53，而“lights”集合中该比值为 4.53。所以我们没有必要对“lights”集合中的图像检测坏点。

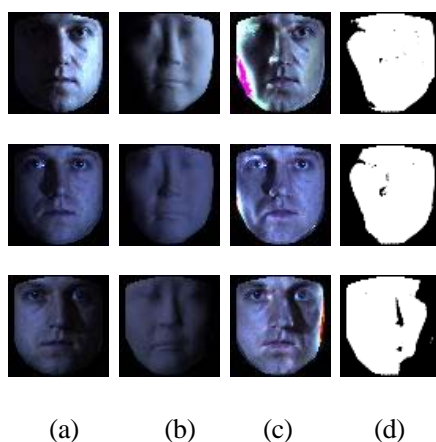


图 4-3: 坏点检测。

(a)输入图像，(b)辐照度 \hat{E} ，(c) 加光到虚拟均匀光照的图像，(d) 检测到的坏点（黑色像素）。

Figure 4-3. Outliers detection. (a) The input images. (b) The estimated face part of the irradiance environment maps. (c) The re-rendered images under the uniform illumination. (d) The detected outliers marked with black pixels.

4.2 差图像光照补偿

如上一节所述，当输入图像的辐照度图中的像素值较低时，基于绘制的光照补偿在图像绘制时会引入噪声。另外均匀标准光照绘制时只有纹理信息，不包含输入人脸的形状信息；而环境光+正面点光源绘制包含的形状信息又依赖于特征点的定位。众所周知，光照变化对特征点的准确定位至今仍是一个很大的挑战。因此本小节我们考虑另外一种光照补偿方式，既能保留形状和纹理信息，同时又对配准精度不敏感。

回到式 (4-2)， $I(x, y) = \lambda(x, y)E(x, y)$ ，其中 $\lambda(x, y)$ 表示图像中点 (x, y) 处的纹理， $E(x, y)$ 为该点对应的辐照度。而辐照度 $E(x, y) \approx \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} A_l Y_{lm}(\alpha(x, y), \beta(x, y))$ ，依赖于光照和输入人脸的形状。

辐照度 E 中只包含低频信息，滤去图像中的低频信息就能将光照的影响去掉。进行滤波操作最合适的方法是将 $I(x, y) = \lambda(x, y)E(x, y)$ 中的乘法变成 \log 域中的加法。对式 (4-2) 两边进行 \log 运算，得到

$$\log I(x, y) = \log \lambda(x, y) + \log E(x, y)。 \quad (4-7)$$

这样， \log 域中输入图像与辐照度图差为纹理图像，即

$$\log \lambda(x, y) = \log I(x, y) - \log E(x, y) \quad (4-8)$$

可以看成是光照不变的特征。

有意思的是，将式 (4-8) 两边取指数运算，则又变成原图像域，纹理图像为

$$\lambda(x, y) = \frac{I(x, y)}{E(x, y)}。 \quad (4-9)$$

将第 4.1 节中标准光照为均匀光照的图像绘制变形，可得到

$$I_0(x, y) = \frac{I(x, y)}{R(x, y)} = I(x, y) \frac{E_0(x, y)}{E(x, y)} = \frac{I(x, y)}{E(x, y)} \times cL_{00}， \quad (4-10)$$

其中 $c = \lambda_{00}A_0Y_{00}$ 为常数。比较式 (4-10) 和式 (4-9)，两式只相差一个常数倍。因此式 (4-8) 中的纹理图像和标准光照为均匀光的图像绘制可以看成是等价的，即纹理图像中没有形状信息，只包含纹理信息。但辐照度图 $E(x, y)$ 的计算还需要知道输入人脸的形状。

由于输入人脸的精确形状很难得到，所以下面的算法中我们不显示地使用输入人脸的形状，只利用编码在输入图像中的形状信息。另外我们希望在处理之后的图像中不同输入人脸形状不同的信息还得到保留。

用平均人脸模型来代替式 (4-8) 中的输入人脸的形状是一个很好的选择：

$$\log \lambda'(x, y) = \log I(x, y) - \log E_{avg}(x, y)。 \quad (4-11)$$

其中 E_{avg} 是平均人脸在输入光照下的辐照度。由于 E_{avg} 的低频性质和 E 类似，上式中的 $\log \lambda'(x, y)$ 受光照的影响也很小。不同的是，上式中 E_{avg} 是用平均人脸形状计算的，所以最后处理得到的图像 $\log \lambda'(x, y)$ 中，不仅包含了输入人脸的纹理信息，同时输入人脸形状同平均人脸形状的差异也得到了保留。上式一个直观的解释是：将输入图像 (\log 域) 减去一个只与光照有关的成分 (平均人脸在输入光照下的辐照度)，最后得到的结果图像中没有了光照的影响，而同时输入人脸的形状和纹理信息没有丢失。

需要指出的是，由于纹理图像和标准光照为均匀光的图像绘制是等价的，当粗略配准时 (只有两个眼睛位置给定)，式 (4-11) 与和标准光照为均匀光的图像绘制也是等价的。

由于二者等价，当输入光照对应的辐照度的亮度过低时，均光光照绘制时遇到的噪声问题，在式 (4-11) 也存在。在实验中我们也确实发现这两种等价的处理方法的人脸

识别结果几乎相同。为了抑制噪声，我们退而求其次，用图像域中的减法来代替式(4-11)中 \log 域的减法，得到

$$I_{diff} = I(x, y) - E_{avg}(x, y)。 \quad (4-12)$$

如果将平均人脸的纹理看成为常数1，则 $E_{avg}(x, y)$ 可以看成为 I_{avg} ，为平均人脸在输入光照下的图像。这样将

$$I_{diff}(x, y) = I(x, y) - I_{avg}(x, y) \quad (4-13)$$

定义为差图像更合适一些。

图 4-4 给出了一些不同光照下的图像、平均人脸在输入光照下的图像及二者的差图像。其中平均人脸的形状用 3D 人脸库[Blanz99]中的形状平均值表示，其纹理假设为常数 1，差图像可视化为差的绝对值。从图中可以看出，差图像中亮度值较大的点主要是有纹理的区域，如眼睛。由于总的亮度值很低，人眼睛可能分辨不出不同人脸之间的差别。但第 4.3 节中的人脸识别的结果表明，差图像方式的光照补偿对光照变化下的人脸识别的性能还是带来了很大提高。

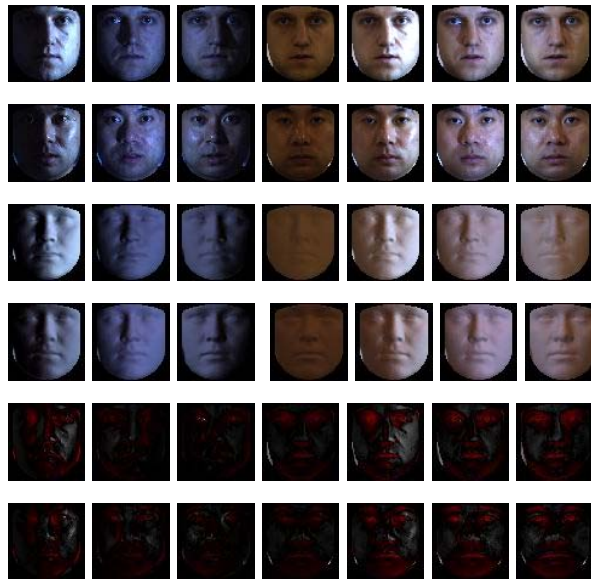


图 4-4：差图像光照补偿实验结果。

前两行为原图像，第三、四行为其对应光照下平均人脸的图像，最后两行为对应的差图像。

Figure 4-4. The results of illumination compensating based on the difference image. The first two rows are the original input images of two persons under different illuminations. The middle two rows are the images of the mean face under the corresponding illuminations and the last two rows are the corresponding difference images.

4.3 光照补偿对配准精度的敏感性

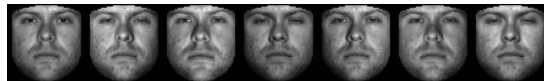
本小节中我们给出光照补偿对人脸识别的实验结果。如前面讨论的那样，人脸识别

需要准确的纹理信息和形状信息，本文给出的三种光照补偿的区别也在于对形状信息的处理上。而形状信息又依赖于特征点的定位。所以这一节我们主要考虑这些光照补偿方法对人脸识别的影响和对配准精度的敏感性。需要说明的是差图像方法在光照补偿时不需要精确配准（光照估计可以在不同的配准精度下进行）。

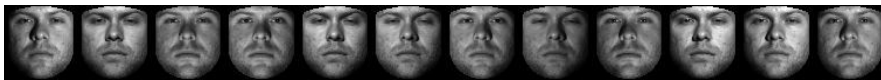
4.3.1 数据的选择

本节中我们选择 Yale B 人脸数据库[Belhumeur97]和 CMU-PIE 人脸数据库[Sim02]来比较本文的三种光照补偿方法。由于暂时没有考虑姿态问题，实验中我们选取的都是正面人脸图像。Yale B 人脸数据库和 CMU-PIE 人脸数据中一些典型的图像分别如图 4-5、图 3-5 所示。

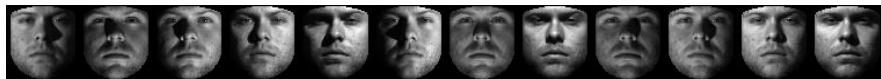
Yale B 人脸数据库是较早发布的一个专门针对光照、姿态和表情变化采集的人脸数据库，其中共有 10 人，64 中光照、9 个姿态的图像下的照片。光照角度从(0,0)变化到($\pm 130, \pm 90$)。该库中 64 种光照又分为 5 个子集如图 4-6 所示。很多算法都在 Yale B 人脸数据库上报告有实验结果，且该库根据光照变化程度分成了几个子集，可以更清楚地观察算法对不同光照变化程度的适应性。



Subset 1: 变化角度 $\leq 12^\circ$ ，共 7 幅图像/人



Subset 2: 变化角度 $\leq 25^\circ$ ，共 12 幅图像/人



Subset 3: 变化角度 $\leq 50^\circ$ ，共 12 幅图像/人



Subset 4: 变化角度 $\leq 77^\circ$ ，共 14 幅图像/人



Subset 5: 变化角度 $\leq 130^\circ$ ，共 19 幅图像/人

图 4-5: Yale B 人脸数据库中光照变化的情况。

Figure 4-5. The illumination variations in Yale B face database.

CMU-PIE 人脸数据库[Sim02]自从 2002 年在 FGR 会议上发布后,被很多算法采用。其一个明显的特点是包含的人数更多,因此更能说明实验结果的统计性。另外 CMU-PIE 人脸数据库中的“lights”集合中还包含了既有背景光又有主光源的光照条件。这种光照条件更接近我们日常生活中的光照环境且该光照条件在 Yale B 人脸数据库中没有。

在进行识别之前,我们需对输入图像进行剪裁,以去掉背景和发饰等对人脸识别的影响。图像剪裁是根据人脸图像中两个眼睛的位置,然后再加上一个掩码,过程如图 4-6 所示。实验中所有算法的裁减方式相同。

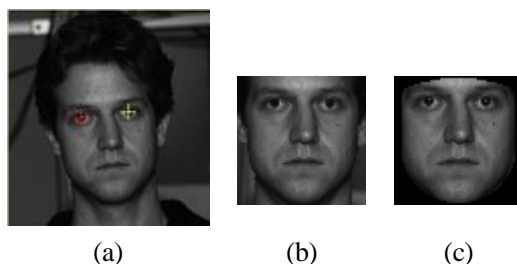


图 4-6: 人脸图像区域裁减。

(a) 输入人脸图像及其两个眼睛位置; (b) 裁减得到的人脸区域部分; (c) 加上掩码之后的人脸图像。
Figure 4-6. The crop of the facial image. (a) The input facial image and the location of the two eyes. (b) The cropped face part. (c) The masked face part with a mask.

4.3.2 人脸识别实验结果

由于光照补偿通常是作为人脸识别的预处理。我们期望经过光照补偿之后,相同人脸在不同光照下的相似度会有很大提高,同时不同人脸之间的相似度较小的性质还得到保留。归一化的相关系数(Normalized Correlation)可认为是最直接的相似性度量方式。所以在本节的实验中我们采用这种简单的相似性度量方式来进行人脸识别。

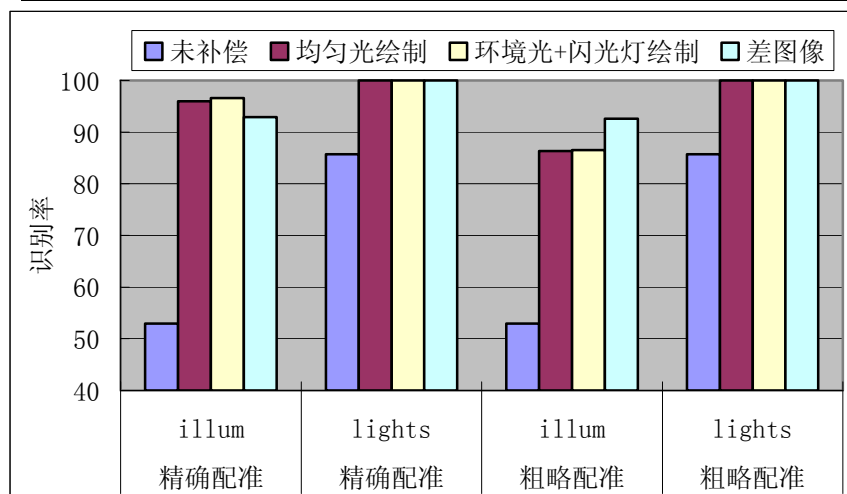
在 CMU-PIE 人脸数据库上的实验中,对“illum”和“lights”集合,我们选取不同的图像集作为原型图像集合。原型图像集合选择的原则是原型图像和测试图像之间只有光照变化。即对“illum”集合中的图像进行识别时,原型图像集合选“illum”中的 f11;对“lights”识别时,原型图像集合选“lights”中的 f11。f11 的主光源位置对人脸相对是正面的。

不同光照补偿方式在 CMU-PIE 人脸数据库上的实验结果在表 4-2 中给出。表中精确配准是指特征点准确定位,而粗略配准是只给定两个眼睛的位置。

表 4-2: 不同光照补偿方法及不同配准精度在 CMU-PIE 人脸数据库上的识别率比较。

Table 4-2. Recognition rate comparisons between the different algorithms with different alignments on the CMU-PIE face database.

光照补偿方法	识别率 (%)			
	精确配准		粗略配准	
	illum	lights	illum	lights
未补偿	52.9	85.7	52.9	85.7
均匀光绘制	96.0	100	86.3	100
环境光+闪光灯绘制	96.6	100	86.5	100
差图像	92.9	100	92.6	100



从表 4-2 我们可以得到如下结论:

(1) 各种光照补偿后, 人脸识别率有了很大提高;

(2) 在基于绘制的光照补偿中, 两种标准光照下人脸识别率差别不大, 这说明主要的有用识别信息为纹理信息; 但总的来说, 标准光照为“环境光+闪光灯”时, 识别率稍高, 说明形状信息还是有些帮助;

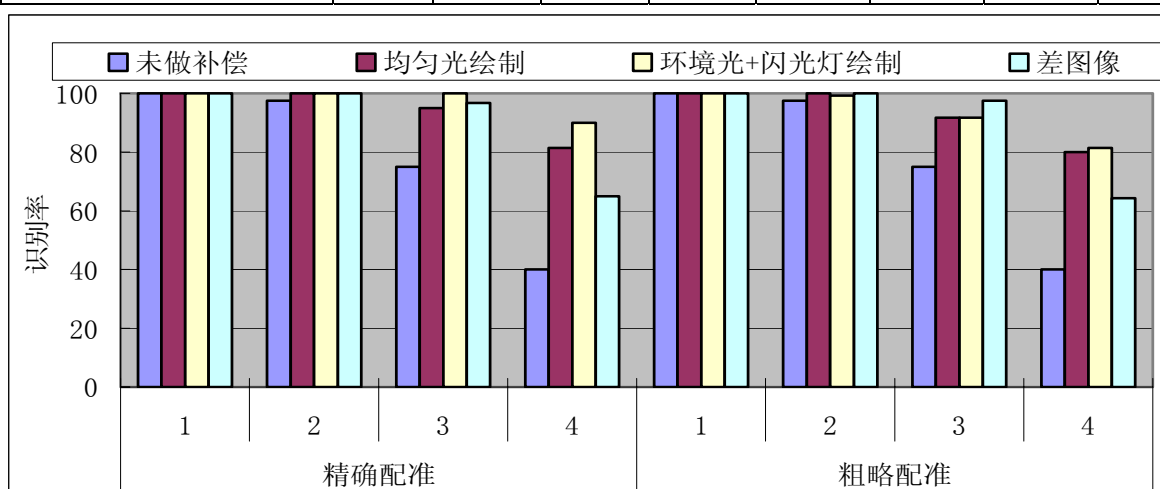
(3) 在基于绘制的光照补偿中, 粗略配准的识别率比精确配准的结果差, 说明在基于绘制的光照补偿中, 特征点的精确定位还是很重要。但即使粗略配准的情况下, 光照补偿后人脸识别的结果也大大提高了。因为差图像方式并不需要精确配准, 差图像在两种配准精度下的结果相差不大, 说明差图像光照补偿方式对配准精度不敏感, 且在粗略配准的情况下, 差图像方式的效果比基于绘制的光照补偿方法的结果更好一些。但精确配准情况下, 差图像方式的结果比基于绘制的光照补偿效果稍差。

对 Yale B 人脸数据库, 原型图像集合为正面点光源(0,0)下的图像, 测试图像集合为其余图像。不同光照补偿方式在 Yale B 人脸数据库上的实验结果如表 4-3 所示。

表 4-3: 不同光照补偿方法和不同配准精度在 Yale B 人脸数据库上的识别率比较。

Table 4-3. Recognition rate comparisons between the different algorithms with different alignments on the Yale B face database.

光照补偿方法	识别率 (%)							
	精确配准				粗略配准			
	Subset No.				Subset No.			
	1	2	3	4	1	2	3	4
未做补偿	100	97.5	75.0	40.0	100	97.5	75.0	40.0
均匀光绘制	100	100	95.0	81.4	100	100	91.7	80.0
环境光+闪光灯绘制	100	100	100	90.0	100	99.2	91.7	81.4
差图像	100	100	96.7	65.0	100	100	97.5	64.3



Yale B 人脸数据库上的识别结果趋势同 CMU-PIE 人脸数据库的结果类似, 精确配准时, 标准光照为“环境光+闪光灯”时的图像绘制的方法识别率最高。而在粗略配准时, 差图像光照补偿方式结果最好。但当光照条件比较恶劣时 (Subset4), 差图像的结果差一些。一个可能的原因是人脸表面不满足凸表面假设, 在 Subset4 中, 投射阴影的影响较大, 产生一些很强的边缘。而这些边缘对差图像的影响很大, 因为差图像中没有低频信号。

总的来说, 与需要多幅图像作原型图像的一些方法, 如光照锥[6]和 9PL [8] 相比, 我们的实验结果要差一些。但与同样只需一幅图像做原型图像的方法, 如光照归一 [WangHT03], 我们的实验结果更好。

4.4 与其他一些经典的光照处理方法的比较

这一节中我们将本文提出的几种光照补偿算法与其他一些经典的光照处理算法作比较。由于实用系统中通常只能比较准确地得到两个眼睛的位置, 所以本小节的结果都是在只有两个眼睛配准情况下的实验结果。

为了公正地与各个算法相比较, 本文我们用统一的测试数据集评测各个算法。本节

中我们一共选择了三个测试数据集，Yale B 人脸数据库、CMU-PIE 人脸数据库和 BVC2004 人脸数据库。由于暂时没有考虑姿态问题，实验中我们选取的都是正面人脸图像。前两个数据集前面已经介绍过了。BVC2004 人脸数据库中人数更多，共有 100 人，每人 10 幅不同光照条件下的图像（每个人的 10 种光照条件不尽相同，所以总的光照变化模式更多）。人脸图像在光照变化的同时还有表情变化。该库的光照采集条件相对比较松弛，光照控制不如前两个数据库那么严格，背景光的影响较大，有些甚至可以看成是在室外环境下拍摄的（室外光照是公认的一个很困难的问题），因而这种采集条件更接近实用系统，可以考察各种算法在可能的实用环境中的性能。BVC2004 人脸数据库的一个例子如图 4-7 所示。



图 4-7: BVC2004 人脸数据库中一个人脸图像示例。

Figure 4-7. The images of one person from BVC2004 face database.

