一种基于高斯过程的数据驱动光度立体视觉全局优化算法

王晰, 任明俊*, 肖高博, 陈明汉, 付琳

上海交通大学机械与动力工程学院, 机械系统与振动国家重点实验室, 上海 200240

* E-mail: renmj@sjtu.edu.cn

收稿日期: 2018-02-11; 接受日期: 2018-06-13; 网络版发表日期: 2018-07-25

国家自然科学基金(批准号: 51505404, 51675456)、上海浦江计划(编号: 16PJ1404300)和深圳市科技计划(编号: GRCK20160415111859786)资助项目

摘要 光度立体视觉技术难点之一在于快速有效地解决非朗伯体材料的复杂反射问题, 针对这一问题, 本文提出了一种基于高斯过程的数据驱动光度立体视觉全局优化算法. 该方法利用高斯过程, 通过贝叶斯推理学习给定材料的实测双向反射分布函数值, 建立用于描述非朗伯体材料非线性反射性质的连续数学模型. 该模型以光源方向和表面反射光强作为输入, 以双向反射分布函数值作为输出, 可在光度立体视觉系统中建立更为精确的反射模型, 从而得到更高的法向量预测精度. 仿真实验和实际实验表明, 该方法显著提升了光度立体视觉的法向量恢复精度, 且拥有较高的计算效率.

关键词 数据驱动, 光度立体视觉, 高斯过程

1 引言

立体视觉是计算机视觉领域中一项重要的研究课题, 其主要方法可以分为两类. 第一类方法通过模拟人眼的工作原理, 使用两个或多个相机对同一场景进行拍摄, 采用相关算法进行像素匹配, 再利用成像系统的几何关系得到场景的深度信息. 第二类方法分析了像素记录值与拍摄物体表面法向量之间的关系, 并根据该关系, 由像素记录值求解物体表面法向量, 比如 20 世纪 70 年代 Horn^[1]提出来的从明暗恢复形状(shape from shading, SFS)方法.

光度立体视觉属于上述的第二类方法, 1980 年 Woodham^[2]最早提出了该方法. 与 SFS 方法不同的是,

光度立体视觉采用不同光照方向下同一场景的多幅图片求解物体表面法向量. 由于光度立体视觉系统简单廉价, 物体表面细节恢复精度高, 所以在缺陷检测、产品质量控制等领域有着广泛的应用, 比如钢板缺陷检测^[3]、地板瑕疵检测^[4]、织物褶皱等级评定^[5]、物体三维数字化^[6]等.

传统光度立体视觉基于朗伯体假设来估计物体表面法向量. 在该假设下, 像素记录值与物体表面法向量之间存在线性关系. 因此, 针对某一个像素, 已知 3 个不共面光照方向的像素记录值, 就可以求取该像素对应物体表面点的法向量. 但在现实世界中, 满足朗伯体假设的材料很少, 这就限制了光度立体视觉的应用空间. 为了解决该问题, 三十多年来发展出了许多算法,

引用格式: 王晰, 任明俊, 肖高博, 等. 一种基于高斯过程的数据驱动光度立体视觉全局优化算法. 中国科学: 技术科学, 2018, 48: 969–979
Wang X, Ren M J, Xiao G B, et al. A data-driven photometric stereo algorithm (in Chinese). Sci Sin Tech, 2018, 48: 969–979, doi: [10.1360/N092018-00035](https://doi.org/10.1360/N092018-00035)

其中主流的解决方法可以分成两种。

第一种方法认为绝大多数像素记录值满足朗伯体假设, 将高光点、投射阴影(cast shadow)、附着阴影(attach shadow)等作为误差处理。Coleman和Jain^[7]提出四光源光度立体视觉法, 通过新加入的光源判断并去除高光点。Soloman和Ikeuchi^[8]在此基础上详细讨论了4个光源的照射区域, 提高了算法的精度与稳定性。Barsky和Petrout^[9]综合了颜色信息, 进一步完善了四光源光度立体视觉法。Mukaigawa等人^[10]利用RANSAC(random sampling consensus)技术对原始图片进行处理, 得到一组满足朗伯体假设的新图片。Miyazaki等人^[11]列举了所有光源的三光源子集, 得到一组法向量集, 对该集取中值并综合领域信息得到最终法向量。中值点具有抗外点的性能, 因此隐含地去除了高光点和阴影点的影响。吴仑等人^[12]分析了像素记录值矩阵的低秩特征, 利用鲁棒的主成分分析技术(robust principal component analysis, RPCA), 分离出噪声矩阵, 从而获得较高的法向量恢复精度。Ikehata等人^[13]从统计概率的角度出发, 采用稀疏贝叶斯学习(sparse Bayesian learning, SBL)方法, 减小高光与阴影的影响。尽管上述方法均使法向量恢复精度得到了提高, 但仍然处理不了非朗伯体漫反射等复杂反射的材料。也就是说, 对于一些材料, 即使是去除了高光和阴影的影响, 像素记录值与对应法向量之间的关系也不满足朗伯体假设。

