

基于多角度单光源图像的人脸重建与渲染

马啸阳, 刘圣禹, 刘坤瓚

清华大学电子工程系

小组分工—马啸阳: 渲染程序框架(迭代法与法向量可积性分析)、T-S 模型实现、报告撰写; 刘圣禹: 通过法向量生成深度图、通过深度图求偏导、报告撰写; 刘坤瓚: 插值算法的实现、重建结果绘图、报告撰写。

I. 实验原理

A. Lambertian 反射模型

首先我们考虑一种简化的反射模型 [1]。假设人脸表面具有完全漫反射特性 [2], 即表面上每个点以向各个方向的反射光强相同。

记人脸表面点 (x, y) 的漫反射系数为 $\alpha(x, y)$, 且具有深度 $z(x, y)$ 。定义 $\mathbf{b}(x, y) \in \mathbb{R}^{1 \times 3}$ 表示点 (x, y) 的漫反射系数与单位法向量的乘积, 即

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

其中 $z_x(x, y)$ 和 $z_y(x, y)$ 表示在 x 和 y 方向的偏导数。

当人脸被一个无穷远处强度为 1 的点光源 $\mathbf{s} \in \mathbb{R}^{3 \times 1}$ 照射时 (\mathbf{s} 表示光源入射方向), 图像中每个像素值可以被认为这个像素的照度, 即 $\hat{x} = \max(\mathbf{b}\mathbf{s}, 0)$, 其中 \hat{x} 表示像素值。定义 $\mathbf{B} \in \mathbb{R}^{n \times 3}$ 表示包含 n 个像素的图片 $\mathbf{x} \in \mathbb{R}^{n \times 1}$ 中每个点对应的行向量 \mathbf{b} 拼成的矩阵, 则

$$\mathbf{x} = \max(\mathbf{B}\mathbf{s}, \mathbf{0}).$$

$\max(\cdot, \mathbf{0})$ 表示将得到的负数元素置零, 这些负数的元素对应着图像中的阴影。如果假设图像中不存在阴影, 那么任意角度光源 $\bar{\mathbf{s}}$ 所照射出的图像 $\bar{\mathbf{x}}$ 将落在由 \mathbf{B} 矩阵列向量张成的子空间 \mathcal{L} 里, 这个子空间 \mathcal{L} 被称为光照子空间 (illumination subspace) [3], 它与第一卦限的交集是一个凸锥 \mathcal{C} , 被称为光照凸锥 (illumination cone) [4], 是一个研究中经常使用的假设。

B. Torrance and Sparrow 反射模型

由于 Lambertian 反射模型假设了人脸表面每个点具有完全漫反射特性, 它在对图像的处理过程中不具

有一般性。另一方面, 光源角度与表面法向量存在着固有的 9-参数模糊 [5]。如果对表面法向量使用了可积性 (integrability) 约束, 这个 9-参数模糊会被化为一个 3-参数广义浮雕模糊 (Generalized Bas-Relief, GBR), 即只能恢复出形如

$$\bar{z}(x, y) = g_3 z(x, y) + g_1 x + g_2 y$$

的 $\bar{z}(x, y)$, 而不能得到实际的 $z(x, y)$, 且这种 GBR 模糊无法进一步利用光度数据解决 [6]。

文章 [6] 因此采用 Torrance and Sparrow (T-S) 反射模型解决 GBR 模糊问题。T-S 模型 [7] 假设反射系数有两个成分决定: 一是由材料本身决定的反射系数, 二是取决于表面的散射。T-S 模型将表面分划为一些微元, 每个微元的表面法向量与所有点表面法向量的偏差是一个零均值高斯随机变量, 方差越大代表表面越粗糙, 进而随着散射得到更钝的表现。据此得到双向反射分布函数 (Bidirectional Reflectance Distribution Function, BRDF)

$$r_{\text{TS}} = a_d + a_s Q F(\theta', \eta) \frac{\exp(-\nu^2 \theta_a^2)}{\cos \theta_i \cos \theta_r},$$

其中 a_d 表示 Lambertian 漫反射系数, a_s 表示镜面反射系数, ν 表示表面粗糙程度, θ_a 表示表面法向量和入射光夹角, θ_i 与 θ_r 的定义与 BRDF 定义相同, Q 表示几何衰减系数, $F(\theta', \eta)$ 表示 Fresnel 反射系数 [8]。由此得到由光源 \mathbf{s} 照射得到的强度为

$$I_{\text{TS}} = a_d \|\mathbf{s}\| \cos \theta_i + a_s \|\mathbf{s}\| Q F(\theta', \eta) \frac{\exp(-\nu^2 \theta_a^2)}{\cos \theta_r}.$$