4.4.1 待比较的光照处理算法

我们选择的几种待比较的光照处理算法如表 4-4 所示。在此我们不再详细介绍各个算法，因为这些算法在第二章和相应文献中已经介绍得很详细了。在此我们只是介绍各个算法中可能涉及的参数选择。在选择待比较的算法时，我们尽量选择不同类型的方法，如表 4-4 中前两个算法，直方图均衡化和 Gamma 灰度校正是最常用的光照预处理方法，接下来的比率图方法与边缘有关。对称 SFS 方法可以看成是基于模型的方法，而相位图是在 2004 年的 ICPR 会议上被提出来用作光照不变的人脸识别。另外一些比较经典的算法在注册时需要多幅图像或输入人脸的 3D 形状，如光照锥[Georghiadess01]、线性子空间[Basri01a]、9 个点光源[Lee01]，或者是其中涉及较复杂的过程不好控制和再现，如统计 SFS[Sim01]和统计谐波基图像[ZhangL03]需要生成很多虚拟样本重新训练，而不在我们的比较之列。另外梯度角度方法[Chen00]和高图像方法[Shashua01]都需要训练集，其性能会依赖于训练集的选择。这样将他们与别的方法相比较有失公平，所以本文也暂不涉及。

下面我们具体讨论各个算法涉及的参数，没有可选参数的算法在此略过：

(1) Gamma 灰度校正：需定义一个“标准”人脸图像 I_0 [Shan03]。本节中我们选用的标准图像同基于绘制的光照补偿相同，如 4-8 所示，其光照为环境光+一个正面点光源。

(2) 比率图：高斯平滑的高斯核 $\sigma = 0.5$ 。

(3) 对称 SFS：两个参数为人脸的 3D 形状和光照估计算法。本章中人脸的 3D 形状我们采用 Vetter 等采集的 3D 人脸库中提供的平均人脸模型[Blanz99]，光照估计采用本

文第三章的光照估计算法，后面再增加如果光源为点光源的话，由 9 个光照频域系数得到参数 P_s, Q_s 的过程。这个过程表示为：

$$P_s = \frac{-L_{11}}{-L_{10}}, \quad Q_s = \frac{-L_{1-1}}{-L_{10}}. \quad (4-14)$$

各个算法对输入原图像处理后得到的形式如图 4-9 所示。

表 4-4：待比较的光照处理算法。

Table 4-4. Some typical illumination processing algorithms compared in the thesis.

算法名称	参考文献	发表测试集
直方图均衡化 (HE)	特别常用的算法	
Gamma 灰度校正 (GIC)	[Shan03]	Yale B DB 和 Harvard DB
比率图 (Ratio Image)	[Belhumeur98]	Harvard DB
对称 SFS (SSFS)	[Zhao00b]	Yale DB, Weizman DB
相位图 (Phase spectrum)	[Savvides04]	CMU-PIE DB
均匀光照绘制 (Relighting1)	本文第 4.1 节	CMU-PIE DB 和 Yale B DB
环境光+正面点光源绘制 (Relighting2)	本文第 4.1 节	CMU-PIE DB 和 Yale B DB
差图像 (Diff)	本文第 4.2 节	CMU-PIE DB 和 Yale B DB



图 4-8：Gamma 灰度校正算法中用的标准图像。

Figure 4-8. The canonical images used in Gamma Intensity Correction.



图 4-9：各种光照预处理算法的直观效果图。

图中的图像按顺序分别为：原始图像、直方图均衡化后的图像、Gamma 灰度校正后的图像、Log 域中的边缘图、对称 SFS 的得到的图像、相位图、标准光照为均匀光照下的图像、标准光照为环境光+正面点光源下的图像和差图像。

Figure 4-9. The results of the lighting processing algorithms. The images are the input image, the image by HE, the image by GIC, the edge image in log domain, the image by SSFS, the phase spectrum, the re-rendered image under uniform illumination, the re-rendered image under a frontal flash with environment illumination and the difference image in order.

4.4.2 对比实验结果

人脸识别时大多数方法我们都是对光照处理后的特征用归一化的相关系数作为两幅图像之间的相似度量，只是其中比率图方法的距离/相似性度量不是相关系数，而是欧氏距离。

表4-5、表4-6和表4-7分别给出了各种光照处理算法在 Yale B、CMU-PIE 和 BVC2004 人脸数据库上的实验结果。

表 4-5: 各种方法在 Yale B 人脸数据库上的实验结果。

Talbe 4-5. The recognition rates of the algorithms on the Yale B face database.

算法名称	识别率 (%)			
	Subset No.			
	1	2	3	4
不做光照补偿 (No)	100	97.5	75	40
直方图均衡化 (HE)	100	95.8	75.8	41.4
Gamma 灰度校正 (GIC)	100	95.8	75.8	49.3
比率图 (Ratio)	98.3	100	99.2	82.1
对称 SFS (SSFS)	93.3	82.5	53.3	25
相位图 (Phase)	100	97.5	88.3	63.6
均匀光照绘制 (Relight1)	100	100	91.7	80
环境光+正面点光源绘制 (Relight2)	100	99.2	91.7	81.4
差图像 (Difference)	100	100	97.5	64.3

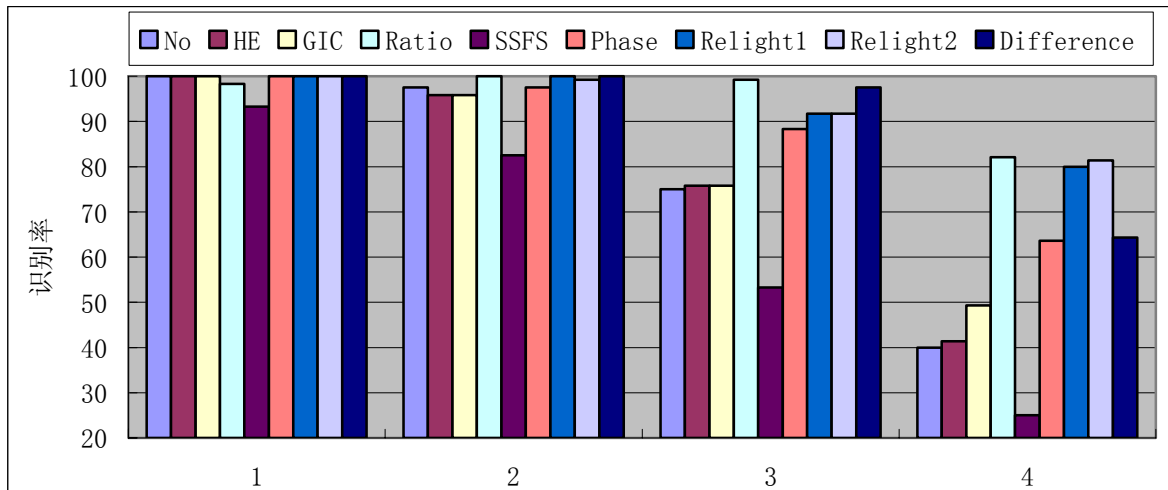


表 4-6: 各种方法在 CMU-PIE 人脸数据库上的实验结果。

Table 4-6. The recognition rates of the algorithms on the CMU-PIE face database.

算法名称	识别率 (%)	
	“illum”集合	“lights”集合
不做光照补偿 (No)	52.9	85.7
直方图均衡化 (HE)	49.9	79.4
Gamma 灰度校正 (GIC)	54.3	75.3
比率图 (Ratio)	97.5	100
对称 SFS (SSFS)	62.6	83.5
相位图 (Phase)	89.7	100
均匀光照绘制 (Relight1)	86.3	100
环境光+正面点光源绘制 (Relight2)	86.5	100
差图像 (Difference)	92.6	100

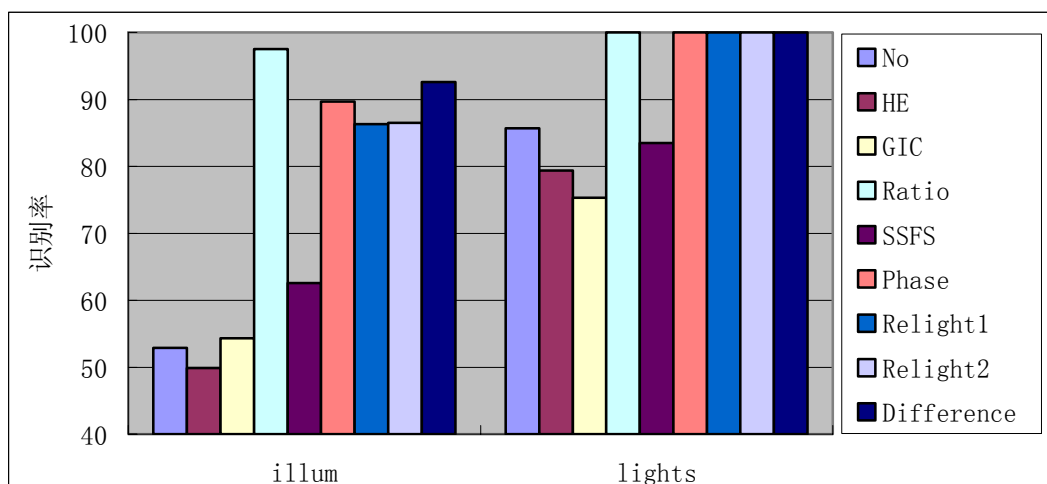
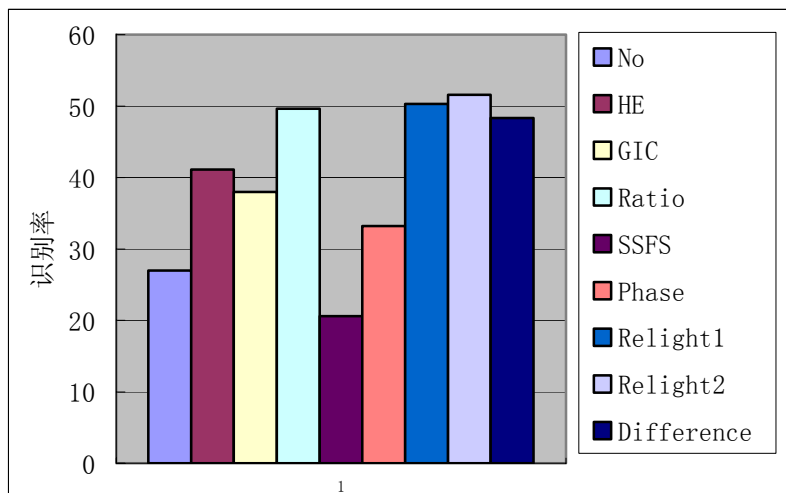


表 4-7: 各种方法在 BVC2004 人脸数据库上的实验结果。

Table 4-7. The recognition rates of the algorithms on the BVC2004 face database.

算法名称	识别率 (%)
不做光照补偿 (No)	27.0
直方图均衡化 (HE)	41.1
Gamma 灰度校正 (GIC)	38.0
比率图 (Ratio)	49.6
对称 SFS (SSFS)	20.6
相位图 (Phase)	33.2
均匀光照绘制 (Relight1)	50.3
环境光+正面点光源绘制 (Relight2)	51.6
差图像 (Difference)	48.3



4.4.3 对比实验结果分析

从上述实验结果我们发现:

(1) 直方图均衡化(HE)和 Gamma 灰度校正(GIC)在 Yale B 人脸数据库和 CMU-PIE 人脸数据库上对人脸识别率的提高不明显, 只在 BVC2004 人脸数据库上对识别率有较大提高。这说明这两种最简单最常用的方法只能在一定程度上削弱光照变化的影响。

(2) 相位图 (Phase Spectrum) 在 Yale B 人脸数据库和 CMU-PIE 人脸数据库上取得了很好的结果, 和这些数据集上与最好的结果相差不大。但其在 BVC2004 人脸数据库上对识别率的提高不大。前两个数据集中图像配准比较准确(所有图像在极短时间内采集, 图像之间的变化只有光照变化), 而 BVC2004 人脸数据库中图像中的表情变化很大, 导致图像间的配准不是很精确。由于相位图特征对配准比较敏感, 所以其在 BVC2004 人脸数据库上表现不佳。

(3) 对称 SFS 在我们实验中的结果不是很好, 其主要问题是对称性假设和点光源假设。其在 CMU-PIE 人脸数据库的“lights”集合上的结果相对没有其在“illum”集合上的结果好。在 BVC2004 数据集上的结果也证明了对称 SSFS 在非点光源下的结果不理想。当然本实验中的光照估计算法加重了点光源假设的负效应, 因为在非点光源环境下用光照的频谱的之间的关系来确定主光源的方向不好, 但在点光源下光照估计结果很准确。

(4) 比率图 (Ratio Image) 方法在所有数据集上都取得了很好的结果, 除了在 BVC2004 人脸数据库外, 该方法的结果几乎是最好的结果。在文献[Qing04]中, 我们从图像成像模型出发(不同于比率图的角度), 也验证了 log 域中图像边缘对光照变化很鲁棒。实验中我们还发现该特征与最小相似性度量 (min) 和归一化的相关系数

(Normalized Correlation) 相似性度量相结合, 比比率图方法中二阶欧式距离 (L_2) 取得了更好一些的结果。但是由于比率图方法本质是边缘图(高频分量), 所以其对配准还是会比较敏感, 其性能会随特征点的定位精确变化较大。

(5) 本文提出的基于绘制的光照补偿 (Relighting1 和 Relighting2) 与差图像光照补偿 (Difference) 在各个数据集上都取得了很好的结果, 平均下来这几种方法的性能相差不大。差图像方法在 CMU-PIE 人脸数据库的 “illum” 集合和 Yale B 人脸数据库的 Subset1~Subset3 上的结果是这三者中是最好的, 说明差图像方法在光照变化合理时会取得很好的结果。标准光照为环境光+正面点光源的绘制比均匀光照绘制稍好, 但二者相差不大。因为在粗略配准的情况下形状信息不是很准确。在有些数据集上, 比率图方法的结果相对这三个光照补偿方法稍好, 这是因为比率图对图像进行了 log 运算, 能在一定程度上抑制噪声, 从而引入的误差较少。

4.5 光照绘制在人脸检测中的应用

现在比较流行的人脸检测算法大多是基于统计学习的方法, 如 SVM 方法[Yang02b] 和 AdaBoost 方法[Viola01][Friedman00][LiSZ04b]。基于统计学习方法的问题之一是训练样本。如果训练样本不够充分, 系统的性能就不太好。在实际情况中, 我们通常能够收集到数量足够多的训练图像, 但还是很难收集到足够多的符合训练要求的图像, 如训练集中图像的光照变化不足以覆盖整个光照变化空间。因此一个比较现实的方法是通过输入人脸的一幅图像, 合成其不同光照的虚拟图像, 用以扩充训练集。本文中基于绘制的光照补偿可以用来绘制不同光照的虚拟图像。而且由于我们可以控制输入图像的光照条件相对较好, 这样合成出来的图像噪声较少。

4.5.1 目标光照的配置

平常能收集到的图像通常是比较正常的光照。为了使训练样本能有效地覆盖整个光照变化空间, 我们需要为其他光照变化合成一些额外的虚拟图像, 因此目标光照的选择是其中一个重要问题。

本文中我们选择了 3 种目标光照: 光源方向分别为头上方 (方向 1)、右方 (方向 2) 和下方 (方向 3), 如图 4-10 所示。人脸检测中对训练样本的处理通常有一步是将人脸图像左右镜像, 因此相当于又多了一个左向光源 (方向 4) 的目标光照。人们已经通过实验[Epstein95]和理论[Ramamoorthi02]证明不同光照下的人脸变化可以用一个 5D 子空间来表示。该 5D 子空间可以用 5 个线性无关的光照下的图像来构造。上述 4 种目标光照下的图像在加上正常光照下的图像正好可以作为覆盖人脸图像光照变化子空间的 5 个基图像。

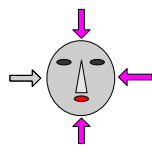


图 4-10: 人脸检测中目标光照的配置。

Figure 4-10. The configuration of the target illuminations for face detection.

在实验中，上述目标光照通过从相应光照环境下的真实人脸图像中用第三章所述的光照估计的方法得到。这些人脸图像及其对应的光照估计的结果如图 4-11 所示。

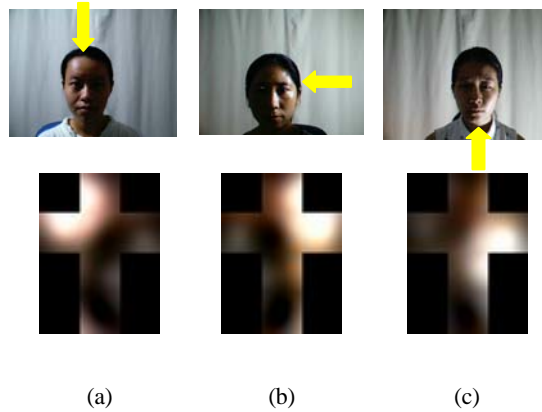


图 4-11：人脸检测中的目标光照。

Figure 4-11. The target illuminations for face detection. (a), (b) and (c) are the lighting conditions of points light source from above, from right and form bottom, respectively.

对任意随机从网上下载的 300 张正面人脸图片，裁剪出人脸区域，并对其大小进行归一化。利用第 4.1 节中的方法，可以将样本绘制到目标光照。这样就可以得到 900 张新的目标光照下的图片。一些生成样本如图 4-12 所示。

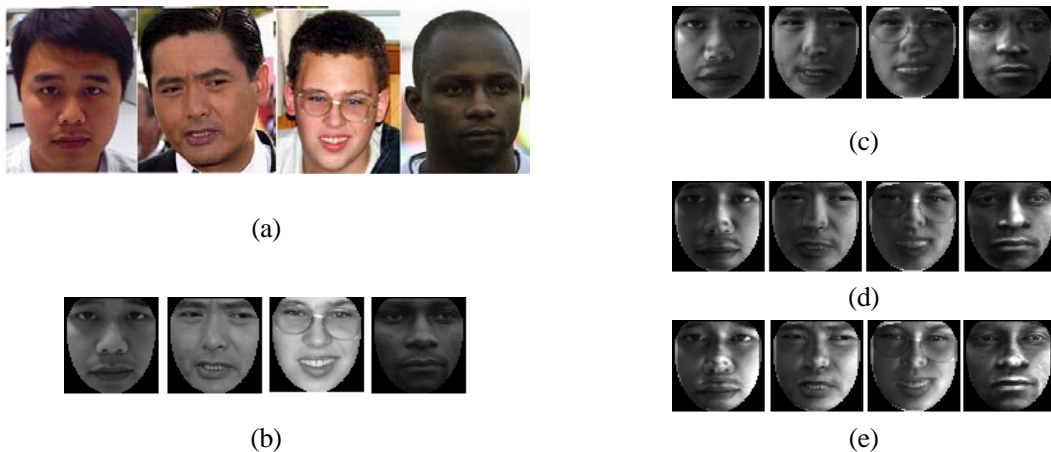


图 4-12：人脸检测中基于图像绘制生成不同光照下的虚拟样本。

(a)为原始图像。(b) 为对(a)中的图像取面部区域标准化到 64x64 大小，加 mask 后的结果。(c)、(d)、(e)分别为对(b)中图像绘制到目标光照 1、目标光照 2 和目标光照 3 下的图像。

Figure 4-12. Face relighting for face detection. (a) The original images. (b) The cropped, normalized and masked face samples; (c), (d) and (e) are the relit images under different lighting conditions.

