

指导教师：         杨涛        

提交时间：         2016/3/13        

# CVPR2015 Paper

## Translation

No:         01        

姓名：         李珮琦        

学号：         2013302490        

班号：         10011302        



# 多反射率对象的光度细化深度地图

Avishek Chatterjee Venu Madhav Govindu

印度科学理工学院

Bengaluru 560012, India

[avishek|venu@ee.iisc.ernet.in](mailto:avishek|venu@ee.iisc.ernet.in)

## 摘要

在本文中,我们提出一种新颖的无标定从消费者深度相机像微软 XBOX360 体感周边外设获得光度的方法来改善多反射率对象深度地图。现有未校准的光度方法假设物体的反射率恒定或依赖于在反射率恒定的地区分割图像。本文的这种方法不需要恒定的反射率的假设,我们相信它是此类第一个可以处理未校准的照明下具有任意反射率的物体。我们第一次使用迭代重新调整的方法全力估测观察亮度矩阵的 3 级近似值。随后,我们因式分解等级降低亮度矩阵成相应的照明、反射率和表面正常组件。被提议的因式分解是收敛的。我们用实验证明我们方法的价值通过提供高度精确的对各种各样对象的三维重建。此外,因为每种测光方法都需要使用一个辐射校准相机,我们也提出了一个采用辐射校准技术的结构光立体深度扫描仪的红外线相机。与现有方法不同,这种校准技术并不依赖于一个已知的校准对象或在使用照明的场景的道具。

## 1. 介绍

来自结构光立体深度照相机像微软 XBOX360 体感周边外设的深度地图遭受非常高的噪声和低分辨率。因此,一个具有精细(高频)结构的细节的对象是不会被这样的深度照相机捕获的。这种必备的高频细节可以通过使用辐射测量的方法获得,比如由明暗重构三维形状[14, 25, 7] 或光度立体[21], 可以与低质量的深度地图融合获得高质量的形状估计。由于关于明暗重构三维形状和光度立体的文献是巨大的,我们约束在下面讨论的相关文献是关于深度融合和法线[23, 9, 18, 10, 5]辐射

估测的特定的上下文。这些方法大致可以分为两类:明暗重构三维形状方法只使用一个强度图像[23, 9, 5]和光度立体方法依赖于在不同照明条件[18, 10]下的观察。明暗重构三维形状可以更简单和快速地获得一个单一的图像,同时也遭受着内在模棱两可的情况包括多个解决方案即使对于已知的照明[3, 8]和通常局限于常数反照率的对象。在表面法线只使用单一图像的情况下区分反射率的变化往往是很困难的。当这种方法在[23, 5]部分观察图像分割成持续的反照率区域,这是很严格的,因为许多对象有着不断变化的反照率或小区用不同的反照率,这使分割任务困难或不可能。

通过使用多个图像获得不同照明条件下,光度立体能够克服明暗重构三维形状所遭受的模棱两可。在本文中,我们提出一个新颖的光度方法来恢复多反射率对象的表面法线,然后和从微软 XBOX360 体感周边外设得到的深度估计融合获得深度细化方法,准确地恢复形状的 3d 对象。在我们的工作中,我们使用微软 XBOX360 体感周边外设的红外(今后红外)相机为辐射测量的方法在[5, 10]。同时在[18, 10]也使用光度立体的原则来正常估计,他们的方法使用限制性假设即对象有一个单一的反照率。当[10]的正常估计步骤是独立的反照率,其照明评估步骤假设持续的反照率。因此,对于正常估计的多反射率对象和最后的细化深度估计是很难的。[18]使用在在[11]中被建议的未校准的光度立体视觉方法。在解决线性模糊观测辐射亮度的分解,[18]使用一个常数反照率的假设。相反,在我们的方法中详细说明 3.2 节我们没有这样严格的假设。当解决这种明亮矩阵分解的线性模糊,而不是假设不变的反照率,我们最小化的一个小

