

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 \_\_\_\_\_

## 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.