4.5.2 人脸检测结果

本文中我们用 SVM 分类器来实现人脸检测器，其中 SVM 分类器的核函数为二阶多项式形式。

为了验证分类器性能更多与图像的光照范围而不是图像的数量有关，我们分别使用了两个训练集，两个训练集中的图像数目相同：

(1) 训练集 1: MIT 的训练库，人脸 2429 个，非人脸 4548 个。

(2) 训练集 2: MIT 的训练库加上重新生成的光照人脸样本共 2429 个，非人脸样本 4548 个，即用新生成的人脸样本替换训练集 1 中的 900 个人脸样本。

本文中测试集我们用的是 MIT 的测试集，其中共包含 24045 幅图像，其中 472 幅人脸图像和 23573 非人脸图像。所有这些图像是 19×19 的灰度图像。训练集和测试集都可以从 MIT 的 CBCL 实验室主页下载 (<http://cbcl.mit.edu/cbcl/software-datasets/FaceData2.html>)。

图 4-13 为分别用两个训练集训练 SVM 分类器在 MIT 的 CBCL 测试集上的人脸检测结果，结果以操作特性曲线 (Receiver Operating Characteristic curve, ROC curve) 的形式给出。图中 No-Light 表示训练集 1 的实验结果，Relight-900 为训练集 2 的实验结果。从图中可以看出在训练图像数目相同的情况下，用合成样本训练的分类器取得了更好的结果。

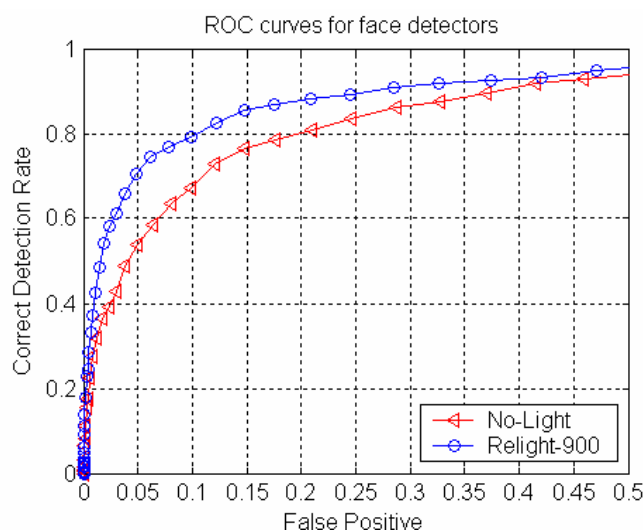


图 4-13: 不同训练集上基于 SVM 的人脸检测器的 ROC 曲线图。

Figure 4-13. The ROC curves of the face detector based on SVM trained with different training sets.

4.6 本章小结

基于球面谐波理论，在光照估计的基础上，本章讨论了人脸图像光照补偿。由于光照补偿是一个局部过程，因此光照补偿需要输入人脸每个像素点的形状和纹理信息。根

据使用中可能出现的情况，本文提出了两种对任意光照下的人脸图像进行识别的算法：基于绘制的光照补偿和差图像光照补偿。基于绘制的光照补偿通过将不同光照下的输入人脸图像都绘制到预先定义的标准光照下，从而使所有的人脸图像在相同的光照条件下进行比较。差图像为原图像与平均人脸在相同光照下的图像的差。这样差图像中既保留了特定人脸和平均人脸的差，又减弱了光照的影响。

其中基于绘制的光照补偿中我们又给出了两种不同的标准光照：均匀光照和“环境光+点光源”。不同的光照补偿方法中需要的形状信息稍有不同，结果也稍有差异。

基于绘制的光照补偿后的标准图像中不再包含输入图像的光照的信息，而差图像中还包含少量的光照信息，因此绘制的光照补偿在精确配准的情况下人脸是别的效果稍好。其中又以标准光照为“环境光+闪光灯”时，识别率稍高，因为该标准形式中不仅包含纹理信息，还包含形状信息。而准确的纹理和形状信息是人脸识别的依据。但两种标准光照下人脸识别率差别不大，这说明主要的有用识别信息为纹理信息。

在粗略配准的情况下（形状信息不够准确），基于绘制的光照补偿相比精确配准的结果都稍差，但下降得并不太多。说明在精确配准不可得的情况下，粗略配准的结果也可以接受。而差图像并不需要精确配准，且在粗略配准的情况下差图像补偿的人脸识别效果比基于绘制的光照补偿稍好一些。因此差图像更适合实用系统且计算也简单。

虽然初步的实验表明，基于不精确的特征配准进行 3D 形状估计进而对人脸识别影响不是太大，但准确的特征点定位有助于得到更准确的 3D 形状，进而提高人脸识别率。因此下一步工作是提高特征点定位的准确性，尤其是在光照条件不佳时。另外降低光照补偿中的噪声也是下一步工作之一。

第五章 特征谐波图像模型：不同光照与姿态下的人脸识别

迄今为止，我们考虑的都是正面姿态、不同光照下的人脸图像识别。由于人脸图像通常同时存在光照变化和姿态变化，本章我们将讨论二者同时变化的情况。

人脸图像之间的差别很大，引起这些差别的因素既包括不同人脸本身的内在变化，也包括不同外部成像条件（如光照、相机的视点等）引起的变化。一个好的模型应该能描述这些变化并有效地区分由外部成像条件引起的人脸图像变化和由内在人脸变化引起的人脸图像变化，从而实现光照和姿态不变的人脸识别。

如前所述（第3.3节），球面谐波基图像（简称谐波基图像）包含了人脸的形状和纹理，因此可作为与光照姿态无关的特征用于光照和姿态变化下的人脸识别。但从一幅任意光照和姿态下的图像恢复人脸的球面谐波基图像是一个病态问题，本文我们用PCA（Principal Components Analysis）模型表示人脸谐波基图像先验分布空间（特征谐波基图像模型），从而使得问题可解。

本章是如下组织的：首先在第5.1节中介绍采用本方法的动机，即为什么用特征谐波基图像模型。然后本章的第5.2节中介绍基于特征谐波基图像的人脸识别系统框架。第5.3节介绍模型的构造。接着在第5.4节和第5.5节分别介绍从正面姿态注册人脸和非正面姿态人脸图像注册人脸，即从输入图像重构输入脸的谐波基图像的PCA系数。这些系数可以看成是人脸本质特征的表示，因此第5.6节介绍利用模型参数进行光照和姿态变化下的人脸识别。最后为本章小结。

5.1 动机

特征谐波基图像模型包括两部分：谐波基图像和PCA模型。下面我们分别从两个方面来解释为什么用特征谐波基图像模型。

5.1.1 为什么用谐波基图像 (Harmonics)

每个人脸能产生很多图像。这些多变的图像使得人脸分析很困难。除不同人脸之间本身引起的人脸图像变化外，引起人脸图像的变化因素大致可分为：

(1) 姿态变化引起人脸图像巨大变化。姿态变化引起人脸的部分可见或不可见，另外两个不同视图中的可见部分之间的空间关系也不同了。

(2) 即使姿态不变，光照变化也影响人脸的外观。光源的位置和分布影响人脸图像的亮度分布、阴影和镜面高光的位置。另外投射阴影还会在人脸图像中产生一些很强的边缘。

(3) 其他如表情、年龄、饰物和化妆等引起人脸外观的变化。

显式描述并分离这些引起人脸图像变化的因素是人脸分析系统的最终目标。我们期望用以描述人脸身份的参数不会受到人脸图像的光照和姿态参数的影响。也就是说，对人脸

身份、光照和姿态等用不同的参数显式地描述。设计这样一个系统的主要挑战是要找到一些合适的参数，既能有效地对图像建模，又能简洁地描述一幅图像。

在上述变化中，光照变化和姿态变化遵循物理成像法则。这些法则反映了由3D人脸及其表面与光线交互的约束。这些约束可以用一个3D表示来限制。如前（第3.3节）所述，球面谐波基图像（简称谐波基图像）模型很适合处理任意光照变化下的人脸图像，且谐波基图像包含了人脸的本质特征：形状和反射率。因此我们可以将谐波基图像作为人脸的3D表示。在谐波基图像的基础上，光照参数用光照的9个低频分量的系数表示即可。这种表示既简单，又能表示任意类型的光照环境。

5.1.2 为什么用 PCA 模型 (Eigen)

从一幅图像中恢复人脸的形状和纹理是一个病态问题[ZhangR99]，从中恢复其谐波基图像也同样困难。借助人脸类的先验知识可以在一定程度上使得该问题可解。

人脸类的先验知识可以通过学习人脸谐波基图像的概率密度分布函数得到。我们采用一个最常用的假设，假设人脸谐波基图像的分布符合高斯模型。这样，我们可以用PCA（Principle Component Analysis）模型来表示人脸谐波基图像的先验概率分布。PCA模型很适合表示一个大类中各个小类别之间的关系。如果将所有人脸看成是一个大类，则不同人脸可看成是其中不同的小类。PCA模型[Turk91][Moghaddam00]在人脸识别中得到了广泛应用，且PCA模型在表示人脸类的先验知识（人脸类的形状[Cootes95][Atick96][Blanz03]和人脸类的纹理[Cootes98][Blanz03]）也取得了很好的实验结果。本文中我们将PCA模型扩展到用来描述人脸的谐波基图像的分布（EigenHarmonics）。谐波基图像中同时包含了人脸的纹理和形状。3D变形模型方法[Blanz03]假设人脸的形状和纹理独立，对纹理和形状分别建模。对纹理和形状参数分别建模使得图像和这些参数之间的关系很复杂，只能用迭代优化的方法来分别求解各个参数。而用谐波基图像将形状和纹理信息统一起来，将大大方便后续的参数估计。

谐波基图像中包含了人脸全部区域的信息，PCA模型描述的是这些全局信息的分布。由于PCA系数数目相对图像像素数目很少，根据姿态变化的人脸图像中的可见部分仍可比较稳定地恢复出该人脸的谐波基图像的PCA系数（由局部推导全局），从而实现光照和姿态同时变化的人脸识别。

5.2 基于球面谐波基图像的人脸识别系统框架

基于EigenHarmonics模型的人脸识别系统框架如图5-1所示。首先通过一个带有纹理的3D人脸数据库，计算库中每个人脸的9个谐波基图像，对库中所有人脸的谐波图像空间用PCA模型进行分析，学习得到人脸谐波基图像的分布模型。这部分将在第5.1节介绍。给定模型参数及成像参数，我们可以合成相应的人脸图像。另外，给定一幅输入人脸图像，我们可以对其进行分析，估计其成像参数及模型参数。我们称之为对输入人脸图像进行注册。

模型参数可以看成是人脸本质特征的表示，从而用于光照和姿态变化的人脸识别。

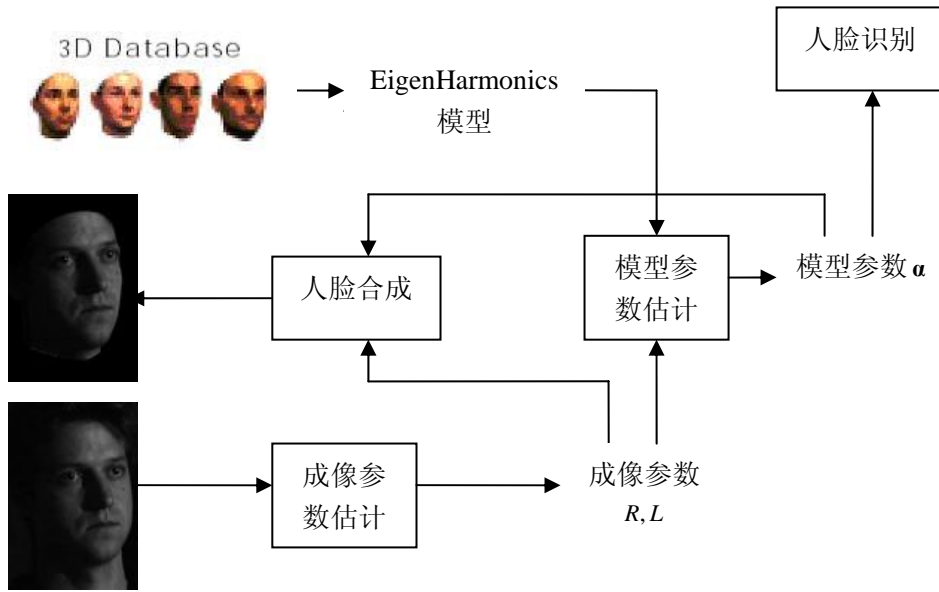


图 5-1: 基于 EigenHarmonics 模型的人脸识别系统框架。

Figure 5-1. The framework of the face recognition based on the EigenHarmonics model.

5.3 特征谐波基图像模型的构造

训练EigenHamonics模型需输入若干人脸的谐波基图像，然后再用PCA模型对其进行分析，学习得到人脸谐波基图像的概率分布。如果知道一个人脸的纹理和3D形状，我们就可以计算其9个谐波基图像。因此我们训练EigenHamonics模型的训练数据集为一个带有纹理的3D人脸数据库。实验中我们选用USF Human ID 3D database [Blanz99]。库中共有100个人脸，其中每个人脸用Cyberware™3030PS激光扫描仪得到其顶点3D几何坐标及纹理。对激光扫描得到的原始数据进行预处理后，再将每个人脸用光流法对齐到一个参考空间 [Blanz99] [Blanz03]。对齐后在参考空间中一个人脸的形状和纹理如图5-2所示。

在计算机图形学中，物体的3D形状通常表示为一些由顶点构成的三角形，每个三角形表示一个平面，可以计算一个法向量。具体的，令三角形的三个顶点分别为 $A(x_A, y_A, z_A)$ 、 $B(x_B, y_B, z_B)$ 和 $C(x_C, y_C, z_C)$ ，则该三角形的法向量为垂直该三角形平面的方向，即 $\vec{n} = AB \times AC$ ，用顶点坐标表示为

$$\begin{aligned} \vec{n} &= (x_n, y_n, z_n), \\ x_n &= (y_B - y_A) \times (z_C - z_A) - (y_C - y_A) \times (z_B - z_A), \\ y_n &= (z_B - z_A) \times (x_C - x_A) - (z_C - z_A) \times (x_B - x_A), \\ z_n &= (x_B - x_A) \times (y_C - y_A) - (x_C - x_A) \times (y_B - y_A). \end{aligned} \tag{5-1}$$

顶点的法向量则为其所在三角形法向量的均值（一个顶点通常在相邻的多个三角形中）。根据上述计算得到的一个人脸法向量图在参考空间中的例子如图5-2所示。

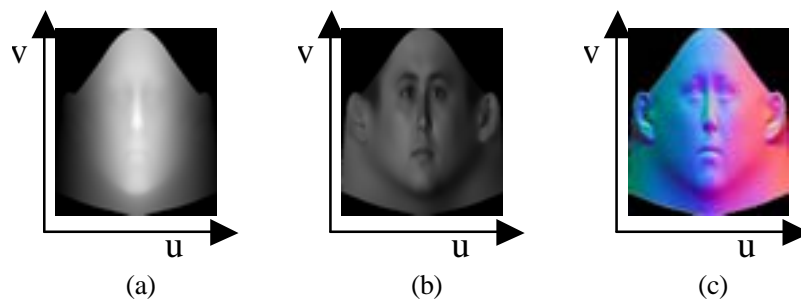


图 5-2: 平均人脸在参考空间 (u, v) 中的表示。

平均人脸的(a)形状(高度图)、(b)纹理图和(c)法向量图。其中高度图为高度线性拉伸至0~255, 法向量图的 R 、 G 、 B 分别表示 X 、 Y 、 Z 方向的法向量, 由 $[-1, 1]$ 线性拉伸至 $[0, 255]$ 。

Figure 5-2. The average face in the reference space (u, v) . (a) The shape of the average face (height map). (b) The texture map of the average face. (c) The normal map of the average face. In the height map, the intensities are linear scaled to $[0, 255]$ according to the height of the vertexes. And in the normal map, the R , G and B channel are the values of the normal in X , Y and Z axis respectively, linear scaled from $[-1, 1]$ to $[0, 255]$.

对训练库中的每个人脸, 我们首先根据其3D形状用式(5-1)计算其法向量, 然后计算其谐波基图像。令 λ_n 为人脸第 n 个顶点的反射率(纹理), (x_n, y_n, z_n) 为该点法向量的笛卡尔坐标表示, 则人脸该顶点的谐波基图像由式(5-2)给出:

$$\begin{aligned}
 B_{00}(n) &= \lambda_n \sqrt{\frac{\pi}{4}}, \\
 (B_{1-1}, B_{10}, B_{11})(n) &= \lambda_n \sqrt{\frac{\pi}{3}} (y_n, z_n, x_n), \\
 (B_{2-2}, B_{2-1}, B_{21})(n) &= \lambda_n \sqrt{\frac{15\pi}{64}} (x_n y_n, y_n z_n, x_n z_n), \\
 B_{20}(n) &= \lambda_n \sqrt{\frac{5\pi}{256}} (3z_n^2 - 1), \\
 B_{22}(n) &= \lambda_n \sqrt{\frac{15\pi}{256}} (x_n^2 - y_n^2).
 \end{aligned} \tag{5-2}$$

设每个人脸有 N_v 个顶点(USF Human ID 3D 人脸数据库中每个人脸用75972个顶点表示, 为计算方便, 本文中我们再对其进行均匀下采样为8955个顶点), 将每个人脸的9个谐波基图像串联成一个 $N_v \times 9$ 的矩阵 \mathbf{H} , 记为

$$\mathbf{H} = \begin{pmatrix} B_{00}(1) & B_{1-1}(1) & \cdots & B_{22}(1) \\ B_{00}(2) & B_{1-1}(2) & \cdots & B_{22}(2) \\ \cdots & \cdots & \cdots & \cdots \\ B_{00}(N_v) & B_{1-1}(N_v) & \cdots & B_{22}(N_v) \end{pmatrix}. \tag{5-3}$$

得到每个人脸的谐波基图像后,我们假设人脸谐波基图像的分布符合高斯模型。这样,我们可以对所有人脸的谐波基图像进行PCA分析,得到人脸谐波基图像的特征向量 $\mathbf{P}_1, \mathbf{P}_2, \dots$ 和特征值 $\sigma_1, \sigma_2, \dots$ 。最后取对应特征值最大的 N 个特征向量,构成人脸谐波基图像分布空间的PCA基。具体的,设训练库中共有 M 个人脸样本,对所有训练样本,减去样本均值后再组成一个数据矩阵 \mathbf{A} ,最后用SVD (Singular Value Decomposition) 方法[傅04]求 \mathbf{A} 的协方差矩阵 \mathbf{C} 的特征向量,并按特征值的大小排序 $\sigma_1^2 \geq \sigma_2^2, \dots$, 得到 $\mathbf{P}_1, \mathbf{P}_2, \dots$ 。写成公式形式为:

$$\begin{aligned}\bar{\mathbf{H}} &= \frac{1}{M} \sum_{i=1}^M \mathbf{H}_i, \\ \mathbf{a}_j &= \text{ves}(\mathbf{H}_j - \bar{\mathbf{H}}), \quad \mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M), \\ \mathbf{C} &= \frac{1}{M} \mathbf{A} \mathbf{A}^T = \frac{1}{M} \sum_{j=1}^M \mathbf{a}_j \mathbf{a}_j^T = \frac{1}{M} \mathbf{U} \mathbf{\Lambda}^2 \mathbf{U}^T,\end{aligned}\tag{5-4}$$