第二种方法采用更加复杂的双向反射分布函数(bidirectional reflectance distribution function, BRDF)模型, 或者利用BRDF的某些性质来提高法向量恢复精度。Goldman等人^[14]认为某种材料的BRDF可以通过几种基本材料的BRDF线性组合表达出来, 他们用Ward模型^[15]来拟合基本材料的BRDF, 用非线性优化方法求解法向量。Higo等人^[16]综合利用了BRDF的单调性、可见性和各向同性, 建立最优化方程, 并使用Levenberg-Marquardt算法求解法向量。Shi等人^[17]针对低频像素记录值信息, 提出一种双变量二次反射模型(Bi-quadratic), 通过迭代求得最优的法向量和反射模型系数。Ikehata和Aizawa^[18]利用BRDF半参数化模型^[19], 综合BRDF单调性、非负性、过原点3个性质, 提出了有约束的双变量回归方法(constrained bivariate regression, CBR), 建立了针对光度立体视觉的像素级逆向模型, 并用伯恩斯坦多项式对该模型进行拟合, 取

得了较好的精度。Han和Shen^[20]建立了符合漫反射、高光点与阴影点特性的数学模型, 将光度立体视觉问题转化为二次锥优化问题, 同时解决了各向同性材料与各向异性材料的法向量恢复问题。最近Shen等人^[21]使用径向基核函数来解决BRDF的高度非线性问题, 建立了岭回归方程, 推导出相应的解析解, 并提出矩阵求逆的优化算法, 实现径向基核函数超参数的自动选择, 进一步提高了法向量的恢复精度。但上述模型的适应性仍然有限, 法向量恢复精度仍旧有提高的空间。此外, 所提出的模型往往相对复杂, 或者最优化迭代求解过程繁琐, 计算速度慢。

本文提出了一种基于高斯过程的数据驱动光度立体视觉全局优化(data-driven photometric stereo, DPS)算法, 它可以更好地解决BRDF高度非线性问题, 而且拥有较高的计算效率。该方法通过高斯过程对某种材料的参照数据进行训练, 得到一个合理的全局推测模型, 对同一材料相同场景下的其他形状物体具有较高的法向量恢复精度。基于MERL(mitsubishi electric research laboratories)数据库^[22]的仿真实验以及真实实验都验证了该方法的优越性。该方法在大批量单一材料表面的缺陷检测方面有较好的应用前景。

2 基本原理

2.1 光度立体视觉基本原理

光度立体视觉可以认为是相机成像过程的逆过程, 因此需要建立相机成像过程的关系, 即某像素的像素记录值 I_p 与物体表面法向量 $\mathbf{n}$ 之间的关系, 如图1所示。推导过程中, 本文假设物体受到平行光照射, 相机线性响应并且拍摄距离远大于物体本身尺寸, 不考虑阴影点、互反射等因素的影响, 上述关系可以表示为

$$I_p = k_p \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \alpha F I_s \mathbf{n}^T \mathbf{l} T, \quad (1)$$

其中, k_p 是相机内部的线性响应系数, d 是光圈直径, f 是相机焦距, α 是物体表面点与对应像素点连线与光轴之间的夹角, 在拍摄距离远大于物体尺寸的假设下, $\alpha \approx 0^\circ$, F 是物体表面点向相机方向反射的辐射强度与物体表面辐照强度的比值, I_s 是光源强度, $\mathbf{l}$ 是入射光线方向, T 是曝光时间。

图 1 (网络版彩图)相机成像模型

对于同一场景, k_p , d , f 均为常数. 令 $k_c = k_p \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \alpha$, 此时式(1)简化为

$$I_p = F k_c I_s T \mathbf{n}^T \mathbf{l}. \quad (2)$$

传统光度立体视觉算法假设物体是朗伯体, 即物体表面任意一点向各个方向反射的辐射强度相同. 此时 $F = \frac{\rho}{\pi}$ 为一常数, 其中 ρ 是反射率. 当给定3张不同光照方向下的图片, 且这3个光照方向不共面时, 对于任意一个像素点, 每个光照方向都有式(2)成立, 由此可以列得拥有3个方程的线性方程组, 其中 I_p 为已知量, $\mathbf{l}$, $k_c I_s T$ 可以通过标定得到. 由此, 求解该线性方程组可以得到 $F \mathbf{n}$, 最后归一化得到法向量 $\mathbf{n}$.