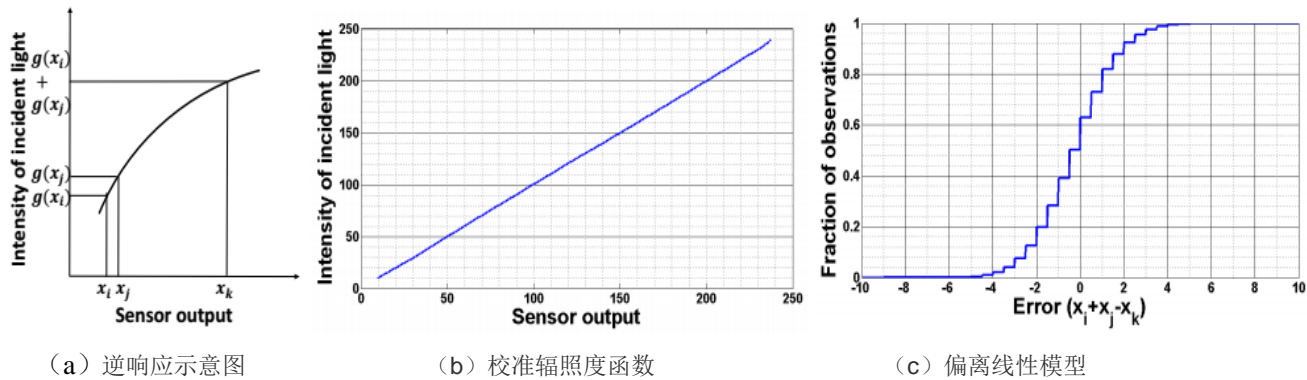


图 1, 红外相机的辐射校准。详情见正文。

说的成本是一个来自未加工的微软 XBOX360 体感周边外设深度地图和用光度立体获得的法线估计的一个函数。这是迭代进行最小化。我们在 3.5 节, 我们的分析证明我们的迭代法在噪声自由的情况是收敛的和证明在一个宽范的条件下的真实数据是收敛的。最后, 第四节中提到, 我们使用深度和正常的估计相结合的融合技术来创建高质量的三维形状表示对象。

## 2. 微软 XBOX360 体感周边外设的红外摄像机的辐射校准

红外相机的辐射校准是一个重要的前提对于任何依赖辐射测量方法的光度估计。在[5], 辐射校准是通过比较预测和观察到的一个传感器的球体在照明控制下的亮度。这个方法要求有一个完美的具有已知几何学的传感器表面用于校准照明源。要克服这样的方法的局限性, 我们提出一个辐射校准方法, 这个方法既不依赖于观察到表面的反射率的知识也不依赖于照明源的知识。在功能示意图如图 1(a) 所示, 横轴代表了测量亮度 (即红外相机测量的像素强度), 纵轴代表了红外相机实际接收的光量, 即  $g(\cdot)$  是相机的逆响应函数, 详情见条目“辐射校准”在 [15]。因此, 恢复  $g(\cdot)$  意味着对任何测量像素强度  $x$  我们可以直接恢复真正的辐照度  $g(x)$ 。

在校准的练习中, 我们寻求恢复  $g(\cdot)$

作为一种非参数函数。让测量亮度或像素强度落在范围  $[\chi_1, \dots, \chi_n]$ , 通常  $[0 \dots 255]$  在步骤 1 中。因此通过辐射校准我们寻求恢复未知值  $\{g(\chi_1), \dots, g(\chi_n)\}$ 。我们用两个光源照亮一个任意的场景。考虑一个像素与观察到的强度等于  $\chi_i$  和  $\chi_j$  分别为第一和第二

光源等于  $\chi_k$  为光源同时打开。因此, 当两个灯都打开的辐照度是单个光源的辐照度之和时我们有

$$g(x_k) = g(x_i) + g(x_j) \quad (1)$$

因此, 方程式 1 代表在  $n$  个中未知的非参数函数  $g(x)$  中三个未知数之间的线性关系。我们通过有着实质性反射率变化的成像场景产生足够的这类约束和解决由此产生的最小二乘意义上的齐次方程组。虽然有 [6] 等基于不同的曝光时间的辐射校准方法, 我们选择使用上述方法因为控制微软 XBOX360 体感周边外设的红外相机的曝光时间是很难的。在我们的实验中, 为了减少传感器噪声的影响, 图像平均超过 10 个连续观察在相同照度下和接近饱和度范围的被丢弃。在图 1(b) 我们看到, 微软 XBOX360 体感周边外设的红外相机的辐射校准是线性的, 这意味着如方程 1,  $\chi_k = \chi_i + \chi_j$  也应该保留。在

图 1(c) 中我们描述观测与错误的分数

$\chi_i + \chi_j - \chi_k$  和注意, 大部分的观测值在 256 值存在  $\pm 4$  的动态范围的偏差。

### 3. 照明、正常和反照率评估

在我们的方法中, 我们假设对象是传感器<sup>1</sup>。我们还假设红外光源在本质上是定向的, 这是通过即将被扫描的对象的来源远远不够来确保的。让照明强度和方向对应第  $i$  个图像, 用  $\lambda_i$  和  $3 \times 1$  单位标准向量  $\mathbf{I}_i$  分别表示。让在像素  $p$  观察到的点的反照率和曲面法线用  $\alpha_p$  和  $3 \times 1$  单位标准向量  $\mathbf{n}_p$  分别表示。对于传感器的表面, 在第  $i$  个图像中像素  $p$  的亮度 ( $B_{ip}$ ) 是由

$$B_{ip} = \lambda_i \mathbf{I}_i^T \mathbf{n}_p \alpha_p = L_i^T N_p \quad (2)$$