其中 $\text{ves}(\mathbf{H})$ 表示将矩阵 \mathbf{H} 按行排列成向量形式,正交矩阵 \mathbf{U} 的 M 列为协方差矩阵 \mathbf{C} 的特征向量, $\sigma_i^2 = \frac{\lambda_i^2}{M}$ 为 \mathbf{C} 的特征值,其中 λ_i 为对角矩阵 $\mathbf{\Lambda}$ 的元素(降序排列)。令 $\mathbf{U}_{:,i}$ 表示矩阵 \mathbf{U} 的第 i 列,即第 i 个主成分,然后再将其重新写成 $N_v \times 9$ 的矩阵 $\mathbf{P}_i = \mathbf{U}_{:,i}^{(9)}$,其中 $\mathbf{a}_{m \times 1}^{(n)}$ 表示将 $m \times 1$ 的向量 \mathbf{a} 折成 $(m/n) \times n$ 的矩阵。

最后我们选取其中 $N(N \leq M)$ 较大的特征值对应的向量作为正交基,任意人脸可以表示成这 N 个主成分的线性组合,即

$$\mathbf{H} = \bar{\mathbf{H}} + \sum_{i=1}^N \alpha_i \mathbf{P}_i.\tag{5-5}$$

若某人脸的谐波基图像为 \mathbf{H} ,将其投影到PCA空间得到系数为 $\boldsymbol{\alpha}$,则其概率为

$$p(\mathbf{H}) \sim \exp\left(-\frac{1}{2} \sum_i \frac{\alpha_i^2}{\sigma_i^2}\right).\tag{5-6}$$

也就是说,如果 $\frac{\alpha_i}{\sigma_i}$ 较大,则这种情况很少见,通常 $\frac{\alpha_i}{\sigma_i} \leq 3$ 。

实验中我们取包含99.97%能量的特征向量, $N = 89$ 。图5-3给出了平均谐波基图像、第一个和第二个特征谐波基图像、及最后一个特征谐波基图像。前面的特征谐波基图像很像人脸的形状,后面的特征谐波基图像的值更随机一些。

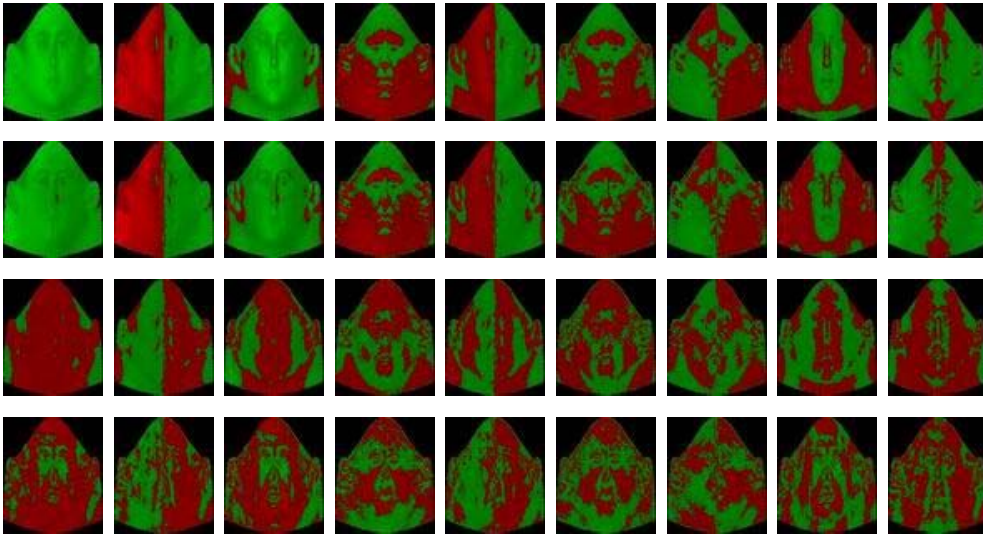


图 5-3: 特征谐波基图像。

第一行为平均谐波基图像，第二行和第三行分别是第一个特征谐波基图像及第二个特征谐波基图像，第四行为最后一个特征谐波基图像。绿色代表正值，红色代表负值（值的范围均为 140~240）。

Figure 5-3. The visualized of the eigen-harmonic images. The first row is the average vector. The second row and the third row are the first two eigenvectors respectively. And the last row is the last eigenvector. Green indicates positive values and Red indicates negative values (the range of intensities is from 140 to 240).

5.4 根据正面人脸图像注册人脸

第5.3节中EigenHarmonics模型是在参考空间 (u, v) 中构造的。为了应用该模型对输入人脸图像进行分析，我们首先需要将输入人脸图像与模型对齐。这可以通过将输入人脸图像变换到参考空间中实现。

将输入图像与模型对齐后，谐波基图像重构通过两步实现：成像参数估计和EigenHarmonics模型参数估计。对于正面人脸图像来说，成像参数估计为光照 \mathbf{L} 的估计。

在将输入图像与模型配准及估计成像参数时，我们用平均3D人脸模型代替特定的输入人脸的模型。在将输入图像与模型配准时，我们用的是平均的3D人脸形状模型加人脸形状的统计模型；在光照估计时，我们用的是平均的人脸谐波基图像。这是基于这样一个假设：在估计成像参数（如光照、姿态参数）时，用一个平均人脸模型就已经足够了，因为成像参数对图像的影响是整体的，用全局信息即可估计出成像参数。且成像参数的维数相对图像的像数数目很少，输入人脸与平均人脸之间的微小差异不会对成像参数估计造成严重影响。

5.4.1 人脸图像与模型配准

通过纹理映射可以将输入图像变换到参考空间，而纹理映射又需要知道输入人脸的3D形状。因此我们下面先讨论输入人脸形状的估计。

给定输入人脸图像, 我们首先在输入图像中定位一些关键特征点。人脸图像上的关键特征点可以用第三章提到的局部纹理约束的主动表现模型 (LTC-AAM) [山04]来定位。这些特征点对应平均3D人脸形状上唯一的顶点, 也对应着EigenHarmonics模型上唯一的顶点标号。

为了得到输入人脸更精确的3D形状, 我们对人脸的3D形状分布用一个统计模型限制, 再结合2D图像中的信息, 我们就可得到输入人脸的3D形状。本文中人脸的3D形状分布也是用一个PCA模型表示[姜05]。

3D人脸的几何数据用向量 $S=(X_1, Y_1, Z_1, X_2, \dots, Y_n, Z_n)^T \in \mathfrak{R}^{3n}$ 来表示, 其中 X, Y, Z 分别表示点的三维坐标。在PCA模型中, 设 \bar{S} 为平均人脸, $P \in \mathfrak{R}^{3n \times m}$ 是前 m 维特征向量的排列矩阵(特征向量按特征值大小降序排列), 那么一个新的人脸几何形状 S 就可以表示成:

$$S = \bar{S} + P\bar{\alpha}, \quad (5-7)$$

其中 $\bar{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)^T \in \mathfrak{R}^m$ 是特征向量的参数。

假设一共有 t 个2D人脸特征点被选中做3D重建, 那么在3D人脸几何模型上也有对应的 t 个特征点。设 $S_f=(X_1, Y_1, X_2, \dots, X_t, Y_t)^T \in \mathfrak{R}^{2t}$ 是这些特征点的平面坐标(只包含 X, Y 坐标)向量, 那么 S_f 就是 S 的子向量。根据公式(5-7), 一个新模型 S_f 特征点的 X, Y 坐标可以表示为:

$$S_f = \bar{S}_f + P_f \bar{\alpha}. \quad (5-8)$$

其中 $\bar{S}_f \in \mathfrak{R}^{2t}$ 和 $P_f \in \mathfrak{R}^{2t \times m}$ 是特征点的 X, Y 坐标分别在 \bar{S} 和 P 上相对应的部分。为了把人脸模型坐标系转化成2D图像的坐标系, 设 S'_f 是坐标转化后人脸几何形状, S'_f 与 S_f 的关系可以表示为:

$$S'_f = cS_f + T. \quad (5-9)$$

其中 $T \in \mathfrak{R}^{2t}$ 是平移向量, $c \in \mathfrak{R}$ 是缩放系数。考虑到2D人脸图像与3D人脸模型都是正面的, 公式中不需要旋转矩阵。因为 S'_f 是由2D图像上的特征点坐标所确定的, 在特征点定位之后 S'_f 就成为一个已知的常量。

设 T_x 和 T_y 是 S_f 上的特征点在 X, Y 坐标轴上的平均位移, 因为PCA训练过程中认为 S_f 点的坐标平均值为0, 那么由(5-9)可以得到:

$$(T_x, T_y)^T = \frac{1}{t} \sum_{i=1}^t (S'_{fi} - cS_{fi}) = \frac{1}{t} \sum_{i=1}^t S'_{fi}. \quad (5-10)$$

因为 S'_f 是常量, 所以 T 也是一个常量。

由公式 (5-9) 又可以变形推导出:

$$c = \frac{\sum_{i=1}^f \langle S_{f_i}' - (T_x, T_y)^T, S_{f_i} \rangle}{\sum_{i=1}^f \|S_{f_i}\|^2}, \quad (5-11)$$

$$S_f = \frac{1}{c}(S_f' - T), \quad (5-12)$$

因为 P_f 是正交矩阵, 根据公式 (5-8), $\vec{\alpha}$ 可以表示成:

$$\vec{\alpha} = (P_f^T P_f)^{-1} P_f^T (S_f - \overline{S_f}). \quad (5-13)$$

为了避免奇异的结果, 需要加入一些先验成分来限制 $\vec{\alpha}$, 那么公式 (2-7) 就变为:

$$\vec{\alpha} = (P_f^T P_f + \lambda \Lambda^{-1}) P_f^T (S_f - \overline{S_f}). \quad (5-14)$$

其中 $\Lambda = \text{diag}(v_1, v_2, \dots, v_m)$, λ 是权重系数, v_i 表示第 i 个特征值。

为了得到人脸几何参数 $\vec{\alpha}$, 本文用循环迭代的方法来计算。在循环之前, 根据提取的 2D 特征点 S_f' 计算出常量 T , 设置 S_f 的初值为 $\overline{S_f}$ 。

第一步: 把常量 S_f' 和 T , 以及初值 S_f 代入公式 (5-11) 求出缩放系数 c ;

第二步: 把 c 代入公式 (5-12) 可以得到新的 S_f ;

第三步: 把 S_f 代入公式 (5-14) 可以计算出特征向量参数 $\vec{\alpha}$;

第四步: 把 $\vec{\alpha}$ 代入公式 (5-8) 就可以得到新的 S_f 。

因为每一步都是求最优解, 最后会达到局部收敛。循环第一到四步 10 次左右, 特征向量参数 $\vec{\alpha}$ 就会逼近一个固定值, 然后把 $\vec{\alpha}$ 代入公式 (5-7) 就可以得到三维人脸几何模型 S 。

此时 3D 人脸几何模型上的特征点坐标与 2D 图像上特征点坐标还有一定的偏差, 因为 3D 几何模型是用特征向量参数恢复的, 而 3D 人脸库所能覆盖的人脸空间比较有限, 所以 3D 几何模型上的特征点无法做到十分准确。为了确保 3D 几何模型上的特征点坐标与 2D 图像上的特征点坐标完全相同, 我们强行使三维模型上的特征点与二维图像上的特征点对齐, 然后再用 **Kriging** 插值方法对非特征点进行插值。在经过插值计算之后, 最后的 3D 几何模型就重构完毕。

得到输入人脸的 3D 形状之后, 假设图像形成为正投影, 我们通过纹理映射 (**Texture Mapping**) 技术将该人脸 3D 形状的每个顶点与人脸图像中的像素对应起来。由于 3D 形状是在参考空间中, 这样输入图像也变换到了参考空间。一个将输入人脸图像及其特征点和变换到参考空间中形式如图 5-4 所示。

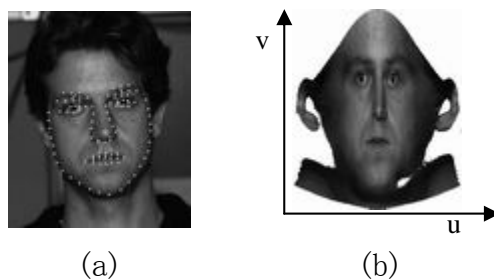


图 5-4: 输入人脸及其与模型对齐后的图像。

(a) 输入人脸。(b) 与模型对齐后的图像，其中白色点表示不可见点或模型中无效的点。

Figure 5-4. The input face image (a) and its aligned image in reference space (b) (the white pixels are invisible or invalid in the model).

5.4.2 光照估计

成像模型在第三章中详细介绍过，为了使本章自成体系，在这我们再给出：

$$\mathbf{I} = \sum_{l=0}^2 \sum_{m=-l}^l L_{lm} \mathbf{B}_{lm} = \mathbf{H}\mathbf{L}, \quad (5-15)$$

其中 $\mathbf{L} = (L_{00}, L_{1-1}, L_{10}, L_{11}, L_{2-2}, L_{2-1}, L_{20}, L_{21}, L_{22})^T$ ，为光照的9个低频分量的系数， \mathbf{H} 为人脸的谐波基图像，如式 (5-3) 所示， \mathbf{I} 为输入图像的列向量表式。

将输入人脸图像与模型对齐后，我们就可以估计其光照条件。在光照估计时，由于输入人脸的谐波基图像是未知的，我们用平均人脸谐波图像 $\overline{\mathbf{H}}$ 替代输入人脸的谐波图像 \mathbf{H} （第三章中我们假设人脸表面的反射系数为常数进行光照估计）。光照估计的目标是使得绘制到该光照下的图像与输入图像最相似，即二者之间的均方差（MSE）最小。这样，通过解最小二乘方程

$$\hat{\mathbf{L}} = \arg \min_{\mathbf{L}} (\overline{\mathbf{H}}\mathbf{L} - \mathbf{I}), \quad (5-16)$$

可求得输入图像的光照系数 $\hat{\mathbf{L}}$ 。需要指出的是光照 $\hat{\mathbf{L}}$ 的坐标系是相对参考空间。

由于光照参数的维数很少，光照估计时可以将图像 \mathbf{I} 和平均人脸的谐波基图像 $\overline{\mathbf{H}}$ 的分辨率都降低。分辨率降低之后，输入人脸的谐波基图像与平均人脸的谐波基图像之间的差别更小。这样用平均模型来代替特定人脸模型的风险也就越小。第三章中我们也已经用实验证明光照估计对图像分辨率不敏感，该结论对本章仍然适用。本章实验中我们取图像的分辨率为 22×22 。在光照估计和后续的谐波基图像重构过程中我们只用到了在人脸面部区域可见部分的像素。

光照估计的几个例子如图5-5所示。图中第1、2列和第3、4列图像的真实光照环境分别相同，第二行中对应的光照估计结果也十分相似。这在一定程度上说明我们的光照估计策略是有效的。图中光照估计的结果用辐照图（Irradiance Environment Map）的形似给出，同第3.3.1节辐照图的计算为

$$E(\theta, \phi) = \sum_{l=0}^2 \sum_{m=-l}^l A_l L_{lm} Y_{lm}(\theta, \phi), \quad (5-17)$$

其中 $A_0 = \pi, A_1 = \frac{2\pi}{3}, A_2 = \frac{\pi}{4}$ 。

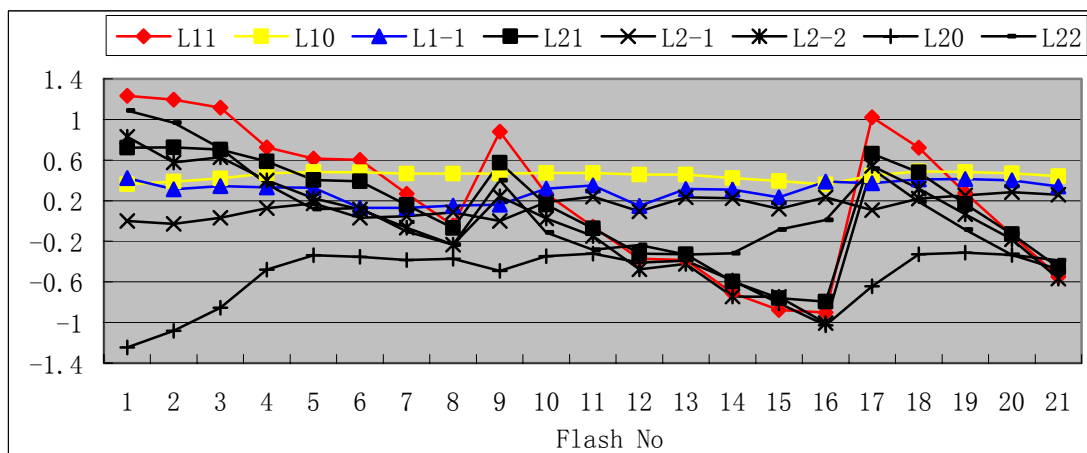


图 5-5: CMU-PIE 人脸数据库上光照估计的结果。

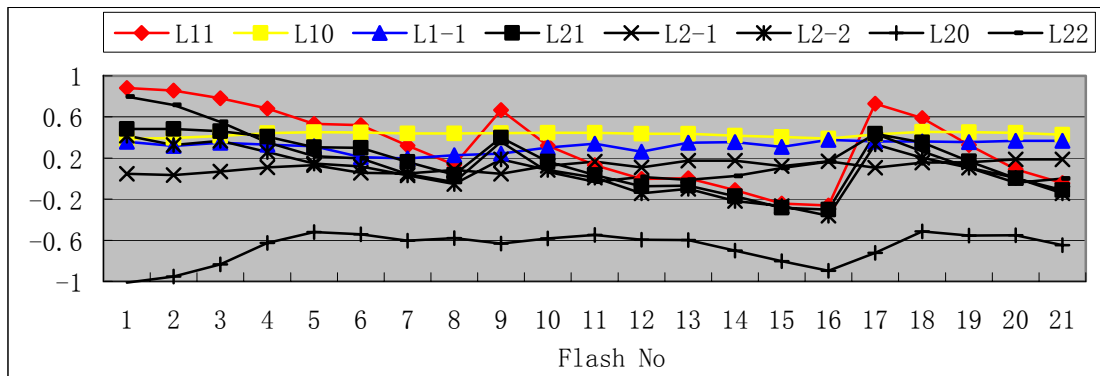
前两列和后两列的光照分别相同而人脸不同，第二行为第一行图像对应的光照估计结果。虽然两个人脸差别较大，但相同光照下的光照估计结果非常相似。光照估计的结果以辐照度图的形式给出。

Figure 5-5. Some examples of the estimated irradiance maps from the CMU-PIE face database. The first two columns and the last columns are the same illuminations with different person respectively. The second row is the corresponding estimated irradiance maps. Evidently, the estimated irradiance maps from the images of same illumination with different persons are fairly similar.

