

# 《视听信息系统导论》第一次课程大作业：

## 基于多角度单光源图像的人脸三维重建与渲染

2020 年 10 月 9 日

### 1. 题目介绍

在已知物体表面反射率、法向量方向等信息的情况下，我们可以利用表面反射模型得到不同光照和观察角度下的渲染结果。随着三维渲染技术的发展，人们提出更精确的反射模型，使得渲染结果尽可能接近真实世界的图像。我们则考虑上述流程的反向过程，即已知不同光照和观察角度的渲染结果，通过表面反射模型推测出物体的表面反射率和法向量等信息，后者可以进一步得到深度信息以实现三维重建。三维重建的结果又可以通过正向的渲染效果得以验证。然而，在一般的自然环境中，光照信息是极为复杂和难以获取的，因此实现上述推测并不容易。因此，在本次大作业中，我们提供了不同角度单一恒定强度光源照射下拍摄的正面角度人脸图像，以及光源的入射角度，从而尽可能简化推测过程。请参考课程内容，根据 1.1 节的任务描述，实现人脸的三维重建与渲染。

#### 1.1 任务描述

在固定方向单一恒定强度光源的条件下，人脸上各点的亮度与该点法向量的方向密切相关，如图 1 所示。



图 1：不同角度单光源人脸图像

为简化问题，我们考虑一种完全漫反射的模型。对于正向人脸图像的每一点 $(x, y)$ ，都有深度 $z(x, y)$ 以及漫反射系数 $k_d(x, y)$ ，令 $\mathbf{b}(x, y)$ 为点 $(x, y)$ 对应的单位法向量与漫反射系数的乘积，即：

$$\mathbf{b}(x, y) = k_d(x, y) \frac{(z_x(x, y), z_y(x, y), -1)}{\sqrt{z_x^2(x, y) + z_y^2(x, y) + 1}}$$

其中， $z_x(x, y)$ 和 $z_y(x, y)$ 分别是 $z(x, y)$ 关于 $x$ 和 $y$ 的偏导数。令光照方向为 $\mathbf{s}$ ，当 $\mathbf{s}$ 与 $z$ 轴角度不大时，得到的人脸图像 $m$ 几乎不含阴影和高光，因此其结果应接近于漫反射得到的渲染结果，即：

$$m(x, y) = \max(\mathbf{b}^T(x, y)\mathbf{s}, 0)$$

此时， $\mathbf{b}$ 的大小为 $w \times h \times 3$ （ $w$ 和 $h$ 分别为图像的宽和高），因此当训练集图像数目超过 3 时，即可通过最小二乘或其它优化算法对 $\mathbf{b}$ 进行优化，再辅以 $z$ 的连续性约束，即可估计出人脸图像各点的深度，实现人脸的三维重建。在得到三维重建结果的基础上，又可以进一步渲

染出新的光线照射角度下的人脸图像。上述思路的详细过程可参考[1]。

在本次作业中，对于每个人脸，我们提供 7 组训练数据和 10 组测试数据，每组训练数据包括人脸图像以及入射光线角度，每组测试数据包括入射光线角度。根据入射光线角度大小，我们将训练和测试数据分为 3 种难度。具体而言，我们以方位角和俯仰角（详见 1.2）的绝对值最大值为阈值，不超过 15 度为简单，超过 15 度不超过 40 度为中等，超过 40 度不超过 90 度为困难，7 组训练数据和 10 组测试数据从简单到困难的三种难度比例分别为 4: 2: 1 和 2: 5: 3。要求对于每个人脸，使用基于**完全漫反射模型**或其它**光照模型**的方法，得到其**深度图像**，并根据入射光线角度得到 10 组测试数据的**渲染结果**。

## 1.2 数据集

本次作业的数据集文件夹被命名为 `dataset`。其中包括线上和线下两个部分，分别命名为 `dataset_online`（不公开）和 `dataset_offline`，各包含三个正面角度的人脸。每个人脸对应的文件夹具体内容如下：