给出的, 其中  $L_i = \lambda_i \mathbf{I}_i$  是按比例缩小的光照强度, 照明方向对应于图像  $i$ ,  $N_p = \alpha_p \mathbf{n}_p$  是在像素  $p$  [21, 14] 观察点的法向量按比例缩小的反照率。[1]发现这亮度不足的线性模型是由于在室内场景的光从不同方向反射, 提出了一种用于[23]的基于二阶球函数的亮度模型。然而, 在[18, 10, 5]找到了中用于光度深度引导足够精确的亮度线性模型。

如果我们有  $k$  照明条件和  $m$  观察到的像素, 然后定义  $\mathbf{L}(3 \times k) = [\mathbf{L}_1 \cdots \mathbf{L}_k]$  和  $\mathbf{N}(3 \times m) = [\mathbf{N}_1 \cdots \mathbf{N}_m]$ 。我们还定义了一个  $k \times m$  亮度矩阵  $\mathbb{B}$ ,  $\mathbb{B}(i, p)$  是在图像  $i$

中像素  $p$  的亮度。因此在所有图像中的所有的像素的亮度约束可以概括为

$$\mathbb{B} = \mathbf{L}^T \mathbf{N} \quad (3)$$

我们的问题是因式分解一个观察亮度矩阵  $\mathbb{B}$  为  $\mathbf{L}$  和  $\mathbf{N}$ 。 $\mathbb{B}$  的等级不应超过 3 当  $\mathbf{L}$  和  $\mathbf{N}$  都不超过三维的时候。但在图像中的噪声, 异常值如阴影, 高光,  $i$  物体间多次反射等等导致  $\mathbb{B}$  等级超过 3。在[11], 在最小二乘意义上使用奇异值分解  $\mathbb{B}$  接近于三维矩阵。但由于最小平方逼近对异常值不是有效的, 我们提出一个有效的队列还原法。

#### 3.1. 等级 3 亮度矩阵的近似

观察到的亮度矩阵  $\mathbb{B}$  可以建模为一个等级 3 矩阵  $\mathbb{B}^*$ 、噪声矩阵  $e$  和异常值稀疏矩阵  $E$  的组合, 即

$$\mathbb{B} = \mathbb{B}^* + e + E \quad (4)$$

[22]忽略了噪声项  $e$  和核范数  $\mathbb{B}^*$  的最小化加权和和使用增广拉格朗日方法 (ALM) [16]的  $E$  的  $\ell_1$  标准。另外, 这个问题可以使用高戴克[26]解决, 它可将噪声分量考虑为因素。在我们的实验中, 我们发现通过异步线路模组或高戴克生成的近似矩阵通常等级小于 3。这种情况的发生是因为在许多实际情况下,  $\mathbf{N}$  在  $\mathbb{R}^3$  相等地不跨越各个方向。通常情况下, 在一个表面观察到的大多数的点有自己的正位点指向相机。同样,  $\mathbf{L}$  也无法很好地跨越  $\mathbb{R}^3$  因为光源通常放置在接近  $Z$  轴, 以避免由于自我遮挡产生的阴影。因此, 当从  $\mathbb{B}$  消除异常值时, 尽管他们包含高价位的计算, 异步线路模组或高戴克经常删除有用的信息和输出等级小于三的矩阵, 当它不能用于进一步的处理的时候。我们现在描述我们的方法, 缓解了这一问题, 并提供一个健壮的等级 3  $\mathbb{B}^*$  估计。我们现在描述我们缓解这一问题和提供一个完善的

<sup>1</sup>虽然实际传感器的表面是罕见的, 我们在实验中显示我们的方法能够很好地用在许多也有大量的镜面反射的非金属材料。

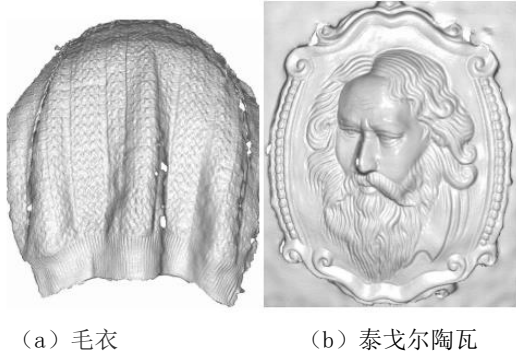


图 2, 我们估计的一个常数反照率物体的三维深度重建。

等级 3 估算  $\mathbb{B}^*$  的方法。

给定一个观测矩阵  $\mathbb{B}(k \times m)$  和  $m \geq k$ , 我们的任务是找到三维标准正交基  $V \in \mathbb{R}^{k \times 3}$  使  $\mathbb{B}$  设置的列间距大部分位于由  $V$  张成的子空间。让  $\mathbb{B}_p$  为  $\mathbb{B}$  的第  $p$  列。我们定义误差  $e_p$  为  $\mathbb{B}_p$  的组件的大小,  $\mathbb{B}_p$  为由  $V$  张成的子空间的正交直线也就是