更多的光照估计结果如图 5-6 所示。图中分别给出了 CMU-PIE 人脸数据库中“illum”集合（无背景光）中和“lights”集合（有背景光）中各种光照下光照系数的均值。同第 3.5.3 节，后 8 个系数都除以直流系 L_{00} 数作归一。本章的结论同第三章类似，如在“illum”集合中，闪光灯编号为“f02”与编号为“f17”时，系数 L_{11} 数值基本相等而符号相反，表示这两种光照的主光源方向相对 YOZ 平面对称。另外有背景光时的对应光照的各光照系数比无背景光的情况要小，说明背景光对直流分量贡献较大。



(a) “illum” 集合

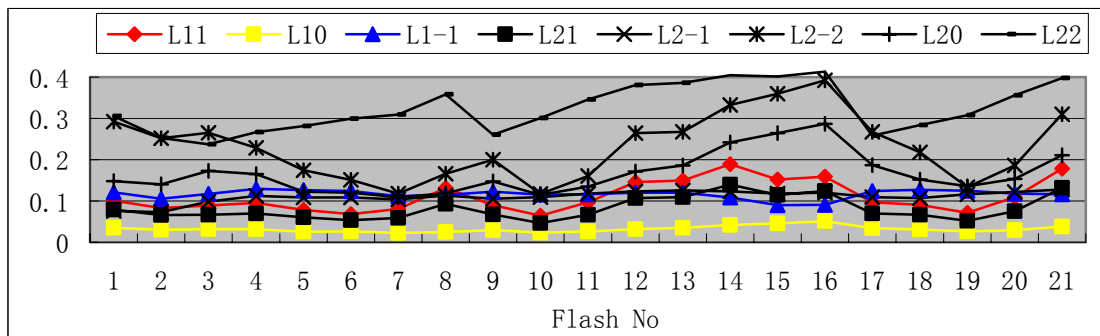


(b) “lights” 集合

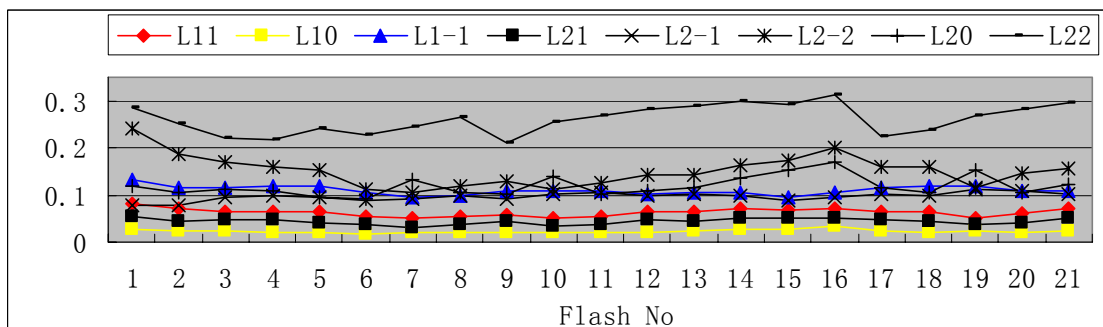
图 5-6: CMU-PIE 人脸数据库中各种光照下光照系数的均值。

Figure 5-6. The mean of the coefficients of the illuminations in the “illum” set (a) and the “lights” set (b) of CMU-PIE face database respectively.

为了考察光照估计结果的稳定性，我们统计了不同人脸在各种光照下的光照系数的方差。CMU-PIE 人脸数据库中各种光照下不同人脸的光照估计的系数的方差如图 5-7 所示。从图中可以看出，各种光照下各系数的方差不大，尤其是一阶的 3 个分量。如在各种光照下 L_{11} 的均值变化范围很大（如图 5-6(a)所示，从 1.4~-1.4），但其方差一直保持在 0.1 左右，也就是说各种光照下光照估计的结果随人脸变化很小，主要随光照变化而变化。8 个系数中有 3 个系数的方差较大，因为图像只是人脸正面可见的部分，无法提供关于人脸背后的光照环境的足够多的信息，所以其结果受临时噪声的影响较大。



(a) “illum” 集合



(b) lights” 集合

图 5-7: CMU-PIE 人脸数据库中各种光照下光照系数的方差。

Figure 5-7. The standard deviations of the coefficients of the illuminations in the “illum” set (a) and the “lights” set (b) of the CMU-PIE face database respectively.

5.4.3 谐波基图像重构

估计出光照后，我们将光照估计结果代入式 (5-15)，以求解输入人脸的谐波基图像。由于谐波基图像为 \mathbf{H} 为 $N_v \times 9$ 的矩阵，已知条件 \mathbf{I} 为 $1 \times N_v$ ，所以即使光照已知，已知条件还是比未知数少，求解输入人脸的球面谐波基图像还是一个病态问题。要使得问题可解，必须加入先验知识辅助。

本文中我们采用的先验知识为第 5.2 节中构造的 EigenHarmonics 模型。在 EigenHarmonics 模型中， $\mathbf{H} = \bar{\mathbf{H}} + \sum_{i=1}^N \alpha_i \mathbf{P}_i$ ，不同人脸的谐波基图像表现为 PCA 系数 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)$ 不同。因此重构输入人脸的谐波基图像就是估计输入人脸的谐波基图像的 PCA 系数 $\boldsymbol{\alpha}$ 。

综合成像公式 (5-15) 和 EigenHarmonics 模型式 (5-5)，可以得到

$$\mathbf{I} = \mathbf{H}\mathbf{L} = (\bar{\mathbf{H}} + \sum_{i=1}^N \alpha_i \mathbf{P}_i)\mathbf{L} = \bar{\mathbf{H}}\mathbf{L} + \sum_{i=1}^N \alpha_i (\mathbf{P}_i\mathbf{L})。 \quad (5-18)$$

将上式中未知数移到等式的一边，得到 $\mathbf{I} - \bar{\mathbf{H}}\mathbf{L} = \sum_{i=1}^N \alpha_i (\mathbf{P}_i\mathbf{L})$ 。这样通过解一个最小二乘方程

$$\boldsymbol{\alpha} = \arg \min_{\boldsymbol{\alpha}} \left\| \mathbf{I} - \bar{\mathbf{H}}\mathbf{L} - \sum_{i=1}^n \alpha_i (\mathbf{P}_i\mathbf{L}) \right\|, \quad (5-19)$$

我们得到给定人脸谐波图像的 PCA 系数 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)$ 。最后该人脸的谐波图像可以通过 (5-5) 式重构。

根据 CMU-PIE 人脸数据库中不同光照下两个不同人脸的图像重构出的谐波基图像的如图 5-8 所示。从图中可以看出，根据相同人脸不同光照下的图像重构出的谐波基图像几乎相同，而不同人脸的谐波基图像仍然保留了输入人脸原有的差别。这说明不同的光照并不对谐波基图像重构造成很大的影响。

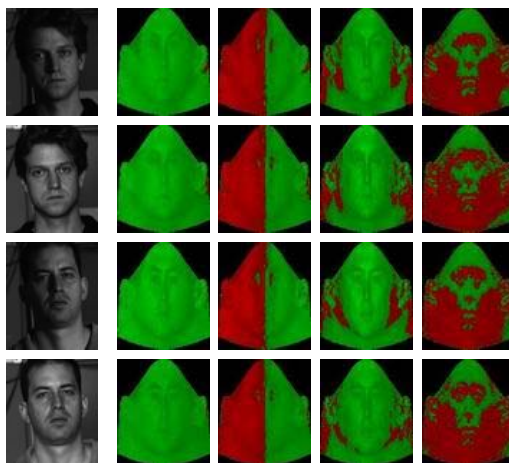


图 5-8: 根据不同输入人脸在不同光照下的图像重构的球面谐波基图像。

其中第一列为输入图像, 接着的四列位重构出的球面谐波基图像 (限于篇幅只给出了前四个)。

Figure 5-8. The results of reconstructed spherical harmonic images recovered from images of the different persons under various lighting. The first column is the images used for the recovery, followed by the set of the recovered spherical harmonic images.

5.5 从任意姿态人脸图像重构谐波基图像

迄今为止, 我们讨论的都是只有光照变化的情况。实际上经常是多种因素一起变化, 尤其是姿态和光照同时变化。在简单回顾前人在姿态和光照同时变化方面的工作后, 我们将介绍EigenHarmonics模型在解决光照和姿态同时变化的过程及结果。

同时解决人脸识别中的光照和姿态变化问题的工作相对较少。FisherLightField [Gross02b]是一个较好地解决了姿态和光照问题的方案。EigenLightFiled[Gross02a]从输入多个姿态的图像中获得该人脸的光场, 然后将训练集中的多个人脸的光场用PCA模型分析, 得到人脸光场的先验分布。对输入任意姿态的图像, 可以看成是整个光场中可见的一部分。用这可见部分可估计其光场的PCA系数, 从而解决姿态变化问题。将Eigen方法扩展到Fisher, 类似将Eigenfaces[Turk91]方法扩展到Fisherfaces[Belhumeur97]方法, 就可以解决光照问题。Zhou[Zhou04]将EigenLightFiled方法扩展, 用一个朗伯反射模型来处理光照变化。

上述方法是直接在2D图像进行分析。由于光照变化和姿态变化都与人脸的3D结构有关, 因此如果3D结构已知, 则很容易实现光照和姿态不变的人脸识别。一些基于模型的方法从2D图像/图像序列重构输入人脸的3D结构, 对引起图像变化的内在的人脸3D结构和外在的成像参数显示地分离。第二章中介绍过的光照锥[Georghiades01]和对称SFS[Zhao00c]能同时解决光照变化和姿态变化问题。3D变形模型[Blanz03]是这方面另一个很好方法。其报告的识别率很高, 且在FRVT2002测试中作为姿态校正, 已经取得了很好的效果 [Phillips03]。

与 EigenLightFiled[Gross02a] 和 3D 变形模型 [Blanz03] 方法有异曲同工之处, EigenHarmonics也可以同时解决光照变化和姿态变化。因为球面谐波基图像 (Harmonics) 中包含了人脸全部区域的信息, 各种姿态下的人脸可以看作人脸区域可见的某部分。由于

EigenHarmonics模型中PCA系数的维数相比人脸图像的像素数还是很少，利用这些可见部分的人脸图像信息，仍可以恢复出输入人脸的PCA系数，从而得到整个人脸区域的信息。另外如本文第5.4节所述，球面谐波基图像很方便处理光照变化问题。

利用EigenHarmonics模型处理姿态和光照同时变化的过程与只有光照变化的情况类似。只是在光照估计之前首先要做姿态估计，将变化姿态下的人脸图像与模型对齐后，其余的过程相同。

5.5.1 姿态估计

对于非正面姿态人脸图像（正面姿态可以看成是任意姿态的一个特例），我们首先估计其姿态参数。

由于人脸可以近似看成三维坐标系下的一个3D刚体，因此其姿态变化也满足刚体运动变化的规律。不同姿态的人脸图像可以认为是3D人脸在自己的空间进行旋转变换之后在3D维图像空间的投影。因此，2D投影图像和3D人脸的旋转变换存在着对应关系。

我们根据2D图像上的几个特征点来重构3D人脸的旋转参数：绕X轴旋转的俯仰角度（pitch），绕Y轴旋转的左右变化的角度（yaw）以及平面内绕Z轴旋转的角度（tilt）。本质上任何一个旋转变换都可以分解为按3个坐标轴分别进行的旋转分别作用。假设绕X、Y、Z轴旋转的角度分别为 α ， β ， γ ，则三个相应的旋转变换矩阵分别是：

$$\begin{aligned} R_x &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix}, \\ R_y &= \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix}, \\ R_z &= \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (5-20)$$

其统一的变换矩阵为：

$$R_x \cdot R_y \cdot R_z = \begin{bmatrix} \cos \beta \cos \gamma & \cos \beta \sin \gamma & -\sin \beta \\ \sin \alpha \sin \beta \cos \gamma - \cos \alpha \sin \gamma & \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \alpha \cos \beta \\ \cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma & \cos \alpha \sin \beta \sin \gamma - \sin \alpha \cos \gamma & \cos \alpha \cos \beta \end{bmatrix} \quad (5-21)$$

$$\text{简记为 } R = \begin{bmatrix} r_0 & r_1 & r_2 \\ r_3 & r_4 & r_5 \\ r_6 & r_7 & r_8 \end{bmatrix}.$$

因此，3D人脸上的任何一个点 (x, y, z) 经过一个旋转变换后的新的坐标为 (x', y', z') ，

满足关系: $(x', y', z') = (x, y, z) \times R$ 。假设3D人脸的成像满足正投影, 则有:

$$(x', y') = (x, y, z) \times s \times \begin{bmatrix} r0 & r1 \\ r3 & r4 \\ r6 & r7 \end{bmatrix} + (tx, ty), \quad (5-22)$$

其中, s 表示尺度 (scale) 变化, tx, ty 分别表示沿 X 和 Y 方向的平移量。

将这个式子重写为:

$$\begin{aligned} (x', y') &= (x, y, z) \times \begin{bmatrix} s \cdot r0 & s \cdot r1 \\ s \cdot r3 & s \cdot r4 \\ s \cdot r6 & s \cdot r7 \end{bmatrix} + (tx, ty), \\ &= (x, y, z, 1) \times \begin{bmatrix} s \cdot r0 & s \cdot r1 \\ s \cdot r3 & s \cdot r4 \\ s \cdot r6 & s \cdot r7 \\ tx & ty \end{bmatrix} = (x, y, z, 1) \times \begin{bmatrix} b0 & b1 \\ b2 & b3 \\ b4 & b5 \\ b6 & b7 \end{bmatrix}. \end{aligned} \quad (5-23)$$

实验中, 我们选择5个特征点, 分别为左、右瞳孔, 鼻尖和左、右嘴角。则2D图像平面上的这些点的坐标和3D人脸上相应点的坐标同样满足上述的关系。因此有:

$$\begin{bmatrix} x0 & y0 & z0 & 1 \\ \dots & \dots & \dots & \dots \\ x4 & y4 & z4 & 1 \end{bmatrix} \bullet \begin{bmatrix} b0 & b1 \\ b2 & b3 \\ b4 & b5 \\ b6 & b7 \end{bmatrix} = \begin{bmatrix} x0' & y0' \\ x1' & y1' \\ x2' & y2' \\ x3' & y3' \\ x4' & y4' \end{bmatrix} \quad (5-24)$$

对其进行求解, 可得到

$$\begin{cases} \alpha = a \tan \frac{b4 \cdot \sin \gamma - b5 \cdot \cos \gamma}{b3 \cdot \cos \gamma - b2 \cdot \sin \gamma} \\ \beta = a \sin \frac{b4 - \sin \gamma \cdot (b4 \cdot \sin \gamma - b5 \cdot \cos \gamma)}{\cos \gamma \cdot (b3 \cdot \cos \gamma - b2 \cdot \sin \gamma)} \\ \gamma = a \tan \frac{b1}{b0} \end{cases} \quad (5-25)$$

这样就得到了输入图像的3个旋转角度参数 α , β , γ 。

由于我们无法得到输入人脸图像对应的特定的3D形状, 我们用平均3D人脸模型的特征点的坐标来代替特定人的3D模型特征点的坐标。这与光照估计基于同样的假设: 姿态参数对图像的影响是整体的, 且参数较少, 因此可用一个平均人脸模型代替输入人脸模型。当然上面的计算是在理想状态, 实际情况中输入人脸形状和平均人脸形状有微小差别, 图像形成过程是透视投影不是正投影, 这些都会带来姿态估计结果的微小的偏差。但是一般的人脸识别算法对姿态的微小偏差都有一定的容忍度, 因此上述方法在实际应用中还是取得了很好的效果。

5.5.2 谐波基图像重构

在估计出输入人脸图像的姿态后，我们将平均人脸形状旋转到该姿态。同正面图像一样，为了得到尽可能与输入人脸接近的3D形状，我们在图像上标注了一些特征点，然后结合这些特征点和人脸3D形状的统计模型，得到输入人脸的3D形状。与正面图像稍有不同的是，式(5-7)两边都乘以旋转式(5-21)中的旋转矩阵 R ，得到

$$SR = \bar{S}R + PR\vec{\alpha}, \quad (5-26)$$

从而后续的公式中都有旋转参数[Chai05]。

最后用纹理映射将输入图像映射到参考空间。两个不同姿态的输入人脸图像及其特征点和其在参考空间中的表示如图5-9所示。与正面姿态的图像相比，只是不可见部分更多了一些。

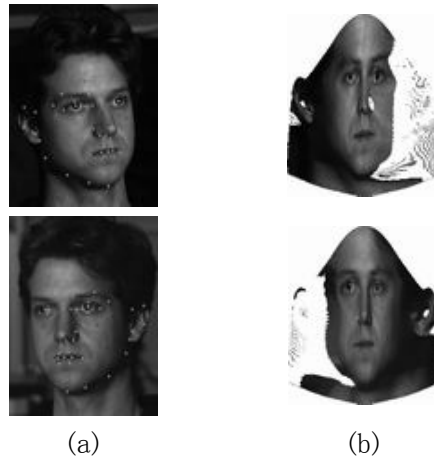


图 5-9: 不同姿态下输入人脸图像及其与模型对齐后的图像。

(a) 输入人脸图像。(b) 与模型对齐后的图像，其中白色点表示不可见点或模型中无效的点。

Figure 5-9. The input face images under different poses (a) and its aligned image in reference space (b) (the white pixels is invisible or invalid in the model).

将输入图像映射到参考空间中后，我们用第5.4.2节中的方法对输入图像进行光照估计。由于非正面姿态的图像中部分人脸区域被遮挡，在光照估计时这些被遮挡的点不参与运算。由于光照的参数只有9个，可见部分的人脸图像仍能鲁棒的估计出光照。两个不同姿态相同光照下的图像的光照结果如图5-10所示。两个光照估计的结果几乎相同，这验证了光照估计不受图像姿态变化的影响。

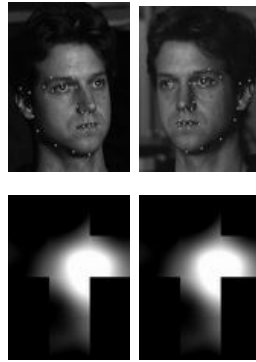


图 5-10: 同一人脸在相同光照、不同姿态下图像的光照估计的结果。

前一行输入图像，后一行是其对应的光照估计的结果，虽然姿态不同，但光照估计的结果非常相似。

Figure 5-10. The estimated irradiance maps from the images of the same person under different poses but the same illumination. The first row is the input images and the last row is the corresponding estimated irradiance maps. Evidently, the estimated irradiance maps are fairly similar though the poses are different.

在光照估计的基础上，我们最后用第5.3.3节中描述的方法估计其谐波基图像的PCA系数。同样，由于非正面姿态的图像中部分人脸区域被遮挡，这些被遮挡的点不参与运算。由于PCA系数的维数较少（本文中为89维），可见部分的人脸图像仍能鲁棒地估计出PCA系数。

估计出PCA系数后，最后根据式（5-5）重构谐波基图像。图5-11给出了根据变化姿态下的人脸图像重构的谐波基图像的例子。从图中可以看出，虽然姿态不同，但其谐波基图像重构的结果非常相似。

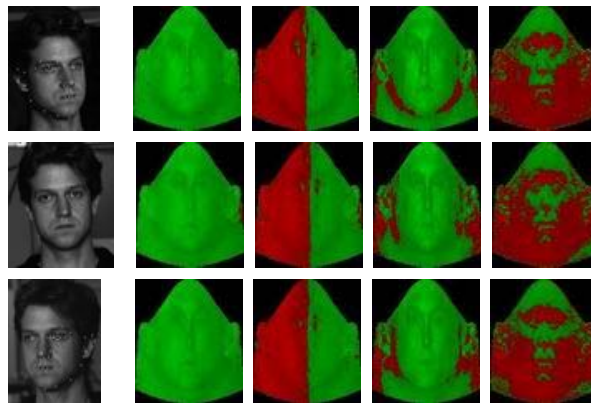


图 5-11: 根据相同人脸在不同姿态下的图像重构的球面谐波基图像。

其中第一列为输入图像，接着的四列为重构出的球面谐波基图像。

Figure 5-11. The results of reconstructed spherical harmonic images recovered from images of the same persons of three different poses. The first column is the images we used for the recovery followed by the set of the recovered spherical harmonic images.