但是, 现实世界中, 满足朗伯体假设的材料很少, 为了表示复杂的反射现象, 人们引入BRDF的概念. BRDF是指物体表面某点反射方向的辐射强度与辐照强度的比值, 与入射光线方向 $\mathbf{l}$ 、物体表面法向量 $\mathbf{n}$ 和观测方向 $\mathbf{v}$ 有关. 在以法向量 $\mathbf{n}$ 为 z 轴的坐标系中, 可以用 $\theta_i, \phi_i, \theta_e, \phi_e$ 表示入射光线方向 $\mathbf{l}$ 和观测方向 $\mathbf{v}$, 其中 θ_i 为入射光线方向的俯仰角, ϕ_i 为入射光线方向的方位角, θ_e 为观测方向的俯仰角, ϕ_e 为观测方向的方位角, 如图2所示. 引入BRDF后, 可得式(3), 其中 $f_{\text{BRDF}}(\mathbf{l}, \mathbf{n}, \mathbf{v})$ 表示BRDF.

$$I_p = f_{\text{BRDF}}(\mathbf{l}, \mathbf{n}, \mathbf{v}) k_c I_s T \mathbf{n}^T \mathbf{l}. \quad (3)$$

2.2 数据驱动光度立体视觉算法

通常, 非朗伯体材料的BRDF是高度非线性的, 而且不同材料的BRDF不一样. 因此, 用某种材料的稠密

图 2 (网络版彩图)BRDF方向表示

测量数据去表示该材料的BRDF, 比统一解析模型更合适. 本文的基本思想就是根据某种材料的测量数据库得到合理的数学模型, 紧接着根据该数据模型对相应材料数据库中的数据进行推断. 具体来说, 本文建立了入射方向 $\mathbf{l}$ 、观测方向 $\mathbf{v}$ 、像素记录值 I_p 与BRDF之间的映射模型, 并根据该映射模型进行数据推断.

为了实现上述思想, 本文利用高斯过程建立数学模型. 高斯过程是一种贝叶斯推理模型, 是一种处理高度非线性问题的强大数学工具. 在拍摄距离远大于物体尺寸的假设下, 观测方向 $\mathbf{v}$ 可以认为是常向量 $(0,0,1)^T$, 因此推导过程中, $\mathbf{v}$ 不作为输入量. 这样高斯过程就建立了入射光线方向 $\mathbf{l}$ 、像素记录值 $\tilde{I}_p$ 与 $\tilde{f}_{\text{BRDF}}$ 之间的映射关系, 如式(4)所示. 为了保证同一像素不同光照方向下像素记录值的变化只是由BRDF和光线入射方向不同引起的, 因此需要消除标定入射光线强度 $k_c I_s T$ 对像素记录值 I_p 的影响. 该模型就使用了消除标定入射光线强度影响的像素记录值 $\tilde{I}_p = \frac{I_p}{k_c I_s T}$. 为了方便后面法向量求解, 该模型使用了BRDF的倒数 $\frac{1}{f_{\text{BRDF}}}$, 记做 $\tilde{f}_{\text{BRDF}}$.

$$GP(\mathbf{X}) \sim N(\mu(\mathbf{X}), K(\mathbf{X}, \mathbf{X})), \quad (4)$$

其中, $\mathbf{X} = (\mathbf{L}, \tilde{\mathbf{I}})$, 是 $n \times 4$ 阶输入量,

$$\mathbf{L} = (\mathbf{l}_{p1}, \mathbf{l}_{p2}, \dots, \mathbf{l}_{pn})^T, \quad \tilde{\mathbf{I}} = (\tilde{I}_{p1}, \tilde{I}_{p2}, \dots, \tilde{I}_{pn})^T,$$

$\mathbf{l}_{pi}$ 为 $\tilde{I}_{pi}$ 对应的入射光线方向, $\mu(\mathbf{X})$ 是均函数方程, $K(\mathbf{X}, \mathbf{X})$ 是 $n \times n$ 阶协方差矩阵, 表达了 $\mathbf{X}$ 之间的相关性, 其中 n 是训练点数量.