$$e_p = \|\mathbb{B}_p - VV^T \mathbb{B}_p\| \quad (5)$$

然后我们最小化  $\sum_{p=1}^m \rho(e_p)$  找到最佳的  $V$ 。

如果我们在我们的成本函数中选择  $\rho(e) = e^2$ , 这个解决方案相当于对异常值不完善的最小二乘法。如果我们使用绝对规范  $\rho(e) = |e|$  来代替, 那么  $\rho(e)$  在  $e = 0$  处是不可微的。因此, 我们使用胡伯成本函数, 定义为:

$$\rho(e) = \begin{cases} e^2 & \forall |e| \leq \alpha \\ 2\alpha|e| - \alpha^2 & \forall |e| > \alpha \end{cases} \quad (6)$$

这确保  $\rho(e)$  是  $C^1$  连续的也具有鲁棒性和凸面的。优化这个函数, 我们使用迭代重加权最小二乘 (IRLS) 方法 [12, 19] 可得到一个有效的, 更重要的是对异常值具有鲁棒性的简单的方法。这个有效的函数的梯度是由

$$\begin{aligned} \nabla(\rho(e)) &= \begin{cases} \nabla(e^2) & \forall |e| \leq \alpha \\ \nabla(2\alpha\sqrt{e^2} - \alpha^2) & \forall |e| > \alpha \end{cases} \\ &= \begin{cases} \nabla(e^2) & \forall |e| \leq \alpha \\ \frac{\alpha}{\sqrt{e^2}} \nabla(e^2) & \forall |e| > \alpha \end{cases} \\ &= \omega^2 \nabla(e^2) \begin{cases} \omega = 1 & \forall |e| \leq \alpha \\ \omega = \frac{\alpha}{|e|} & \forall |e| > \alpha \end{cases} \end{aligned} \quad (7)$$

给出的。

我们可以相同地把我们的成本函数的梯度为零的最优解为

$$\nabla\left(\sum_{p=1}^m \rho(e_p)\right) = \sum_{p=1}^m \omega_p^2 \nabla(e_p^2) = 0 \quad (8)$$

对于一个给定的  $\omega$ , 解方程 8 相当于最小

化  $\sum_{p=1}^m (\omega_p e_p)^2$  即一个加权最小二乘问题。

我们迭代计算, 直到收敛, 权重  $\omega_p$  基于误差  $e_p \forall p$  然后保持  $\omega_p$  不变, 我们解决了加权最小二乘问题——在本例中是主成分分析的问题——如算法 1 所示。

**算法 1** 等级 3  $\mathbb{B}$  的近似值使用鲁棒性主成分分析法。

输入:  $\mathbb{B}$ : 观察到的亮度矩阵 ( $k \times m$ )

输出:  $\mathbb{B}^*(k \times m)$ ,  $V(k \times 3)$ ,  $\mathbb{X}(3 \times m)$ ,

在  $\mathbb{B}^* = V\mathbb{X}$  处即为等级 3  $\mathbb{B}$  的近似。

初始设定:  $\omega_p = 1, \forall p = 1, 2, \dots, m$

1.  $V = \mathbb{B}W$  主要的 3 个主成分, 其中  $W$  是一个对角线加权矩阵,  $W(p, p) = \omega_p$
2.  $X = V^T \mathbb{B}$  和  $\mathbb{B}^* = VX$
4.  $e_p =$  队列  $(\mathbb{B} - \mathbb{B}^*) \forall p = 1, 2, \dots, m$  的第  $p$  列
5. 使用方程 7  $\forall p = 1, 2, \dots, m$  计算  $\omega_p$
6. 如果不是收敛的, 转到状态 1

我们注意到这里亮度矩阵  $\mathbb{B}$  的所有的列, 在算法 1 中作为输入使用之前要首先规范化。这将确保  $e_p$  的值有下界 2 即既确保鲁棒性也防止高反照率像素导致的误差, 从而控制解决方案。

### 3.2. 亮度矩阵的分解

从算法 1 我们有

$$\mathbb{B} \approx \mathbb{B}^* = VX \quad (9)$$

但这分解不是唯一的一个常规的  $\mathbb{B}^*$  的因式分解可以写成

$$\mathbb{B}^* = \mathbb{L}^T N = VAA^{-1}X \quad (10)$$

$A$  可以是任何是非奇异  $3 \times 3$  矩阵。因此, 存在一个  $A$  这样的

$$\mathbb{L} = A^T V^T \quad (11)$$

$$N = A^{-1}X \quad (12)$$

因此, 我们的问题是估计必备的让我们恢复  $N$  的矩阵  $A$ 。在 [11], 利用方程 12 解决  $A$ , 问用户选择相同反照率的 6 像素还是相同照明强度的 6 图像。相比之下, 我们解决