5.6 人脸识别

当对原型人脸图像 (Gallery) 进行注册时, 我们首先估计其姿态参数和光照参数, 然后恢复出该人脸谐波图像的PCA系数, 进而可以构造该人脸的谐波基图像。在测试时, 测试图像 (Probe) 可以进行与原型图像相同的处理, 然后用PCA系数作为特征进行人脸识别; 也可以直接将测试图像用9D线性子空间的方法[Basri01a]进行识别。每个原型图像集中的人脸用一个根据原型人脸图像构造的谐波基图像为基的9D线性子空间表示。

用9D线性子空间的方法只需对人脸在注册时进行谐波基图像重构, 然后用9个谐波基图像作为子空间的基。识别时光照变化只需求与输入图像距离最近的子空间。姿态变化时可以将输入图像对应到参考空间, 然后用可见部分参与运算, 求与输入图像差别最小的子空间。其优点是对测试图像不必作太多处理。当原型图像集合中的图像的光照和姿态条件较好时, 其重构的谐波基图像相对更准确一些。而输入的测试图像的条件千变万化, 不必对其重构谐波基图像。这样系统的识别率可能会更好一些。但9D线性子空间的方法需要对每个人脸存储9个谐波基图像, 内存空间耗费大。

用人脸谐波图像的PCA系数直接作识别, 这类似用Eigenfaces[Turk91]的方法中的人脸识别。Probe中的图像处理过程同型图像集合中的图像, 得到其谐波图像的PCA系数。用 $\mathbf{a}_1, \mathbf{a}_2$ 表示两个人脸谐波图像的PCA系数向量, 这两个人脸的相似性度量为

$$d_n = \frac{\langle \mathbf{a}_1, \mathbf{a}_2 \rangle}{\|\mathbf{a}_1\| \cdot \|\mathbf{a}_2\|}, \quad (5-27)$$

以区别于原始图像之间的相似性度量 $d_I = \frac{\langle \mathbf{I}_1, \mathbf{I}_2 \rangle}{\|\mathbf{I}_1\| \cdot \|\mathbf{I}_2\|}$ 。这种识别方法的优点是对所有图像的处理过程相同, 程序编写对称, 且每幅图像对应的特征维数非常少 (为PCA系数的维数)。

最后识别时速度很快。当然从图像恢复其谐波基图像的PCA系数的过程仍然需要较长时间。另外虽然光照变化和姿态变化对谐波基图像的重构影响不大, 但毕竟还是有些影响。所以这种方法的人脸识别精度相对9D线性子空间的方法可能要低一些。

由于存储空间的限制, 本文我们选用谐波基图像的PCA系数作为人脸识别特征。本节我们选用CMU-PIE人脸数据库作为测试数据[Sim02]来验证算法的有效性。我们一共设计了两组实验: 只有光照变化的正面姿态下的人脸识别、光照和姿态都变化的人脸识别。其中后者包含了只有姿态变化而光照不变的情况。只有光照变化部分单独拿出来是因为CMU-PIE人脸数据库中光照变化部分有两个子集: “illum” 和 “lights”, 其中 “lights” 集合中姿态变化较少, 所以姿态变化的实验集我们在 “illum” 集合上进行。本文的实验图像中特征点的标注是手动完成的。

5.6.1 只有光照变化的实验结果

光照变化部分的实验条件同本文第4.3.2节。我们分别对CMU-PIE人脸数据库中正

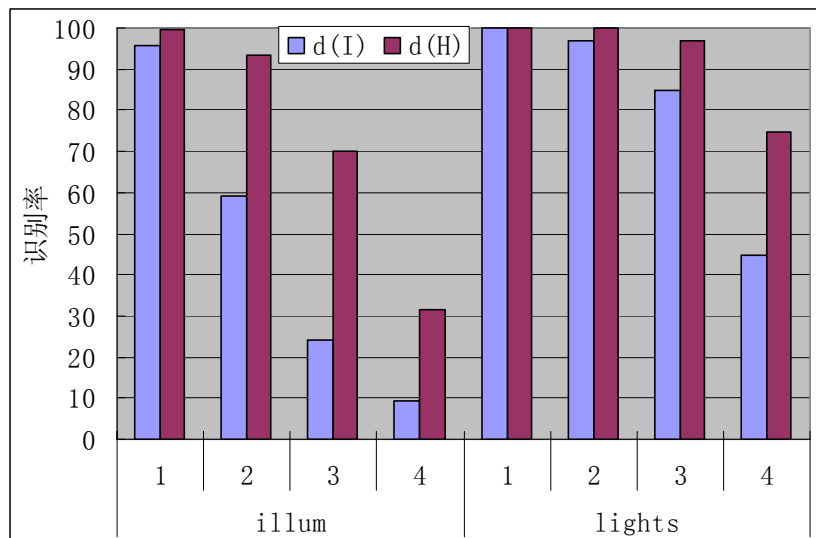
面姿态 (p27) 中的两个光照集合: “illum”集合和“lights”集合实验。实验的型图像集合分别为正面主光源 (f11) 中的图像。根据主光源偏光程度, 测试数据被分成 4 个子集, Subset1 (f06, f07, f08, f09, f12, f20)、Subset2 (f05, f10, f13, f14, f19, f21)、Subset3 (f04, f15, f18, f22) 和 Subset4 (f02, f03, f16, f17)。实验结果如表 5-1 所示。

从表中的结果可以看出, 用输入人脸的谐波基图像的 PCA 系数作特征比用原始图像直接作为特征的人脸识别率有了很大提高, 尤其在前三个光照条件稍好的子集上 (Subset1、Subset2 和 Subset3)。在光照条件较差的 Subset4 子集结果稍差, 原因之一是光照条件较差时, 由于人脸表面不是严格的凸表面, 此时投射阴影的影响较大。而这些投射阴影中的点参与谐波基图像重构时会引入噪声。如果在光照估计之后, 进行检测投射阴影的工作, 然后这些投射阴影中的像素不参与运算的话, 我们期待谐波基图像的重构结果会更好, 从而人脸识别的结果更好一些。

表 5-1: CMU-PIE 人脸数据库上不同光照下的人脸识别结果。

Table 5-1. Recognition rates on the CMU-PIE face database with different illuminations.

相似性 度量	识别率 (%)									
	illum					lights				
	Subset No.					Subset No.				
	1	2	3	4	平均	1	2	3	4	平均
d_I	95.8	59.1	24.3	9.2	52.9	100	96.8	84.9	44.9	85.7
d_H	99.5	93.4	69.9	31.6	78.2	100	100	96.7	74.6	94.5



5.6.2 姿态和光照同时变化的实验结果

在姿态和光照同时变化的这部分实验中, 我们选取 CMU-PIE 人脸数据库中的两个姿态变化子集: c05 (左转 22.5°) 和 c29 (右转 22.5°) 作为测试集, 其中每个姿态下又包括 21 种光照条件 (“illum”集合)。人脸识别时原型图像集合选择正面姿态图像 (c27) 中 “illum”集合的正面点光源 (f11) 下的图像。实验结果如表 5-2 所示, 其中 f11 所在行可以看成是只有姿态变化的情况。

从表中的结果可以看出：

(1) 当只有姿态变化时，用输入人脸的谐波基图像的 PCA 系数作特征的人脸识别率大约为 80% 左右。

(2) 当光照相对较差（投射阴影较多）时，人脸识别率的较低（小于 20%），如姿态 c05 和姿态 c29 中的(f02, f03, f16, f17)。不过这在真实情况中可能较少。

(3) 当姿态可见部分的光照相对较好（投射阴影较少）时，人脸识别率约在 50%。如姿态 c05 和姿态 c29 其余的光照条件。

综合这种情况，如果存储空间允许，只对正面姿态下光照条件较好的图像注册，重构其谐波基图像，人脸识别时用 9D 线性子空间的方法，人脸识别的性能应该会更好一些。

表 5-2: CMU-PIE 人脸数据库上不同光照和不同姿态下的人脸识别结果。

Table 5-2. Recognition rates on the CMU-PIE face database with different illuminations and different poses.

F\C	c05 (d_l)	c05 (d_H)	c29 (d_l)	c29 (d_H)
f02	0.044	0.118	0.015	0.162
f03	0.059	0.191	0.029	0.176
f04	0.103	0.353	0.059	0.265
f05	0.397	0.441	0.103	0.529
f06	0.735	0.735	0.162	0.75
f07	0.676	0.559	0.118	0.662
f08	0.544	0.691	0.588	0.809
f09	0.235	0.603	0.676	0.706
f10	0.324	0.368	0.088	0.426
f11	0.676	0.780	0.838	0.853
f12	0.309	0.662	0.765	0.706
f13	0.074	0.441	0.221	0.441
f14	0.088	0.5	0.235	0.485
f15	0.029	0.279	0.059	0.309
f16	0.029	0.162	0.044	0.235
f17	0.015	0.176	0.029	0.147
f18	0.250	0.368	0.074	0.382
f19	0.647	0.632	0.118	0.735
f20	0.662	0.779	0.838	0.809
f21	0.265	0.662	0.706	0.632
f22	0.074	0.412	0.118	0.338
平均	0.296	0.457	0.280	0.485

5.7 总结和讨论

本章提出的EigenHarmonics模型能处理人脸识别中的光照和姿态同时变化的问题，而光照变化和姿态变化被公认是人脸识别中的两大难点。实验结果表明EigenHarmonics模型对光照变化和姿态变化下人脸识别性能都有了一定提高。与其他这方面基于模型的方法相比，EigenHarmonics模型有如下特点：

(1) 输入简单，只需一幅任意光照、任意姿态下的输入图像，而光照锥方法[Georghiades01]需要输入7幅正面姿态下不同光照的图像。

(2) 训练简单，只需训练人脸的3D结构。而[ZhangL03]中需要训练人脸的3D结构和训练人脸在多个已知光照下的图像。

(3) 光照参数很少，且可以表示任意类型的光照，包括室外光照。而对称SFS[Zhao00c]中只能表示一个点光源的光照。且参数少使得我们可以用平均人脸模型就能得到很好的光照估计结果，从而不必作迭代优化过程。

(4) 参数求解简单。在光照和姿态参数确定后，求解谐波基图像的参数是一个线性问题，参数求解简单。而在3D变形模型[Blanz03]中对形状和纹理分别建模，两个模型的参数之间的关系复杂，只能用迭代优化的方法来分别优化各个参数。

但我们毕竟只是在光照和姿态同时变化方面做了一些初探性的工作，一个方法的成熟需要更细致的工作。所以本章中实验结果还不是非常令人满意，只是让我们看到了一线曙光。这方面的工作可以从以下多个方面做进一步探索：

(1) 本文中我们的EigenHarmonics模型是从100个3D人脸上统计学习而得。这样还不足够覆盖人脸的丰富变化。由于谐波基图像的模型的维数很高，这一矛盾显得更突出。通过采集更多的3D人脸样本扩大训练集，模型的表达能力应该会更好。但是当前的3D扫描仪相对还是比较昂贵，在很多情况下获得更多的样本可能是一个过于奢侈的要求。为了解决训练样本有限的问题，3D变形模型[Blanz03]采用分区域(region-based)的做法。这种做法在EigenHarmonics模型中应该也能发挥类似作用。另外为了减少原始数据的维数，计算中我们将顶点进行了下采样，这也损失了一些精度。

(2) EigenHarmonics很方便扩充到多幅图像。如果某个人脸有不同光照或姿态下的多幅图像，随着输入条件的增多，最后求解EigenHarmonics的PCA系数的结果也应该越稳定。EigenLightFields[Gross02a]已经证明多幅图像有利于提高人脸识别系统的性能，这也和人的认知学习类似，从更多的信息能更准确地认识事物。

(3) 3D变形模型[Blanz03]为了克服计算中的误差，还对PCA系数进行LDA判别分析，并用实验结果证明这种方法比直接用PCA系数进行识别的效果更好一些。本文中相同人脸的不同姿态和不同光照下的图像估计出的PCA系数也会稍有差别，如果通过训练学习这些差别的分布，用LDA进行判别分析的话，我们期望结果应该也会更好一些。

(4) 同前几章一样，谐波基图像模型本身没有考虑投射阴影，因此在投射阴影较多时，

基于谐波基图像模型的方法效果欠佳，最终导致识别效果不是很好。这需要在模型中加入额外的参数或者后处理以更好地处理投射阴影问题。Ramamoorthi等在[Ramamoorthi04]中对投射阴影用FFT变换的基图像表示，得到了一些有意思的结论。将这些结论融合到人脸识别中来也是下一个研究内容。

(5) 我们只是求出了球面谐波基图像，根据球面谐波基图像可以计算形状和纹理，从而使得结果可以应用在更广的领域。

第六章 结束语

由于人脸识别研究有着重要的理论价值和应用价值，得到了研究者的重视。经过近四十年的发展，人脸识别技术已经取得了长足进步，已有很多实验室系统和商用系统出现。但最近 FERET 项目[Phillips00]对一些算法的评测和 FRVT[Phillips03]对一些商用系统的评测结果表明：虽然现有人脸识别系统在用户配合的理想情况下取得了比较令人满意的结果，但如在更大规模人脸数据库、摄像环境不可控、用户不配合的情况下使用，目前最好的系统的识别性也能急剧下降。因此，现有人脸识别系统还尚未成熟，尤其需要解决在非理想摄像条件（光照变化、背景变化、摄像设备差异）和用户不配合（视角变化、表情变化、饰物乃至化妆）等问题。

上述问题反映在技术层面上对算法的影响主要表现在：(1)人脸图像变化太大。人脸图像受到图像拍摄时的光照、视角、遮挡等影响。人脸还会有表情、年龄等变化。另外还有摄像机的成像参数，如相机的焦距、曝光等。这些变化使得我们很难从图像中提取人脸的本质特征。(2)人脸图像分布在高复杂的非线性流形中。上述很多变化因素导致人脸图像分布非线性流形中，因此线性子空间的方法表现不是太好。这可能有两种方法来解决：一是对图像作归一化，使得人脸图像的流形变得更简单；另一个办法是设计更强大的学习和分类算法，使得系统可以有效地对非线性流形分类。(3)人脸图像的高维性与人脸样本的矛盾。诸多变化因素使得人脸图像在一个高维空间内，但我们可得到的每个人脸的样本通常很少。因此很多算法很难泛化到未见过的样本。

针对上述问题，本文主要考虑人脸识别中的光照变化问题，同时兼顾姿态变化。

6.1 本文的主要工作

本文采用基于成像模型的方法，从人脸图像的成像模型出发，在估计输入图像的光照参数和姿态参数的后，得到与光照无关或对光照变化不敏感的特征（人脸的本质特征或根据本质特征绘制得到的标准光照下的图像）用于人脸识别，以提高人脸识别系统在光照变化下的性能。具体的，本文的主要研究内容包括以下几个部分。

(1) 任意光照下人脸图像的光照估计

光照问题可分解成两个问题：光照估计和光照补偿。光照估计是解决光照问题的第一步。基于球面谐波理论，朗伯反射对光照而言是一个低通滤波器。因此表示影响朗伯反射的光照用其低频分量表示即可。光照估计为估计光照的这些低频分量的系数。这种光照表示可表示任何类型的光照环境，包括室外光照环境。

由于光照参数的维数很少（9 维），且光照对人脸图像的影响是全局的，在光照估计时我们可以采用很多假设：(1)人脸表面可近似为朗伯凸表面；(2)不同人脸的形状基本相

同；(3)人脸的各点反射率近似相等。在这些假设的基础上，本文给出了任意人脸图像的光照估计，即光照环境的低三阶，共 9 个频域分量的系数。

为了能有效地估计光照，我们还给出了尽量符合上述假设的操作：(1)形状用基于特征点的变形，以得到尽可能与输入人脸一致的人脸形状；(2)降低分辨率。尤其是降低分辨率后人脸的形状差异更小了，使得光照估计不十分依赖特征点定位的结果。这对实用系统很重要，因为实用系统在光照变化剧烈时特征点的定位通常都不太准确。且由于光照的参数很少，光照估计结果对图像分辨率变化不敏感。

总而言之，本文可对任意输入人脸图像进行光照估计，光照估计对人脸图像分辨率及配准精度的不敏感，且可处理任何类型的光照，包括室外光照。

(2) 人脸图像光照补偿

为了实现光照不变的人脸识别，在光照估计的基础上，我们对输入人脸图像进行光照补偿，然后用于人脸识别（光照估计的另一个作用是告诉我们在合适的光照变化范围内可以不作光照补偿直接进行人脸识别）。光照补偿的目的是将不同光照下的人脸图像做归一化，使得不同光照下相同人脸的图像变得更相似，从而使得人脸图像的流形分布更简单。这样人脸识别更容易。

与光照估计的全局过程不同，光照补偿是一个局部过程，即光照补偿是对图像中每个像素单独进行，因此需要输入人脸每个像素点的形状和纹理信息。

本文提出了两种光照补偿算法：基于绘制的光照补偿和差图像光照补偿。基于绘制的光照补偿通过将不同光照下的输入人脸图像都绘制到预先定义的标准光照下，从而使所有的人脸图像在相同的光照条件下进行比较。本文讨论了两种标准光照：均匀光照和“环境光+点光源”，其中均匀标准光照下的图像只保留了输入人脸的纹理信息，而“环境光+点光源”标准光照下的图像既包含纹理信息，又包含了形状信息。差图像为原图像与平均人脸在输入光照下的图像的差。由于二者光照相同，这样差图像中光照变化的影响得到了减弱，同时又保留了用于识别的输入人脸与平均人脸的差别。

不同的算法对人脸形状的处理稍有差异。主要表现在：(1)基于绘制的方法在计算中需要显式地利用每点的形状信息，尤其是标准光照为“环境光+点光源”时标准图像中还保留了形状信息。(2)差图像方法不需要显式地利用输入人脸的纹理和形状信息，只是隐式地利用编码在输入图像中的纹理和形状信息。由于输入人脸的形状信息是通过将平均人脸模型用基于特征点的变形得到，输入人脸的形状主要依赖于特征点的定位。因此基于绘制的光照补偿对特征点定位（配准）要求更严格。

基于绘制的光照补偿后的标准图像中不再包含输入图像的光照的信息，而差图像中还包含少量的光照信息，因此绘制的光照补偿在精确配准的情况下人脸识别的效果稍好，其中又以标准光照为“环境光+闪光灯”时，识别率稍高，因为该标准形式中不仅包含纹理信息，还包含形状信息。差图像并不需要精确配准，因此差图像更适合实用系统，且其计算也简单。

本文还给出了基于图像绘制方法在人脸检测中用于扩充训练样本集的应用。因为我们虽然能收集到数量足够多的样本，却很难得到质量足够好的样本，如训练集中样本的光照覆盖不充分。初步的实验结果证明了用合成虚拟样本扩充训练集的方法能提高人脸检测器的性能。

(3) 基于特征谐波基图像模型的光照和姿态不变的人脸识别

由于人脸图像通常同时存在光照变化和姿态变化，本文还提出了一个特征谐波基图像（EigenHarmonics）模型，用于光照和姿态同时变化的人脸识别。EigenHarmonics模型包含两部分：谐波基图像和PCA模型。

谐波基图像包含了人脸的本质特征：形状和反射率，且谐波基图像适合处理人脸图像中的光照问题。同时谐波基图像中包含了人脸全部区域的信息，各种姿态下的人脸可以看作人脸区域可见的某部分。因此我们可以用人脸的谐波基图像用来表示人脸的本质特征并用于不同光照和姿态下的人脸识别。

从一幅未知光照的人脸图像恢复人脸的形状和纹理是一个病态问题，从中恢复其谐波基图像也同样困难。因此我们对所有人脸的谐波基图像空间用PCA建模（Eigen模型）。由于PCA系数数目相对图像像素数目很少，根据姿态变化的人脸图像中的人脸的可见部分仍可恢复出该人脸的谐波基图像的PCA系数，因此该模型可兼顾姿态变化。