在实际求解中, 几何衰减系数 Q 和 Fresnel 反射系数 $F(\theta', \eta)$ 可以不用考虑。相比于 Lambertian 反射模型, T-S 模型的照射强度为

$$I_{\text{TS}}(\bar{\mathbf{b}}(x, y), \bar{\mathbf{s}}_i) = \bar{\mathbf{b}}^\top \mathbf{G}^{-1} \mathbf{G} \bar{\mathbf{s}}_i + \frac{a_s \|\mathbf{G} \bar{\mathbf{s}}_i\| \exp \left(-\nu^2 \left[\cos^{-1} \left(\frac{\bar{\mathbf{b}}^\top \mathbf{G}^{-1} \left(\frac{\mathbf{G} \bar{\mathbf{s}}_i}{\|\mathbf{G} \bar{\mathbf{s}}_i\|} + \hat{\mathbf{v}} \right)}{\|\mathbf{G}^{-\top} \bar{\mathbf{b}}\| \left\| \frac{\mathbf{G} \bar{\mathbf{s}}_i}{\|\mathbf{G} \bar{\mathbf{s}}_i\|} + \hat{\mathbf{v}} \right\|} \right) \right]^2 \right)}{\frac{\hat{\mathbf{v}}^\top (\mathbf{G}^{-\top} \bar{\mathbf{b}})}{\|\mathbf{G}^{-\top} \bar{\mathbf{b}}\|}},$$

其中 $\hat{\mathbf{v}}$ 表示单位观测向量,

$$\mathbf{G}^{-\top} = (\mathbf{G}^\top)^{-1}, \quad \mathbf{G}^{-1} = \frac{1}{g_3} \begin{bmatrix} g_3 & 0 & 0 \\ 0 & g_3 & 0 \\ -g_1 & -g_2 & 1 \end{bmatrix}.$$

C. 表面法向量的可积性

表面法向量将 z 方向坐标归一化后, 法向量 \mathbf{x} 、 \mathbf{y} 方向坐标相反数即应为表面上当前位置深度的 \mathbf{x} 、 \mathbf{y} 方向偏导 z_x 与 z_y 。而对一个二阶可微函数, 应有

$$\frac{\partial}{\partial y} z_x = \frac{\partial}{\partial x} z_y$$

因此我们需要由先前估算的偏导 \hat{z}_x 、 \hat{z}_y (由估计的表面法向量转换而来) 生成一组可积的偏导 \tilde{z}_x 与 \tilde{z}_y , 使得两组偏导如下的距离最小

$$d\{(\hat{z}_x, \hat{z}_y), (\tilde{z}_x, \tilde{z}_y)\} = \iint |\hat{z}_x - \tilde{z}_x|^2 + |\hat{z}_y - \tilde{z}_y|^2 dx dy$$

我们采用文献 [9] 中所述的算法 (Frankot-Chellappa 算法) 最小化如上距离函数。其基本方法为, 考察深度 z 在一组正交基 $\phi(x, y, \omega)$ 上的展开, 通过将偏导 \hat{z}_x 、 \hat{z}_y 在正交基的偏导 $\phi_x(x, y, \omega) = \frac{\partial \phi}{\partial x}$ 与 $\phi_y(x, y, \omega) = \frac{\partial \phi}{\partial y}$ 上展开, 并将两组系数平均, 取逆变换得到深度 z 将使得上述距离最小。

具体地, z 、 z_x 、 z_y 展开式为

$$\begin{aligned} z(x, y) &= \sum_{\omega \in \Omega} C(\omega) \phi(x, y, \omega) \\ z_x(x, y) &= \sum_{\omega \in \Omega} C(\omega) \phi_x(x, y, \omega) \\ z_y(x, y) &= \sum_{\omega \in \Omega} C(\omega) \phi_y(x, y, \omega) \end{aligned}$$

对于估算值

$$\begin{aligned} \hat{z}_x(x, y) &= \sum_{\omega \in \Omega} \hat{C}_1(\omega) \phi_x(x, y, \omega) \\ \hat{z}_y(x, y) &= \sum_{\omega \in \Omega} \hat{C}_2(\omega) \phi_y(x, y, \omega) \end{aligned}$$