$A$  用来自微软 XBOX360 体感周边外设的原始深度图作为参考。我们生成一个粗略的估计的表面法线  $\hat{n}_p$  在来自原始深度地图的每个像素和收集在一个  $3 \times m$  矩阵, 如  $\hat{N} = [\hat{n}_1 \dots \hat{n}_m]$ 。

当曲面法线 ( $N$ ) 是  $X$  的线性变换, (见方程 12),  $X$  的所有线性变换是无效的曲面法线。这是因为要求任何有效的表面法线应可积或旋度自由 [13]。然而, 即使这种约束的实施也不会产生一种独特的线性变换  $A$ 。相反, 对于每一个给出一个可积的解决方案的  $A$ , 有  $A$  的变化的三个参数族也满足可积性的约束。这是著名的浅浮雕模糊 [2, 4]。

在本文中, 我们忽略了可积性约束在解决  $A$  的时候。我们注意到因为  $A^{-1}X$  应该与微软 XBOX360 体感周边外设的深度图的法线估计对齐, 我们有一个足够数量的约束来解决  $A$ , 尽管这种  $A$  不一定满足可积性约束。然而, 由于在我们的方法解出的表面是微软 XBOX360 体感周边外设的深度图估计以及正常的地图进行融合的, 我们不需要在法线的估计施加可积性约束。

在我们发展我们的多反射率因式分解解决方案之前, 我们在 [18, 10] 举例说明了用  $\hat{N}$  来解决常数反照率的简单问题。

### 3.3. 因式分解常数反照率表面

为一个表面反照率不变, 我们可以分配相对反照率  $\alpha_p = 1, \forall p$ 。这让我们在方程式 12 中用微软 XBOX360 体感周边外设估算的法矩阵  $\hat{N}$  来代替  $N$ 。因此, 我们需要使用方程 12 解出  $A$ 。在 [18]  $A$  是在最小二乘思想上解决的, 如

$$A = ((\hat{N}^T)^{\dagger} \hat{X}^T)^T \quad (13)$$

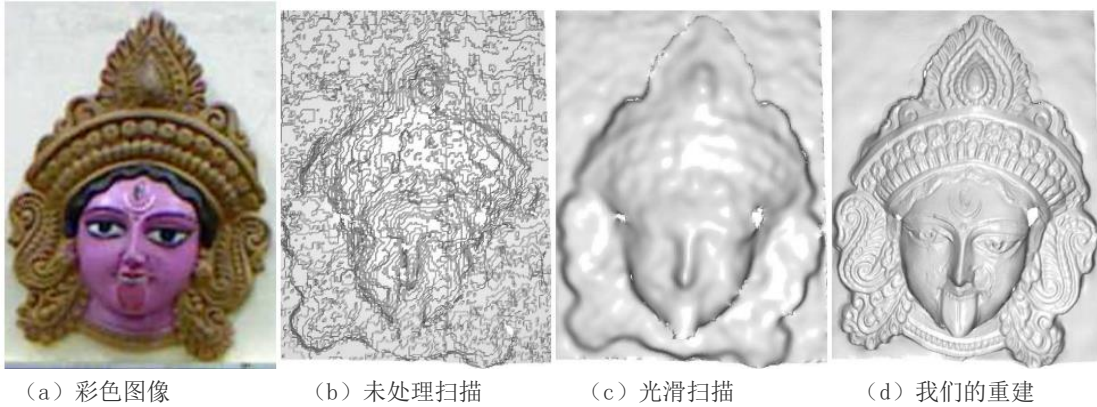


图 3, 使用我们的方法精制卡莉着色陶匾深度地图。

† 表示伪逆运算。但是接近高曲率区域包括一个物体的深度边缘的  $\hat{n}_p$  的估算是不可靠的, 因为微软 XBOX360 体感周边外设深度地图非常嘈杂。因此, 最小二乘解是受糟糕的或不正确的  $\hat{n}_p$  深度边缘附近地估算的影。为了鲁棒性地解决  $A$ , 我们定义  $e_p =$

$\| (A\hat{N} - \mathbb{X})$  的第  $p$  列  $\|$  和最小化  $\sum_{p=1}^m \rho(e_p)$ , 其中  $\rho(\cdot)$  定义在方程式 6 中。

这最小化是使用迭代重加权最小二乘法执行的, 详情见算法 2。

**算法 2** 使用迭代加权最小二乘鲁棒性成本最小化

初始设定:  $\omega_p = 1 \quad \forall p$

1.  $A = \left( (W\hat{N}^T)^\dagger (W\mathbb{X}^T) \right)^T$ , 其中  $W$  是一个对角线加权矩阵,  $W(p, p) = \omega_p$