根据输入的人脸图像，我们在成像参数的基础上，恢复该人脸的谐波基图像。在成像参数估计（包括光照参数估计和姿态参数估计）时，我们平均的3D人脸模型来代替特定的输入人脸的3D模型。因为成像参数对图像的影响是整体的且参数较少，用一个平均人脸模型代替输入人脸模型是可行的。其中姿态参数估计主要是为了将输入图像与模型配准。在估计出成像参数后，重构输入人脸的谐波基图像通过最小二乘法解一个线性方程组即可得到。

实验结果表明EigenHarmonics模型对光照变化和姿态变化下人脸识别，性能都有了一定提高。与其他这方面基于模型的方法相比，EigenHarmonics模型有如下特点：(1)输入简单，只需一幅任意光照、任意姿态下的输入图像。(2)光照参数很少，且可以表示任意类型的光照，包括室外光照。且参数少使得我们可以用平均人脸模型就能得到很好的光照估计结果，不必作迭代优化过程。(3)参数求解简单。在光照和姿态参数确定后，求解谐波基图像的参数是一个线性问题，参数求解简单。

6.2 将来的工作

要彻底解决人脸识别中的光照变化和姿态变化问题，基于3D分析的人脸识别是一个最佳的选择。近些年来3D人脸分析发展很快，这得益于计算机视觉和计算机图形学两方面的共同推动。因为3D人脸分析的核心是从图像中恢复输入人脸的本质特征（形状和纹理）和成像参数（光照参数和姿态参数），这也是计算机视觉和计算机图形学共同关心的内容。

由于人脸识别的输入通常只有一幅图像，要从一幅图像中恢复其3D信息必须借助先验知识。因此寻找更合适的人脸类的先验知识的表示是人脸识别的问题之一。现在的方法大多是直接使用人脸类的各种限制或假设人脸类的分布符合高斯模型，通过训练学习高斯模型的参数。

另外人脸识别的输入通常是光照和姿态不理想的图像，因此需要从非理想输入条件得到人脸的模型。现有的基于3D分析的人脸识别方法中大多只考虑简单的成像模型（包括本文的谐波基图像模型），如没有考虑投射阴影和镜面反射等。由于人脸形状和反射属性的复杂性，这些因素在人脸图像中不可避免地存在。尤其在偏光严重时，投射阴影较多，这时基于谐波基图像模型的方法效果欠佳，最终导致识别效果不是很好。这需要在模型中加入额外的参数或者后处理以更好地处理投射阴影问题。投射阴影是人脸识别领域至今一个较少涉及的问题，虽然在图形学中用来估计光照已经有了一些工作[Sato99a] [Sato99b]。

本文给出的光照估计已经比较准确。准确的光照估计为后续的精确的3D形状和纹理重构提供了依据。如在光照估计之后，可检测出人脸图像中的镜面反射高光区和投射阴影区，为脸型形状和纹理恢复提供有效依据。

准确的特征点定位有助于得到更准确的3D形状，进而提高人脸识别率。因此提高特征点定位的准确性也是下一步工作之一，尤其当光照条件欠佳时。第四章中 attached 阴影区中光照补偿中的噪声较为严重，降低光照补偿中的噪声也是下一个研究内容。另外在光照和姿态同时变化方面我们还只是做了一些初探性的工作，如第五章所述这方面的工作在多个方面都应做进一步探索。

另外，有些应用需要更高级别的安全性，而输入可以更多，如输入的是人脸的3D结构、不同拍摄环境的多幅图像、图像可以是高分辨率的等[FRGC05]。怎样利用这些信息得到识别率更高的人脸识别系统也是下一步工作。

参考文献

- [Adini97] Y. Adini, Y. Moses, S. Ullman. Face Recognition: The Problem of Compensating for changes in illumination Direction. IEEE Trans. on PAMI, Vol. 19, No.7, pp.721-732, 1997.
- [Atick96] J. Atick, P. Griffin, A. Redlich. Statistical Approach to Shape from Shading: Reconstruction of 3D Face Surfaces from Single 2D Images. Journal of Neural Computation, Vol.8, No.6, pp.1321-1340, 1996.
- [Bailliere03] E .B. Bailliere, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Poree, B. Ruiz, and J.P. Thiran. The BANCA database and evaluation protocol. Proc. 4th Int. Conf. on Audio- and Video-based Biometric Person Authentication (AVBPA03), pp.625-638, 2003.
- [Basri01a] R. Basri, D. Jacobs. Lambertian Reflectance and Linear Subspaces. Proc. 8th IEEE Int. Conf. on Computer Vision (ICCV01), pp.383-390, 2001.
- [Basri01b] R. Basri, D. Jacobs. Photometric Stereo with General, Unknown Lighting. Proc. 2001 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR01), Vol.2, pp.374-381, 2001.
- [Basri03] R. Basri, D. Jacobs. Lambertian Reflectance and Linear Subspaces. IEEE Trans. on PAMI, Vol.25, No.2, pp.218-233, 2003.
- [Belhumeur96] P. N. Belhumeur and D. J. Kriegman. What is the Set of Images of an Object Under All Possible Lighting Conditions?. Proc. 1996 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR96), pp. 270-277, 1996.
- [Belhumeur97] P. N. Belhumeur, J. P. Hespanha and D.J. Kriegman. Eigenfaces vs Fisherfaces: recognition using class specific linear projection. IEEE Trans. on PAMI, Vol.20, No.7, pp.711-720, 1997.
- [Belhumeur98] P. N. Belhumeur and D. W. Jacobs. Comparing images under variable illumination. Proc. 1998 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR98), pp. 610-617, 1998.
- [边 01] 边肇祺, 张学工 等. 模式识别 (第 2 版). 清华大学出版社. 2001 年 5 月.
- [Blackburn01] D. M. Blackburn, J. M. Bone, and P. J. Phillips. FRVT 2000 Evaluation Report. February 2001. http://www.frvt.org/DLs/FRVT_2000.pdf.
- [Blanz99] V. Blanz and T. Vetter. Morphable Model for the synthesis of 3D faces. Proc. ACM Siggraph99, pp.187-194, 1999.
- [Blanz02] V. Blanz, S. Romdhani, and T. Vetter. Face Identification across different Poses and Illuminations with a 3D Morphable Model. Proc. 5th IEEE Int. Conf.on Automatic Face and

- Gesture Recognition (FGR02), pp.202-207, 2002.
- [Blanz03] V. Blanz and T. Vetter. Face Recognition Based on Fitting a 3D Morphable Model. IEEE Trans. on PAMI, Vol.25, No.9, pp.1-12, 2003.
- [Blinn76] J. Blinn and M. Newell. Texture and Reflection in Computer Generated Images. Communications of the ACM, Vol.19, pp.456-547, 1976.
- [Chai05] X. J. Chai, L. Y. Qing, S. G. Shan, W. Gao. Pose Invariant Face Recognition under Arbitrary Illumination based on 3D Face Reconstruction. Proc. 5th Int. Conf. Audio- and Video-based Biometric Person Authentication (AVBPA05), 2005 (to appear).
- [Chellappa95] R. Chellappa, C. L. Wilson, S. Sirohey. Human and Machine Recognition of Faces: A Survey. Proc. The IEEE, Vol.83, No.5, pp. 705-740 1995.
- [Chen00] H. F. Chen, P. N. Belhumeur, D. W. Jacobs. In Search of Illumination Invariants. Proc. 2000 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR00), Vol.1, pp.1254-1261, 2000.
- [Cootes95] T. F. Cootes, C. J. Taylor, D. Copper, J. Graham. Active Shape Models--Their Training and Application. Computer Vision and Image Understanding (CVIU), Vol.61, pp.38-59, Jan. 1995.
- [Cootes98] T. F. Cootes, G. J. Edwards, C. J. Taylor. Active Appearance Models. Proc. 5th European Conf. on Computer Vision (ECCV98), Vol.2, pp.484-498, 1998.
- [Debevec97] P. E. Debevec, J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. Proc. ACM Siggraph97, August 1997.
- [Debevec98] P. E. Debevec. Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography. Proc. ACM Siggraph98, pp.189-198, July 1998.
- [Debevec00] P. Debevec et al. Acquiring the reflectance field of a human face. Proc. ACM Siggraph00, pp.145-156, 2000.
- [Dror01] R. O. Dror, T. K. Leung, E. H. Adelson, and A. S. Willsky. Statistics of real-world illumination. Proc. 2001 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR01), Vol.2, pp.164-171, 2001.
- [Duda01] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern Classification, second edition, John Wiley & Sons, 2001.
- [Epstein95] R. Epstein, P. Hallinan and A. Yuille. 5+/-2 eigenimages suffice: An empirical investigation of low-dimensional lighting models. Proc. IEEE Workshop on Physical-Based Modeling in Computer Vision with CVPR95, pp.108-116, 1995.
- [FRGC05] <http://www.frvt.org/FRGC/>.
- [Friedman00] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. The Annals of Statistics, Vol. 28, No.2, pp.337-374, 2000.

- [傅 04] 傅祖芸, 赵梅娜, 丁岩石译, W. H. Press 等著. C 数值算法 (第二版). 国外计算机科学教材系列, 电子工业出版社, 2004.
- [Georghiades98] A. S. Georghiades, D. J. Kriegman and P. N. Belhumeur. Illumination Cones For Recognition Under Variable Lighting: Faces. Proc. 1998 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR98), pp.52-58, 1998.
- [Georghiades99] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. Illumination-based Image Synthesis: Creating Novel Images of Human Faces under Differing Pose and Lighting. Proc. IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes, pp.47-54, 1999.
- [Georghiades01] A.S. Georghiades, P.N. Belhumeur and D.J.Kriegman. From Few to Many: Illumination Cone Models for Face Recognition under Differing Pose and Lighting. IEEE Trans. on PAMI, Vol.23, No.6, pp643-660, 2001.
- [Gross02a] R. Gross, I. Matthews and S. Baker, "Eigen Light-Fields and Face Recognition Across Pose", Proc. 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition (FGR02), 2002.
- [Gross02b] R. Gross, I. Matthews, S. Baker. Fisher Light-Fields for Face Recognition across Pose and Illumination. Proc. German Symposium on Pattern Recognition (DAGM02), pp.481-489, 2002.
- [Guo00] G. Guo, S. Z. Li and K. Chan. Face Recognition by Support Vector Machines. Proc. 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition (FGR00), pp.196-201, 2000.
- [Haddon98a] J. A. Haddon, D. A. Forsyth. Shape Representations from Shading Primitives. Proc. Proc. 5th European Conf. on Computer Vision (ECCV98), Vol.2, pp.415-431, 1998.
- [Hallinan, 94] P. Hallinan. A Low-Dimensional Representation of Human Faces for Arbitrary Lightening Conditions. Proc. 1994 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR94), pp. 995-999, 1994.
- [Hertzmann01] A. Hertzmann et. al. Image analogies Proc. ACM Siggraph01, pp.327-340, 2001.
- [Horn86] B. K. P. Horn and M. J. Brooks. The Variational Approach to Shape from Shading. Computer Vision, Graphics, and Image Processing, Vol.33, No.2, pp.174-208, 1986.
- [Huang00] J. Huang. Statistics of Natural Images and Models. PhD thesis, Brown University, May 2000.
- [黄 01] 黄杏元. 地理信息系统概论(修订版). 高等教育出版社, 2001 年 12 月.
- [Inui90] T. Inui, Y. Tanabe, and Y. Onodera. Group theory and its applications in physics. 1990.
- [Ishiyama02] R. Ishiyama, S. Sakamoto. Geodesic Illumination Basis: Compensating for Illumination Variations in any Pose for Face Recognition. Proc. 16th Int. Conf. on Pattern Recognition 2002 (ICPR02), 2002.

- [姜 05] 姜大龙, 真实感三维人脸动画方法研究, 中国科学院计算所博士论文, 2005 年 1 月.
- [Jonsson00] K. Jonsson, J. Matas, J. Kittler, Y. P. Li, Learning Support Vectors for Face Verification and Recognition, Proc. the 4th Int. Conf. on Face and Gesture Recognition (FGR00), pp.208-213, Grenoble, France, 2000.
- [Jonsson02] K. Jonsson, J. Kittler, Y. P. Li, J. Matas. Support vector machines for face authentication. Image and Vision Computing. Vol.20, No.5-6, pp.369-375, 2002.
- [Kanade77] T. Kanade. Computer Recognition of Human Faces. Birkhuser Verlag, Stuttgart, 1977.
- [Lades93] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. V. D. Malsburg, R. P. Wurtz, W. Konen. Distortion Invariant Object Recognition in the Dynamic Link Architecture. IEEE Trans. on Computers, Vol.42, No.3, pp.300-311, 1993.
- [Lee01] K. C. Lee, J. Ho and D. Kreigman. Nine Points of Light: Acquiring Subspaces for Face Recognition under Variable Lighting. Proc. 2001 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR01), Vol.I, pp.519-526, 2001.
- [LiSZ04a] S. Z. Li, A. K. Jain. Handbook of Face Recognition. Springer Verlag. New York. ISBN# 0-387-40595-x, 2004
- [LiSZ04b] S. Z. Li, Z. Q. Zhang. FloatBoost Learning and Statistical Face Detection. IEEE Trans. on PAMI, Vol.26, No.9, pp. 1112-1123, 2004.
- [LiQ04] Q. Li, J. P. Ye, C. Kambhamettu. Linear Projection Methods in Face Recognition under Unconstrained Illuminations: A Comparative Study. Proc. 2004 IEEE Conf. Computer Vision and Pattern Recognition (CVPR04) Vol.2, pp.474-481, 2004.
- [LiuCJ02] C. J. Liu and H. Wechsler. Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition. IEEE Trans. on Image Processing, Vol. 11, No. 4, pp.467-476, 2002.
- [LiuCJ03] C. J. Liu and H. Wechsler. Independent Component Analysis of Gabor Features for Face Recognition. IEEE Trans. on Neural Networks, Vol.14, No.4, pp.919-928, 2003.
- [LiuZC01] Z. C. Liu, Y. Shan, Z. Zhang. Expressive Expression Mapping with Ratio Images. Proc. ACM Siggraph01, pp.271-276, Los Angeles, August 12-17, 2002.
- [马 97] 马颂德, 张正友. 计算机视觉——计算理论与算法基础. 科学出版社, 1997.
- [Marschner97] S. R. Marschner, D. Greenberg. Inverse Lighting for Photograph. The 5th Color Imaging Conference, 1997.
- [Marschner98] S. R. Marschner. Inverse Rendering for Computer Graphics. PhD thesis, Cornell University, Ithaca, NY, 1998.
- [Messer04] K. Messer, K., J. Kittler, M. Sadeghi, et al. Face Authentication Test on the BANCA Database. Proc. the 17th International Conference on Pattern Recognition (ICPR04),

pp.523-532, 2004.

[Moghaddam97] B. Moghaddam, and A. Pentland. Probabilistic Visual Learning for Object Representation. IEEE Trans. on PAMI, Vol.19, No.7, pp.775-779, July, 1997.

[Moghaddam98] B. Moghaddam, W. Wahid, and A. Pentland. Beyond eigenface: probabilistic matching for face recognition. Proc. 3rd IEEE Int. Conf.on Automatic Face and Gesture Recognition (FGR98), pp. 30-35, April, 1998.

[Moghaddam00] B. Moghaddam, T. Jebara, and A. Pentland. Bayesian Face Recognition. Pattern Recognition, Vol.33, pp.1771-1782, 2000.

[Moghaddam01] B. Moghaddam, C. Nastar, and A. Pentland. Bayesian Face Recognition with Deformable Image Models. Proc. Image Analysis and Processing, pp.26-35, 2001.

[Moghaddam02] B. Moghaddam. Principal Manifolds and Probabilistic Subspaces for Visual Recognition. IEEE Trans. on PAMI, Vol.24, No.6, pp.780-788, June 2002.

[Moses92] Y. Moses and S. Ullman. Limitation of Non-Model-Based Recognition Schemes. Proc. 2nd European Conf. on Computer Vision (ECCV92), pp.820-828, 1992.

[Moses94] Y. Moses, Y. Adini, S. Ullman. Face Recognition: the Problem of Compensating for Changes in Illumination Direction. Proc. 3rd European Conf. on Computer Vision (ECCV94), Vol.I, pp. 286-296, 1994.

[Nilius03a] P. Nilius, J.O. Eklundh. Low-Dimensional Representations of Shaded Surfaces under Varying Illumination. Proc. 2003 IEEE Conf. Computer Vision and Pattern Recognition (CVPR03), Vol.2, pp.185-192, 2003.

[Nilius03a] P. Nilius, J. O. Eklundh. Phenomenological Eigenfunctions for Image Irradiance. Proc. 9th IEEE Int. Conf. on Computer Vision (ICCV03), Vol.2, pp.568-575, October 2003.

[Olshausen96] B. A. Olshausen and D. J. Field. Emergence of simplecell receptive field properties by learning a sparse code for natural images. Nature, Vol.381, pp.607-608, 1996.

[Phillips96] P. J. Phillips, P. Rauss, S. Der. FERET (Face REcognition Technology) Recognition Algorithm Development and Test Report, ARL-TR-995, U.S. Army Research Laboratory, 1996. <http://www.frvt.org/DLs/FERET3.pdf>.

[Phillips98a] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi. The FERET database and Evaluation procedure for Face Recognition Algorithms. Imaging and Vision Computing, Vol.16, No.5, pp.295-306, 1998.

[Phillips98b] P. J. Phillips. Support vector machines applied to face recognition. In Advances in Neural Information Processing Systems, pp.803. Editors: M.C.Mozer, M.I.Jordan, and T. Petsche, MIT Press, 1998.

[Phillips00] P. J. Phillips, H. Moon, et al.. The FERET Evaluation Methodology for Face-Recognition Algorithms. IEEE Trans. on PAMI, Vol.22, No.10, pp.1090-1104, 2000.

[Phillips03] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M.

- Bone. FRVT 2002: Evaluation Report. http://www.frvt.org/DLs/FRVT_2002_Evaluation_Report.pdf, March 2003.
- [Qing04] L. Y. Qing, S. G. Shan, W. Gao. Face Recognition under Varying Lighting Based on Derivates of Log Image. The 5th Chinese Conf. on Biometric Recognition (Sinobiometrics04), Guangzhou, China. December 2004.
- [Ramamoorthi01a] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: Determining the illumination from images of a convex lambertian object. *Journal of Optical Society of America A (JOSA)*, Vol.18, No.10, pp.2448-2459, 2001.
- [Ramamoorthi01b] R. Ramamoorthi and P. Hanrahan. An efficient representation for irradiance environment maps. *Proc. ACM Siggraph01*, pp.497-500, August 2001.
- [Ramamoorthi01c] R. Ramamoorthi, P. Hanrahan. A signal-processing framework for inverse rendering. *Proc. ACM Siggraph01*, pp.117-128, August 2001.
- [Ramamoorthi02a] R. Ramamoorthi. Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object. *IEEE Trans. on PAMI*, Vol.24, No.10, pp.1-12, 2002.
- [Ramamoorthi02b] R. Ramamoorthi, P. Hanrahan. Frequency space environment map rendering. *Proc. ACM Siggraph02*, 2002.
- [Ramamoorthi04] R. Ramamoorthi, M. Koudelka, and P. Belhumeur. A Fourier Theory for Cast Shadows. *Proc. 8th European Conf. on Computer Vision (ECCV04)*, pp.146-162, 2004.
- [Raviv99] T. R. Raviv, A. Shashua. The Quotient Image: Class Based Recognition and Synthesis under Varying Illumination Conditions. *Proc. 1999 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR99)*, Vol.2, pp.2566-2571, 1999.
- [Romdhani03] S. Romdhani, T. Vetter. Efficient, Robust and Accurate Fitting of a 3D Morphable Model. *Proc. 9th IEEE Int. Conf. on Computer Vision (ICCV03)*, Vol.1, pp.59-66, 2003.
- [Sato99a] I. Sato, Y. Sato, and K. Ikeuchi. Illumination distribution from brightness in shadows: adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions. *Proc. 7th IEEE Int. Conf. on Computer Vision (ICCV99)*, pp.875-882, September 1999
- [Sato99b] I. Sato, Y. Sato, and K. Ikeuchi. Illumination distribution from shadows. *Proc. 1999 IEEE Conf. Computer Vision and Pattern Recognition (CVPR99)*, pp.306-312, June 1999.
- [Sato01] I. Sato, Y. Sato, and K. Ikeuchi. Stability issues in recovering illumination distribution from brightness in shadows. *Proc. 2001 IEEE Conf. Computer Vision and Pattern Recognition (CVPR01)*, pp.II-400-407, December 2001.
- [Sato03a] I. Sato, Y. Sato, and K. Ikeuchi. Illumination from shadows. *IEEE Trans. on PAMI*, Vol.25, No.3, pp.290-300, March 2003.