有展开系数 \hat{C}_1 、 \hat{C}_2 的均值 \hat{C} 是使得距离 $d\{(\hat{z}_x, \hat{z}_y), (\tilde{z}_x, \tilde{z}_y)\}$ 的 \tilde{z} 的展开系数。

$$\hat{C}(\omega) = \frac{P_x(\omega) \hat{C}_1(\omega) + P_y(\omega) \hat{C}_2(\omega)}{P_x(\omega) + P_y(\omega)}$$

此处

$$\begin{aligned} P_x(\omega) &= \iint |\hat{C}_1(\omega)|^2 dx dy \\ P_y(\omega) &= \iint |\hat{C}_2(\omega)|^2 dx dy \end{aligned}$$

为归一化系数。

具体地, 正交基的选取可以使用 DCT (离散余弦变换) 或 DFT (离散傅里叶变换), 同时由于取值离散, 因此偏导都需要用差分替代。以傅里叶变换为例:

$$\phi(x, y, \omega) = \exp(j(\omega_x x + \omega_y y))$$

有 $\phi_x = j\omega_x \phi$, $\phi_y = j\omega_y \phi$, 可得到估计:

$$\hat{C}(\omega) = \frac{-j\omega_x \hat{C}_x(\omega) - j\omega_y \hat{C}_y(\omega)}{\omega_x^2 + \omega_y^2}$$

取逆变换即得到 z 的估计, 完成表面法向量到深度的转换。

II. 实现方法

代码的实现主要由以下部分组成:

- 1) 读入训练图像和光线照射角度
- 2) 设定阈值去除阴影和高光部分
- 3) 训练 T-S 模型参数
- 4) 通过插值恢复阴影和高光的反射系数与法向量
- 5) 通过表面法向量生成深度图
- 6) 通过深度图与 T-S 模型参数, 根据测试图像光线照射角度渲染人脸图像

以下详细介绍 2-6 步的实现方法 (注意实际代码中步骤 3-5 是耦合的)。

A. 设定阈值去除阴影和高光部分

根据文献 [1], 高光部分指的是饱和光强, 即像素值达到最大的 255, 而阴影部分通常需要设定阈值进行衡量。阈值过大将使得去除的非实际阴影点过多, 难以通过插值恢复。而阈值过小将使得一些阴影未能成功去除。我们实验认为阈值设定为 5 能较好渲染。同时, 对于每一个像素, 若训练集中有至少一张图该像素为阴影或高光, 则整个像素不再考虑。具体实现中, 我们直接将拼接的训练图像矩阵中阴影或高光像素置为 NaN, 可在之后的最小二乘中不再考虑该像素, 相当于将整个像素位置所有七张图像忽略。同时需要注意, 之后最小化范数距离时, 需要去除阴影和高光部分, 因此实际实现中范数乘以一个由 `~isnan` 生成的掩膜。

B. 训练 T-S 模型参数

T-S 模型需要训练的参数除 Lambertian 反射模型中具有的各点表面法向量 $(z_x(x, y), z_y(x, y), -1)$ 以及

漫反射系数 $\alpha(x, y)$ 外, 还需训练镜面反射系数 a_s 、表面粗糙程度 ν 、GBR 矩阵 \mathbf{G} 。为模型简便起见, 我们假定 a_s 与 ν 在整个图像中不变。

对于基础的 Lambertian 反射模型, 只需根据其渲染公式, 通过最小二乘解出 $\mathbf{x} = \mathbf{B}\mathbf{s}$ 中的 \mathbf{B} 。这里不需要考虑图像阴影中将负数置 0 的 $\max(\cdot, 0)$, 而 $\mathbf{s} \in \mathbb{R}^{3 \times 7}$ 是 7 张训练集光线照射角度拼成的矩阵, $\mathbf{x} \in \mathbb{R}^{n \times 7}$ 是 7 张包含 n 个像素的训练图片拼成的矩阵, $\mathbf{B} \in \mathbb{R}^{n \times 3}$ 为每个点对应的行向量 \mathbf{b} 拼成的矩阵。然后可由每个点对应的 \mathbf{b} 的模长得到漫反射系数 $\alpha(x, y)$, 并将 z 坐标归一为 -1 得到表面法向量 $(z_x(x, y), z_y(x, y), -1)$ 。

注意这里 \mathbf{x} 中有像素为 NaN 不影响通过最小二乘法解出 \mathbf{B} , 只是 \mathbf{B} 中对应点的最小二乘解也为 NaN, 这将在之后通过插值弥补。在 Matlab 中, 左除号 “\” 自动包含了最小二乘的含义。

在由 Lambertian 反射模型改进至 T-S 模型时, 我们在每一个迭代周期中通过如下的方法进行训练:

- 1) 将 \mathbf{x} 扣除上一轮迭代中计算得到的镜面反射强度 \mathbf{I}_m (初始 \mathbf{I}_m 置 0) 得到 $\tilde{\mathbf{x}}$ 。
- 2) 将 $\tilde{\mathbf{x}}$ 代入方程, 通过已知的 \mathbf{s} 最小二乘解得 Lambertian 模型的参数 (表面法向量与漫反射系数)。
- 3) 保持表面法向量与漫反射系数不变, 通过拟牛顿法迭代计算 a_s 、 ν 与 \mathbf{G} , 使得非阴影、高光像素与训练图像的差范数最小。
- 4) 由如上求得的参数更新深度, 代入 T-S 模型渲染得到镜面反射强度 \mathbf{I}_m 。

这里拟牛顿法通过 `fminunc` 函数实现, T-S 模型的渲染实现见后。虽然 \mathbf{I}_m 与表面法向量也相关, 但在迭代中, 我们选择轮流优化两组参数。代码中, 我们进行十次这样的迭代周期, 基本可以收敛。具体实现详见代码。

C. 通过插值恢复阴影和高光的反射系数与法向量

实验中我们采用线性插值的方法来恢复阴影和高光的反射系数与法向量, 其中的难点在于寻找待恢复像素的所有邻接像素。我们主要利用了 `ind2sub` 函数将每个像素构造成一个 (序号, 行序号, 列序号) 的三元组方便其邻接像素的确定, 最后得到了所有待恢复像素邻接的已知像素的集合, 利用 `sparse` 函数实现利用邻接像素对待插像素的线性插值。

D. 通过表面法向量生成深度图

我们使用实验原理中所述的 Frankot-Chellappa 算法, 使用 FFT 作为具体实现, 将表面法向量转换为深度图。具体公式已在前述实验原理中给出, 这里需要注意的是, 从人的角度与相机的角度, x 方向呈镜面对称, 两坐标系有所区别, 因此需要将公式中的 $j\omega_x$ 前改为正号。具体实现详见代码, 主要使用 `fft2` 函数和 `ifft2` 实现。

E. 渲染人脸图像

这一部分使用先前实验原理中所述的 T-S 模型渲染公式, 由深度图和 T-S 模型参数获得渲染图象 (这一步骤在迭代中也存在)。

III. 结果分析与思考

A. 结果分析

三组图像在不同光照下的渲染重建结果如图 1 所示。发现在光照方向在正面时重建效果较好, 当光从人脸侧面照来时重建存在失真, 在鼻翼部分会出现高光。同时发现该模型的鲁棒性有待提升, 对于第二个人脸在光照不充足的部分出现了些许扭曲。

三组人脸的漫反射系数分别如图 2 所示。第一、三组人脸的鼻翼部分出现高光。这大致是二人鼻子较高, 鼻翼时常为阴影, 扣除阴影像素时没有扣除完全, 导致最小二乘及插值计算得到的漫反射系数过高。

三组人脸模型建模后深度图分别如图 3 所示。发现总体建模效果较理想, 部分位置 (如眼睛) 效果不佳, 这可能是因为眼睛较为接近镜面, 虹膜部分呈现深色而瞳孔部分存在反光导致。使用 T-S 模型后深度图局部细节略微变得精细, 但总体效果区别不大。

B. 改进措施

总体来看, 重建的深度图较为真实, 问题主要集中在重新渲染的图象。渲染结果中, 大角度有时不够真实, 鼻翼有时出现高光。这主要是阴影部分的漫反射系数训练不够好, 模型本身不够精细。改进措施可能包括: 训练时设置更高的阴影阈值, 而后续使用更复杂的插值方法恢复阴影、高光区域的漫反射系数和表面法向量; 采用光线追踪的方法, 利用初步恢复的深度图计算鼻子遮挡的鼻翼区域, 更精确地扣除阴影部分; 更精细地调整 T-S 模型中地镜面反射系数和表面粗糙度 (这可能因为 T-S 模型表达式二次可微, 可使用牛顿迭代法求解)。