2. 计算误差  $e_p$  和权重  $\omega_p$  对所有像素  $p$

### 3. 如果不是聚合, 转到状态 1

在图 2 中, 我们展示我们的估算三维重构 (a) 一件羊毛衫和 (b) 描绘印度诗人泰戈尔的陶匾。可以看到, 我们的方法非常准确地恢复高频形状细节。

### 3.4. 因式分解多反照率表面

我们现在可以继续解决更一般的问题恢复  $A$ , 具有多个或任意不同的反照率的对象。让  $x_p$  表示规范化的  $\mathbb{X}$  的第  $p$  列, 也就是对应于第  $p$  个像素。那么我们有

$$A\alpha_p \hat{n}_p \parallel x_p \Rightarrow A\hat{n}_p \parallel x_p \quad \forall p \quad (14)$$

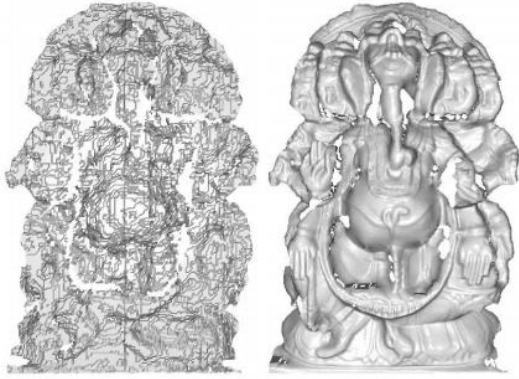
其中  $\alpha_p$  是第  $p$  个像素的反照率, 符号  $\parallel$  表示对所有像素, 矢量  $A\alpha_p \hat{n}_p$  和  $x_p$  是平行的。我们定义误差  $e_p$  为向量  $A\hat{n}_p$  和  $x_p$  之间的张角式测量, 如

$$e_p = \left\| x_p - \frac{A\hat{n}_p}{\|A\hat{n}_p\|} \right\| \quad (15)$$

$A$  的最优值现在可以鲁棒性地估算为

$$A^* = \arg \min_A \sum_{p=1}^m \rho(e_p) \quad (16)$$

方程 16 的优化问题是使用梯度下降法



(a) 未处理的扫描 (b) 我们的重建  
图 4, 使用我们的方法精制象鼻神深度地图估算。

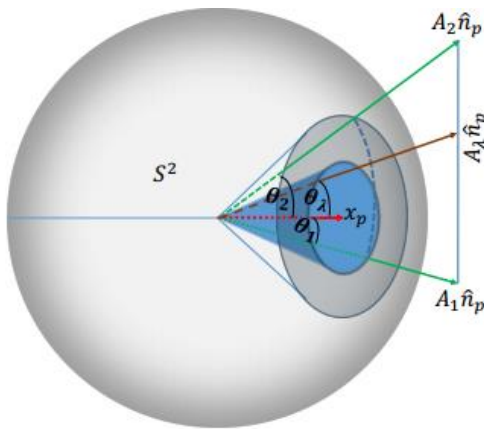


图 5, 我们的多反射率分解方法的收敛性。有关详细信息, 请参阅文本。

来解决。一旦我们找到  $A$  的最优值, 每个像素的曲面法线都可以恢复。

图 3 和图 4 中我们展示了使用这种方法获得的深度估计, 两个多反射率物体: 一个画有印度教女神卡莉的陶匾 (图 3) 和一个有着长象首的印度预言神图像的塑料 (图 4)。这两种模式都是镜面反射的。如图 3 (b) 和 4 (a), 初始的微软 XBOX360 体感周边外设深度地图有着很高的噪音, 使用我们的方法获得的合成改进 (图 3 (d) 和 4 (b)) 是相当可观的。

### 3.5. 多反射率因式分解的收敛性

在本节, 我们提出一个 3.4 节梯度下降

法收敛的证明, 我们用此来解在多反射率场景的  $A$ 。具体来说, 我们证明方程式 16 中的成本函数在无噪声场景是单调非递减远离全局最小点。因此, 只要观察到的亮度矩阵的讹误不是太大, 我们的梯度下降法保证收敛到最优解。

**引理 3.1** 让  $A_\lambda$  为  $A$  的两个可能值的凸组

合, 分别用  $A_1$  和  $A_2$  表示, 即

$$A_\lambda = (1-\lambda) A_1 + \lambda A_2, 0 \leq \lambda \leq 1. \text{ 然后,}$$

$$\rho(e_p(A_\lambda)) \leq \max(\rho(e_p(A_1)), \rho(e_p(A_2)))$$

(17)

在图 5 中向量  $x_p, A_1 \hat{n}_p, A_2 \hat{n}_p$  和  $A_\lambda \hat{n}_p$  都来自单位球  $S^2$  的中心。此外, 我们也显示  $A_1 \hat{n}_p$  和  $A_2 \hat{n}_p$  在  $x_p$  附近的回转锥面。假设,  $A_1 \hat{n}_p, A_2 \hat{n}_p$  和  $A_\lambda \hat{n}_p$  相对  $x_p$  角度分别为  $\theta_1, \theta_2$  和  $\theta_\lambda$ 。不失一般性地, 让  $\theta_2 > \theta_1$ 。