- [Sato03b] I. Sato, Y. Sato, and K. Ikeuchi. Appearance Sampling for Obtaining A Set of Basis Images for Variable Illumination. Proc. 9th IEEE Int. Conf. on Computer Vision (ICCV03), pp.800-807, 2003.
- [Savvides04] M. Savvides, B. V. Kumar and P.K. Khosla. Eigenphases vs. Eigenfaces. Proc. 17th Int. Conf. Pattern Recognition (ICPR04), pp.810-813, 2004.
- [Shan02] S. G. Shan, W. Gao, W. Wang, et al.. Enhanced Active Shape Models with Global Texture Constraints for Face Image Analysis. Proc. 14th Int. Symposium on Methodologies for Intelligent Systems (ISMIS03), pp.593-597, 2003.
- [Shan03] S. G. Shan, W. Gao, B. Cao, D. B. Zhao. Illumination Normalization for Robust Face Recognition against Varying Lighting Conditions. Proc. IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures (AMFG03), pp.157-164, 2003.
- [山 04] 山世光. 人脸识别中若干关键问题的研究. 中国科学院计算所博士学位论文. 2004年5月.
- [Shashua94] A. Shashua. Geometry and Photometry in 3D Visual Recognition. PhD Thesis, MIT, 1994.
- [Shashua97] A. Shashua. On Photometric Issues in 3D Visual Recognition from a Single 2D Image. Int. Journal of Computer Vision (IJCV), Vol.21, No.1-2, pp.99-122, 1997.
- [Shashua01] A. Shashua and T. R. Raviv. The Quotient Image: Class-Based Re-Rendering And Recognition With Varying Illuminations. IEEE Trans. on PAMI, Vol.23, No.2, pp.129-139, 2001.
- [石 02] D. F. Rogers 著, 石教英, 彭群生等译. 计算机图形学的算法基础. 第二版, 机械工业出版社, 2002.
- [Sim01] T. Sim, T. Kanade. Combining Models and Exemplars for Face Recognition: An Illuminating Example. Proc. Workshop on Models versus Exemplars in Computer Vision with CVPR01, 2001.
- [Sim02] T. Sim, S. Baker and M. Bsat. The CMU Pose, Illumination, and Expression (PIE) Database. Proc. 5th IEEE Int. Conf.on Automatic Face and Gesture Recognition (FGR02), Washington, DC, May 2002.
- [Simakov03] D. Simakov, D. Frolova, R. Basri. Dense Shape Reconstruction of a Moving Object under Arbitrary, Unknown Lighting. Proc. 9th IEEE Int. Conf. on Computer Vision (ICCV03), pp. 1202-1209, October 2003.
- [Sloan02] P.P. Sloan, J. Kautz, J. Snyder. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. Proc. ACM Siggraph02, pp.527-536, 2002.
- [Tannoudji77] C. C. Tannoudji and et al. Quantum Mechanics. John Wiley & Sons, 1977.
- [Tefas01] A. Tefas, C. Kotropoulos, and I. Pitas. Using Support Vector Machines to Enhance the Performance of Elastic Graph Matching for Frontal Face Authentication, IEEE Trans. on

- PAMI, Vol.23, No.7, pp.735-746, July 2001.
- [Turk91] M. Turk and A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, Vol.3, No.1, pp.71-86, January 1991.
- [Vetter97] T.Vetter and T.Poggio. Linear Object Classes and Image Synthesis from A Single Example Image. *IEEE Trans. on PAMI*, Vol.19, pp.733-742, 1997.
- [Viola01] P. Viola, M. Jones. Rapid Object Detection using a Boosted Cascade of Simple. *Proc. 2001 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR01)*, Vol.1, pp.511-518, 2001.
- [WangW02] W. Wang, S. Shan, W. Gao, et. al.. An Improved Active Shape Model for Face Alignment. *Proc. 4th Int. Conf. on Multimodal Interfaces (ICMI02)*, pp.523-528, 2002.
- [WangHT03] H T. Wang, S. Z. Li, Y. S. Wang, W. W. Zhang. Illumination Modeling and Normalization for Face Recognition. *Proc. IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures (AMFG03)*. 2003.
- [WangHT04a] H. T. Wang, S. Z. Li, Y. S. Wang. Face Recognition under Varying Lighting Conditions Using Self Quotient Image. *Proc. 6th IEEE Int. Conf.on Automatic Face and Gesture Recognition (FGR04)*, pp.819-824, 2004.
- [WangHT04b] H. T. Wang, S. Z. Li, Y. S. Wang. Generalized Quotient Image. *Proc. 2004 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR04)*, Vol. 2, pp.498-505, 2004.
- [WangY02] Y. Wang, D. Samaras. Estimation of Multiple Directional Light Sources for Synthesis of Mixed Reality Images. *Proc. 10th Pacific Conference on Computer Graphics and Applications (PG02)*, pp.38-47, 2002.
- [WangXG03] X. G. Wang, and X. O. Tang. Unified Framework for Subspace Face Recognition. *Proc. 9th IEEE Int. Conf. on Computer Vision (ICCV03)*, Vol.1, pp.679-686, 2003.
- [WangXG04] X. G. Wang, and X. O. Tang. A Unified Framework for Subspace Face Recognition. *IEEE Trans. on PAMI*, Vol.26, No.9, pp.1222-1228, 2004.
- [Wei04] S. D. Wei, S. H. Lai. Robust Face Recognition under Lighting Variations. *Proc. 17th Int. Conf. on Pattern Recognition (ICPR04)*, Vol.1, pp.354-357, 2004.
- [Weiss01] Y. Weiss. Deriving intrinsic images from image sequences. *Proc. 8th IEEE Int. Conf. on Computer Vision (ICCV01)*, Vol.II, pp.68-75, Vancouver, Canada, July 2001.
- [Wen03] Z. Wen, Z. Liu and T. S. Huang. Face Relighting with Radiance Environment Maps. *Proc. 2003 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR03)*, Vol.II, pp.158-165, 2003.
- [Wiskott97] L. Wiskott, J. M. Fellous, N. Kr uger, and C. von der Malsburg. Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. on PAMI*, Vol.19, No.7, pp.775-779, 1997.
- [Worthington01] P. L. Worthington, E. R. Hancock. Object Recognition Using Shape-from-

- Shading, IEEE Trans. on PAMI, Vol.23, No.5, pp.535-542, May 2001.
- [Xie05] X. Xie, K. Lam. Face Recognition under Varying Illumination Based on a 2D Face Shape Model. Pattern Recognition, Vol.38, No.1, pp.221-230, 2005.
- [Yang02] M. H. Yang, D. Kriegman, and N. Ahuja. Detecting Faces in Images: A Survey. IEEE Trans. on PAMI, Vol.24, No.1, pp.34-58, Jan. 2002.
- [Yuile97] A. Yuile and D. Snow. Shape and Albedo from Multiple Images Using Integrability. Proc. 1997 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR97), pp.158-164, 1997.
- [张00] 张翠平, 苏光大. 人脸识别技术综述. 中国图形图像学报, Vol.5, 2000.
- [ZhangL03] L. Zhang, D. Samaras. Face Recognition Under Variable Lighting using Harmonic Image Exemplars. Proc. 2003 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR03), Vol. I, pp.19-25, June 2003.
- [张 03] 张军平, 流形学习及应用, 中科院自动化所博士学位论文, 2003.
- [ZhangR99] R. Zhang, P.S. Tai, J. E. Cryer, M. Sha. Shape From Shading: A Survey. IEEE Trans. On PAMI, Vol.21, No.8. pp.690-706, 1999.
- [ZhangYF00] Y. F. Zhang, Y. H. Yang. Illumination Direction Determination for Multiple Light Source. Proc. 2000 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR00), pp. 269-276, 2000.
- [ZhangYF01] Y. F. Zhang, Y. H. Yang. Multiple Illumination Direction Detection with Application to Image Synthesis. IEEE Trans. on PAMI, Vol.23, No.8, pp.915-920. August 2001.
- [Zhao98] W. Y. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant Analysis of Principal Components for Face Recognition. Proc. 3rd IEEE Int. Conf.on Automatic Face and Gesture Recognition (FGR98), pp.336-341, 1998.
- [Zhao00a] W. Y. Zhao, R. Chellappa, A. Rosenfeld, et al. Face recognition: A literature survey. UMD CfAR Technical Report CAR-TR-948, 2000.
- [Zhao00b] W. Y. Zhao and R. Chellappa. Illumination-insensitive Face Recognition using Symmetric Shape-from-Shading. Proc. 2000 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR00), Vol.1, pp.286-293, 2000.
- [Zhao00c] W. Y. Zhao, R. Chellappa. SFS Based View Synthesis for Robust Face Recognition, Proc. 4th Int. Conf.on Face and Gesture Recognition (FGR00), pp.285-292, 2000.
- [Zheng91] Q. Zheng and R. Chellapa. Estimation of Illuminant Direction, Albedo, and Shape from Shading", IEEE Trans. on PAMI, Vol.13, No.7, pp.680-702, July 1991.
- [周00] 周杰, 卢春雨, 张长水, 李衍达. 人脸自动识别方法综述. 电子学报, 2000年第4期.
- [Zhou04] K. Zhou and R. Chellappa. Illuminating Light Field: Image-based Face Recognition across Illuminations and Poses. Proc. 6th IEEE Int. Conf. on Automatic Face and Gesture

Recognition (FGR04), 2004.

致 谢

值此论文完成之际，首先衷心感谢我的导师高文教授对我的亲切关怀和悉心指导。从博士论文选题到整个论文的撰写，高老师给予我大量的指导和帮助，使得我的研究工作得以顺利的开展并取得一定的成绩，我从中受益颇深。在高老师指导下的几年的博士研究生的学习和研究使得我对科学研究工作有了更深的体验。高老师广博的科学知识，精深的学术思想，高瞻远瞩的方向感，严谨求实的科学精神，宽大包容的胸襟，高标准严要求的工作作风，给我打下了深刻的烙印，永远是我学习的榜样。

诚挚地感谢赵德斌教授、陈熙霖教授、黄庆明教授、黄铁军研究员、贺思敏博士、崔国勤副研究员以及姚鸿勋教授。这些老师是我在联合实验室工作学习的良师益友，我能够时刻感受到他们对我的工作、学习、生活诸方面的关心、支持和帮助。特别感谢陈熙霖教授对我的学术研究、工作方法、日常生活诸方面的指导和帮助。陈教授精深的学术思想、广博的知识面、严谨的科学精神，一直是我的楷模，鞭策着我前进。

感谢在联合实验室与我朝夕相处多年的学长或同学们：王伟强博士、王春立博士、周德龙博士、廖频博士、段立娟博士、王实博士、陈益强博士、吕岩博士、孙晓艳博士、庞斌博士、曾炜博士，以及博士生马斯伟、孙书韬、王耀威、张远、张京芬、付岩和苏荔等。感谢人脸组和我一起并肩战斗的的同伴们：山世光、苗军、富小薇、柴秀娟、杜波、陈杰、李月敏、张文超、曹波、杨澎、唱轶钲、张晓华、张洪明、王建宇、阎胜业、刘心、赵三强、张宝昌、马丙鹏，唐杰、杨非、苏煜、牛志恒、王乐珩。大家的辛勤劳动使得我的研究有一个良好的基础平台，人脸组融洽的氛围让我在愉快中完成了我的论文工作。我尤其感谢我的组长山世光博士，他对我有知遇之恩并在论文的各个环节上都给予了我大力的帮助和支持。

感谢联合实验室办公室的张青老师，羌滨健老师，蔡光辉老师，李丹老师，刘敬老师，董强老师，李伟老师和张亚冬女士。他们的辛勤劳动为联合实验室提供了良好的研究环境，也为我个人的工作、学习和生活提供了极大的便利和许多帮助，在此表示衷心的感谢。感谢研究生院信息学院、教务处和学生处的老师们，在我几年的工作学习生活中，他们给了我大力的支持，在此表示深深的感谢。

感谢我的父母和兄弟姐妹对我追求学业、从事科研工作的理解和支持，他们的关爱是我前进的动力源泉。最后，让我将最诚挚的感谢献给我的爱人，他始终一贯地支持我的求学历程，陪我一起走过求学中最艰苦的日子，是我能够顺利完成学业的强大后盾和有力保障。还有我稚嫩乖巧的小宝贝，在我几年的博士学习中一直照顾不周，心怀愧疚。

无以为报，谨以此文敬献内心的歉疚和感激之情。再一次深深地感谢所有培养、支持、关心、帮助过我的善良的人们，并衷心地祝福他们！

作者简介

姓名：卿来云 性别：女 出生日期：1974.04.02 籍贯：湖南邵东

2000.09 -- 2005.07 中科院研究生院计算机软件与理论专业 博士研究生
1999.04 -- 2000.08 广州金鹏集团有限公司 软件工程师
1996.09 -- 1999.03 东北大学计算机系计算机软件与理论专业 硕士研究生
1992.09 -- 1996.07 东北大学计算机系计算机应用专业 本科生

【攻读博士学位期间发表的论文】

国际期刊：

- [1] Laiyun Qing, Shiguang Shan, Wen Gao. Face Recognition under Generic Illumination Based on Harmonic Relighting. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.19, No.4, 2005.
- [2] Yuemin Li, Jie Chen, Laiyun Qing, Wen Gao, Baocai Yin. Face Detection under Variable Lighting Based on Nine Point of Light. *Asian Journal of Information Technology*. Vol.4, No.3, pp.49-55, 2005.

国际会议：

- [3] Laiyun Qing, Shiguang Shan, Wen Gao. Eigen-Harmonics Faces: Face Recognition under Generic Lighting. *Proc. 6th IEEE International Conference on Automatic Face and Gesture Recognition (FGR04)*, pp.296-301, May 2004.
- [4] Laiyun Qing, Shiguang Shan, Wen Gao. Face Recognition with Harmonic De-lighting. *Proc. Asian Conference on Computer Vision (ACCV04)*, pp.824-829, Jan. 2004.
- [5] Laiyun Qing, Shiguang Shan, Xilin Chen. Face Relighting for Face Recognition under Generic Illumination. *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP04)*. Vol.V, pp.733-736, May 2004.
- [6] Laiyun Qing, Wen Gao, Shiguang Shan. Surface Reflectance Classifying under Natural Illumination, *Proc. 4th IEEE Pacific-Rim Conference on Multimedia (PCM03)*, pp.753-758, Dec. 2003.
- [7] Laiyun Qing, Weiqiang Wang, Wen Gao. Illumination Invariant Shot Boundary Detection. *Proc. 4th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL03)*, pp.1097-1101, March, 2003.
- [8] Laiyun Qing, Weiqiang Wang, Wen Gao. A Framework for Background Detection in Video. *Proc. 3rd IEEE Pacific-Rim Conference on Multimedia (PCM02)*, pp.799-805, Dec. 2002.
- [9] Xiujuan Chai, Laiyun Qing, Shiguang Shan, Wen Gao. Pose Invariant Face Recognition under Arbitrary Illumination based on 3D Face Reconstruction. *Proc. 5th Int. Conf. Audio- and Video-based Biometric Person Authentication (AVBPA05)*, (to appear).
- [10] Bo Du, Shiguang Shan, Laiyun Qing, Wen Gao. Empirical Comparisons of Several

- Preprocessing Methods for Illumination Insensitive Face Recognition. *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP05)*, Vol.II, pp.981-984, 2005.
- [11] Jie Chen, Yuemin Li, Laiyun Qing, Baocai Yin, Wen Gao. Face Samples Re-lighting for Detection Based on the Harmonic Images. *Proc. 5th Pacific Rim Conference on Multimedia (PCM04)*, Part II, pp.585-590, Nov. 2004.
- [12] Yuemin Li, Jie Chen, Laiyun Qing, Baocai Yin, Wen Gao. Face Detection under Variable Lighting Relighting Based on Resample by Face. *Proc. Int. Conf. on Machine Learning and Cybernetics (ICMLC04)*, Vol.7, pp.3775-3780, 2004.
- [13] Weiqiang Wang, Laiyun Qing, Yan Fu, Wen Gao. Locating Anchor Shots in Compression Domain Based on Neural Networks. *Proc. 5th Asian Conference on Computer Vision (ACCV02)*, pp.50-55, Jan. 2002.

国内会议:

- [14] Laiyun Qing, Shiguang Shan, Wen Gao. Face Recognition under Varying Lighting Based on Derivates of Log Image. *Advances in Biometric Personal Authentication (Sinobiometrics'04)*, pp.196-204, Dec. 2004.
- [15] Laiyun Qing, Shiguang Shan, Wen Gao. Face De-lighting under Natural Illumination. 第四届中国生物识别学术会议. 2003 年 12 月.
- [16] 卿来云, 山世光, 高文. 基于商图像的人脸图像光照补偿算法. 第四届中国生物识别学术会议. 2003 年 12 月.
- [17] Shiguang Shan, Wen Gao, Laiyun Qing. Generating Canonical Form Face Images for Illumination Insensitive Face Recognition. 第四届中国生物识别学术会议, 2003 年 12 月.
- [18] Yuemin Li, Jie Chen, Laiyun Qing, Wen Gao, Baocai Yin. Face Detection under Variable Lighting Based on Nine Point of Light. 第九届中国机器学习会议 (CCML'04).
- [19] 卿来云, 王伟强, 高文. 自动抽取文字及其在视频索引中的应用. 第七届中国科学院计算机科学与技术研讨会. 2002 年 7 月.

已投稿论文:

- [20] 卿来云, 山世光, 高文, 杜波. 基于球面谐波基图像的任意光照下的人脸识别. 计算机学报

【攻读博士学位期间参加的科研项目】

- [1] 基于生物特征的身份识别研究, 国家自然科学基金委重点项目, 2004-2007, 负责人脸识别中的光照变化部分, 编号: 60332010
- [2] 生物特征识别核心技术与关键问题研究, 863 计划, 2001-2003, 负责人脸识别中的光照变化部分, 编号: 2001AA114190
- [3] 多媒体掌上电脑与通信掌上电脑, 国家 863 计划, 2001-2003, 负责移动身份验证子课题中的光照变化部分, 编号: 2002AA118010
- [4] 面像识别研究, 银晨网讯科技有限公司资助项目, 2003-2006, 负责人脸识别研究中的光照变化部分
- [5] 面向 4C 融合的数字媒体处理与检索系统, 中科院知识创新方向性项目, 2000-2002
- [6] 网络视频流内容检测技术研究及系统, 国家信息安全办公室项目, 2002-2002