- `train` 文件夹，包含 7 张人脸在不同角度单光源照射下的 .bmp 格式图像，图像大小为 168\*168；
- `train.txt`，包含训练图像的名称和入射光线角度，每行对应一张训练图像，共 7 行。每行格式为“名称, 入射光线方向方位角 $\alpha$ , 入射光线方向俯仰角 $\beta$ ”，例如，“1, +020, +30”，代表图像名称为 1.bmp，入射光线朝向的方位角为 z 轴正向起始朝 x 轴正向旋转 20 度，俯仰角为 y 轴正向 30 度，方向的具体定义如图 2 所示；
- `test.txt`，包含测试图像的名称和光线角度，每行对应一张测试图像，共 10 行。每行格式同 `train.txt`。

另外，本次大作业提供一份简短的样例代码用于定义接口，以及待补全的算法核心代码，详见 `demo.m/py` 以及 `rendering.m/py`。

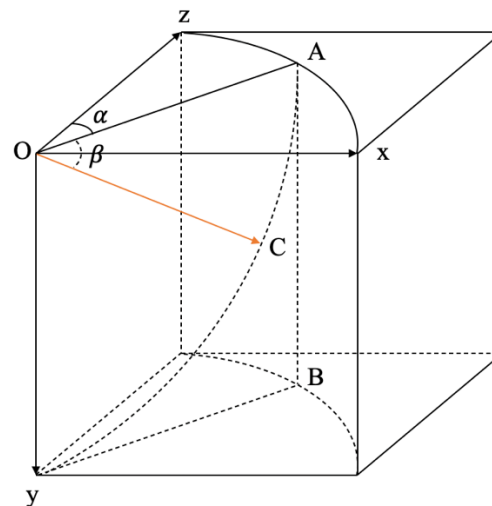


图 2：方位角与俯仰角示意，此时 OC 为入射光线方向向量

## 2. 设计要求

### 2.1 设计报告

每组提交一份设计报告，报告篇幅不得超过 **4 页 A4 纸**，报告应至少包含以下内容：

- 小组成员名单及分工情况：小组成员评分可能会因分工及完成情况产生差异。
- 提交文件清单。

- 工作开展及研究情况：应至少包含原理、实现方法、结果展示、结果分析、问题与不足，也可以包含其他任何对于解决问题有益的思考和讨论。

## 2.2 提交清单

每小组提交一份以“提交同学学号\_提交同学姓名.zip/rar”命名的压缩文件，压缩文件内至少包含：

- 设计报告（.pdf/docx/doc）。
- 补全的三维重建及渲染函数 `rendering.m/py`。
- `dataset_offline` 中每一个人脸的深度信息（保存于人脸文件夹内，命名为 `z.mat/npz`）和全部测试数据的渲染结果图像（保存于人脸文件夹内，命名为 `i.bmp`，`i` 与 `test.txt` 中的图像序号一一对应），具体格式可以参考 `demo.m/py`。
- 其他代码文件和依赖库文件。

请于网络学堂规定截止时间之前将作业提交至网络学堂，逾期提交的作业将根据逾期天数予以扣分，以满分 100 分计，每逾期一天扣除 10 分（不满一天按一天计）。

## 2.3 编程语言

本次作业要求使用 MATLAB R2019a 或 python3.6。请确保使用正确的版本以免测试失败。若使用 python3.6，请提供用于配置环境的 `setup.sh/py` 文件，以确保环境一致性。

# 3. 评分标准

本次大作业满分 100 分，占期末总评的 15%，一般情况下，组内成员得分相同。具体评分标准如下：

## 3.1 设计报告及测试代码

设计报告及测试代码总计 85 分，报告需满足以下条件：

- 符合 2.1 的要求。
- 提交的代码应能对 `dataset_offline` 数据集中的 3 组人脸数据正确运行。
- 对 3 组人脸数据的测试结果应与报告中的结果一致。
- 估计的深度结果和渲染结果应在视觉上符合真实情况。
- 渲染结果应尽可能接近真实的拍摄结果，在评判过程中会综合使用几种度量手段（如范数距离、余弦距离等）若代码无法正确运行则以报告中结果为测试结果。

不满足上述要求的报告均会被酌情扣分。

## 3.2 线上数据集的测试结果

助教使用提交的代码对 `dataset_online` 中 3 组未提供的人脸数据进行测试，每组数据 5 分，共计 15 分，根据测试结果与真实值的接近程度评分。代码无法运行则不得分。

## 参考文献

[1] Georgiades A S, Belhumeur P N, Kriegman D J. From few to many: Illumination cone models for face recognition under variable lighting and pose[J]. IEEE transactions on pattern analysis and machine intelligence, 2001, 23(6): 643-660.