**证明** 当  $\lambda$  从 0 增加到 1,  $A_\lambda \hat{n}_p$  点在  $\mathbb{R}^3$  直线从  $A_1 \hat{n}_p$  移动到  $A_2 \hat{n}_p$ 。相应地,  $A_\lambda \hat{n}_p$  在

$S^2$  上的投影即  $\frac{A_\lambda \hat{n}_p}{\|A_\lambda \hat{n}_p\|}$  沿着在  $S^2$  上测量的

圆弧移动, 如蓝虚线所示。因为  $A_1 \hat{n}_p$  和  $A_2 \hat{n}_p$  都在由  $A_2 \hat{n}_p$  围绕  $x_p$  运行产生的凸集以内,

任何  $A_1 \hat{n}_p$  和  $A_2 \hat{n}_p$  的凸组合即

$$(1-\lambda) A_1 \hat{n}_p + \lambda A_2 \hat{n}_p \text{ 必须在 } A_2 \hat{n}_p \text{ 围绕 } x_p$$





图 6, 一些复杂的多反射率对象的重建。左: 红外图像, 中间: 我们的三维重建, 右: 我们估计的相对反射率。

$\forall \lambda \in [0, 1]$  的回转锥面以内。因此,  
 $\theta_\lambda \leq \theta_2 = \max(\theta_1, \theta_2)$ 。既然  $\rho(e_p(A_\lambda))$  是  
 一个  $\theta_\lambda$  单调递增的函数, 我们可以知道  
 $\rho(e_p(A_\lambda)) \leq \rho(e_p(A_2)) \leq$   
 $\max(\rho(e_p(A_1)), \rho(e_p(A_2)))$ 。反过来这意

味着方程 17 总是适用。

**定理 3.2**  $\sum_{p=1}^m \rho(e_p)$  在无噪音情况下是单调

非递减远离全局最小值的。

**证明** 对所有像素, 我们得到求和公式 17,  

$$\sum_p \rho(e_p(A_\lambda x_p)) \leq \sum_p \max(\rho(e_p(A_1)), \rho(e_p(A_2)))$$
 (18)

如果我们考虑无噪声的情况，如果  $A_1$  是真



(a) 木制的盘子

图 7, 重建的对象。左: 红外图像, 中间: 我们的三维重建, 右: 我们估计的相对反射率。

正的解决方案, 即  $\rho(e_p(A_1)) = 0 \forall p$ , 那么

$$\sum_p \rho(e_p(A_\lambda)) \leq \sum_p \rho(e_p(A_2)) \quad (19)$$

因此, 我们的成本是单调非递减从真正的解决方案引入  $A_1$ 。

我们注意到[20]使用类似的方法来解决全局转换估算, 在由移动构造的情况下。

#### 4. 深度和正常估算的融合

鉴于微软 XBOX360 体感周边外设的深度图和估计表面法线, 我们需要一个有原则的方法融合他们来获得一个高质量的成像表面的重建。在我们的工作中, 我们使用一个深度标准 1 融合法, 最初是在 [17] 提出和在 [24] 延伸至微软 XBOX360 体感周边外设的环境。这种方法也被用于 [9, 18, 10] 和读者可能涉及到的 [17, 24, 10] 用于细节, 因为由于空间的限制我们不能在这里提供他们。

#### 5. 结果

除了常数和多反照率物体的结果分别如图 2, 3 和 4 所示, 在这一部分中, 我们将额外展示复合多反照率物体的结果。在图 6 我们展示我们的结果, 一件羊毛夹克, 一双

物体	图 片 编号	材料	多 反 射 率	反 射 比
毛衣	2 (a)	毛线	×	传 感 器
泰 戈 尔	2 (b)	陶瓦	×	传 感 器
卡莉	3	着 色 陶瓦	√	镜 面 反 射
象 鼻 神	4 (b)	塑料	√	高 度 镜 面 反 射
夹克	6 (a)	毛线	√	传 感 器
短裤	6 (b)	棉线	√	传 感 器
鞋子	6 <sup>©</sup>	混 合 物	√	轻 度 镜 面 反 射
盘子	7	木头	√	中 度 镜 面 反 射