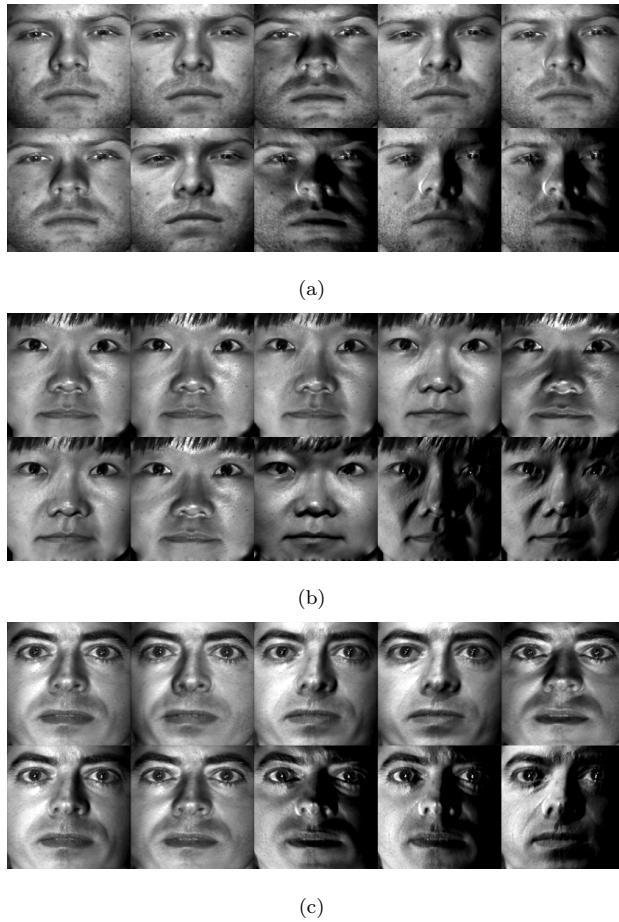


图 1. (a)-(c) 分别为 P1-P3 数据集中的人脸重建渲染结果, 在光从大角度照来时会失真。

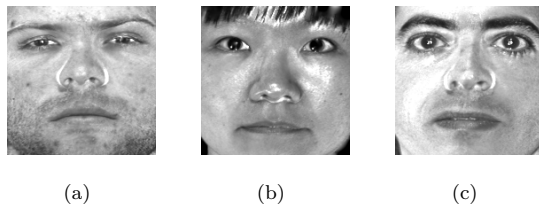


图 2. 三组人脸的漫反射系数。

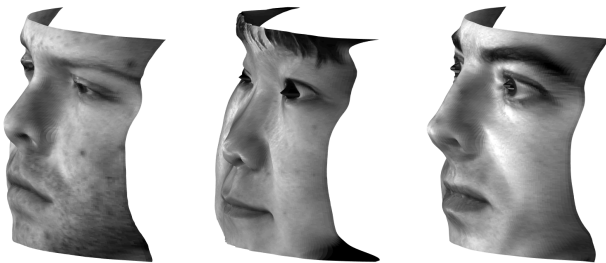


图 3. 三组人脸的深度图重建结果。

IV. 总结

我们考虑了使用 Lambertian 反射模型实现人脸的三维重建和渲染, 进一步我们使用了更一般的 Torrance and Sparrow 反射模型解决 Lambertian 反射模型中固有的参数模糊。实现过程中我们根据渲染效果完善了参数训练、插值恢复等细节, 完成了从表面法向量生成深度图等任务。实验结果表明我们采用的模型和实现方法取得了较好的重建与渲染效果, 同时我们还给出了这个问题中可以进一步改善的方向。

V. 文件清单

dataset_offline	原数据集 (含渲染图像与深度信息)
angleConversion.m	光照角度转换
demo.m	主脚本
Inpaint.m	插值函数
NeighborsIdentification.m	邻点判定函数 (插值函数辅助)
normalToDepth.m	表面法向量转换深度函数
partial.m	偏导函数
rendering.m	渲染函数
TSmodel.m	T-S 模型渲染公式

参考文献

- [1] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 6, pp. 643–660, 2001.
- [2] J. H. Lambert, "Photometria sive de mensura et gradibus luminis," *Colorum et Umbrae*, Eberhard Klett, Augsburg, Germany, 1760.
- [3] P. W. Hallinan *et al.*, "A low-dimensional representation of human faces for arbitrary lighting conditions," in *CVPR*, vol. 94, 1994, pp. 995–999.
- [4] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible illumination conditions?" *International Journal of Computer Vision*, vol. 28, no. 3, pp. 245–260, 1998.
- [5] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *JOSA A*, vol. 11, no. 11, pp. 3079–3089, 1994.
- [6] A. S. Georghiades, "Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo," in *null*. Ieee, 2003, p. 816.
- [7] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," *Josa*, vol. 57, no. 9, pp. 1105–1114, 1967.
- [8] J. R. Howell, M. P. Menguc, and R. Siegel, *Thermal radiation heat transfer*. CRC press, 2010.
- [9] R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 10, no. 4, pp. 439–451, 1988.