表 1, 我们的方法对有着不同反射率和其他材料属性的广泛的对象的适用性。在所有这些情况下我们的方法准确。

棉短裤和一双鞋子。从夹克的红外图像可以注意到反照率变化是和高度与深度边缘非常相关的, 如夹克的垂直深度条纹。然而, 我们可以准确地恢复三维结构和相对反照率。

例如, 条纹短裤(中间行), 我们可以观察到条纹反照率变化是完全恢复的, 这种不同的反照率模式的存在不是因为形状变化的曲解。此外, 弹性腰带短裤的高频细节也正确地恢复了。可以看到在图 6 的最后一行, 尽管这双鞋的反照率有着大量的变化幅度, 我们还是能够准确地恢复它们的三维形状, 包括精细缝合的细节。在图 7 中, 我们描绘了一个非常具有挑战性的对象的结果, 即展览一个木制雕刻板由于所使用的木头的纹理导致反照率的连续变化。在这种情况下, 任何依赖于分段反照率转化常数反照率区域的方法都会失败。在我们的例子中, 我们能够准确地恢复高度不同反射率的模式, 我们可以看到我们的三维模型重建一点也不受反照率变化的影响。我们还要注意重建工作, 尽管这个木雕板是适度镜面。最后, 在表 5 中, 我们总结了所有我们已经准确地重建和在本文中展示成果的对象属性。我们相信这些物体属性的广泛变化是我们三维重建方法性能的证据。

## 6. 结论

在本文中, 我们提出一个多反射率物体精炼深度估算的新颖的方法。我们的方法能够观察亮度矩阵因式分解成适当的照明, 反照率和曲面法线分量。我们的方法可以不需要任何预先校准照明配置或反照率的分割, 我们提出了一套更广泛的结果证明其作用于各种物体的准确性。

## 致谢

第一作者的工作由 TCS 研究奖学金支持。作者感谢匿名评论者的有用反馈。

## 参考文献

[1] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general unknown lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007.

[2] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The basrelief ambiguity. *International journal of computer vision*, 35(1):33–44, 1999.

[3] M. J. Brooks. Two results concerning ambiguity in shape from shading. In *AAAI*, pages 36–39, 1983.

[4] M. Chandraker. Bas-relief ambiguity. In *Computer Vision*, pages 43–46. Springer, 2014.

[5] G. Choe, J. Park, Y. Tai, and I. Kweon. Exploiting shading cues in kinect IR images for geometry refinement. In *Computer Vision and Pattern Recognition (CVPR)*, pages 3922–3929. IEEE, 2014.

[6] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '97*, pages 369–378, 1997.

[7] J.-D. Durou, M. Falcone, and M. Sagona. Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding*, 109(1):22–43, 2008.

[8] A. Ecker and A. D. Jepson. Polynomial shape from shading. In *Computer Vision and Pattern Recognition (CVPR)*, pages 145–152. IEEE, 2010.

[9] Y. Han, J. Lee, and I. Kweon. High quality shape from a single rgb-d image under uncalibrated natural illumination. In *International Conference on Computer Vision (ICCV)*, pages 1617–1624. IEEE, 2013.

[10] S. M. Haque, A. Chatterjee, and V. M. Govindu. High quality photometric reconstruction using a depth camera. In

- Computer Vision and Pattern Recognition (CVPR), pages 2283–2290. IEEE, 2014.
- [11] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *JOSA A*, 11(11):3079–3089, 1994.
- [12] P. W. Holland and R. E. Welsch. Robust regression using iteratively reweighted least-squares. *Commun. Statist. Theory Methods*, 9, 1977.
- [13] B. K. P. Horn. *Robot vision*. MIT press, 1986.
- [14] B. K. P. Horn. *Obtaining Shape from Shading Information*. MIT Press, 1989.
- [15] K. Ikeuchi. *Computer Vision: A Reference Guide*. Springer, 2014.
- [16] Z. Lin, M. Chen, and Y. Ma. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. arXiv preprint arXiv:1009.5055, 2010.
- [17] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently combining positions and normals for precise 3d geometry. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 536–543. ACM, 2005.
- [18] J. Park, S. N. Sinha, Y. Matsushita, Y. Tai, and I. Kweon. Multiview photometric stereo using planar mesh parameterization. In *International Conference on Computer Vision (ICCV)*, pages 1161–1168. IEEE, 2013.
- [19] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag, 2010.
- [20] K. Wilson and N. Snavely. Robust global translations with 1dsfm. In *Computer Vision–ECCV 2014*, pages 61–75. Springer, 2014.
- [21] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139–191139, 1980.
- [22] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Computer Vision–ACCV 2010*, pages 703–717. Springer, 2011.
- [23] L. Yu, S. K. Yeung, Y. Tai, and S. Lin. Shading-based shape refinement of RGB-D images. In *Computer Vision and Pattern Recognition*, pages 1415–1422. IEEE, 2013.
- [24] Q. Zhang, M. Ye, R. Yang, Y. Matsushita, B. Wilburn, and H. Yu. Edge-preserving photometric stereo via depth fusion. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2472–2479. IEEE, 2012.
- [25] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Shapefrom-shading: a survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(8):690–706, 1999.
- [26] T. Zhou and D. Tao. Godec: Randomized low-rank & sparse matrix decomposition in noisy case. In *International Conference on Machine Learning (ICML)*, pages 33–40, 2011.