建立模型时, 由于不知道任何先验信息, 所以均函

数方程认为是零均值函数. 对于任一想要预测的 $\mathbf{x}^* = (\mathbf{l}^{*T}, \tilde{I}_p^*)$, $\tilde{f}_{\text{BRDF}}^*$ 数值可以通过与已知数据相关的联合分布预测得到, 其中 $\mu(\tilde{f}_{\text{BRDF}}^*)$ 为BRDF预测结果, 具体表达式如式(5)所示.

$$\begin{aligned} \tilde{f}_{\text{BRDF}}^* \mid \mathbf{X}, \mathbf{x}^*, \tilde{f}_{\text{BRDF}} &\sim N(\mu(\tilde{f}_{\text{BRDF}}^*), \text{COV}(\tilde{f}_{\text{BRDF}}^*)), \\ \mu(\tilde{f}_{\text{BRDF}}^*) &= K(\mathbf{x}^*, \mathbf{X}) (K(\mathbf{X}, \mathbf{X}) + \sigma_e^2 \mathbf{I})^{-1} \tilde{f}_{\text{BRDF}}, \\ \text{COV}(\tilde{f}_{\text{BRDF}}^*) &= K(\mathbf{x}^*, \mathbf{x}^*) \\ &\quad - K(\mathbf{x}^*, \mathbf{X}) (K(\mathbf{X}, \mathbf{X}) + \sigma_e^2 \mathbf{I})^{-1} K(\mathbf{X}, \mathbf{x}^*), \end{aligned} \quad (5)$$

其中, $\mathbf{I}$ 是单位矩阵, σ_e^2 表示噪声, 本文所选用的核函数是Matérn核函数.

2.3 数据集

要实现上述的过程, 需要有数据集. 生成数据集, 也就是要获得掩模图、入射方向 $\mathbf{l}$ 、标准法向量 $\mathbf{n}$ 、观测方向 $\mathbf{v}$ 与像素记录值 I_p . 本文涉及两种生成数据集的方法: 一种方法用于生成仿真实验的数据集, 而另一种方法为了获得实际实验的数据集. 仿真实验是基于MERL数据库^[22]展开的, 该库记录了100种材料的BRDF. 本文使用了公开的物体掩模图、物体标准法向量以及入射光线方向. 在拍摄距离远大于拍摄物体尺寸的假设下, 虚拟相机方向即观测方向 $\mathbf{v}$ 是常向量 $(0,0,1)^T$. 这样就得到了入射方向 $\mathbf{l}$ 、标准法向量 $\mathbf{n}$ 和观测方向 $\mathbf{v}$, 下一步要根据上述信息计算像素记录值 I_p .

在求解像素记录值 I_p 时, 需要得到每个像素点在每个入射方向下的BRDF. BRDF的4个输入参数都是在局部坐标系下计算得到的, 该局部坐标系是以每个像素点对应的物体表面点为原点, 以该点表面法向量为 z 轴建立的, 如图2所示. 因此在求解BRDF过程中, 需要将虚拟相机坐标系转化为局部坐标系, 随后求得 $\theta_i, \phi_i, \theta_e, \phi_e$ 四元参数, 再通过MERL数据库提供的函数接口得到BRDF. 最后根据渲染方程 $I_p = f_{\text{BRDF}}(\mathbf{l}, \mathbf{n}, \mathbf{v}) k_c J_s T \mathbf{n}^T \mathbf{l}$, 求得 I_p , 其中当 $\mathbf{n}^T \mathbf{l} < 0$ 时, $I_p = 0$ 表示阴影, 同时令 $k_c J_s T = 1$ 以简化问题. 求解 I_p 流程如图3所示.

真实实验时, 我们需要分割得到掩模图; 通过激光扫描或者标准三维模型得到标准法向量值; 场景中放

输入: 入射光线方向 $\mathbf{l}$,
标准法向量 $\mathbf{n}$,
虚拟相机方向 $\mathbf{v}=(0,0,1)^T$
输出: 像素记录值 I_p
1) 虚拟相机坐标系与局部坐标系坐标变换;
2) 计算BRDF四元输入参数 $\theta_i, \phi_i, \theta_e, \phi_e$ 并求得BRDF;
3) 根据渲染方程求得像素记录值 I_p .

图3 渲染算法

入标准镀铬钢球, 标定光源入射方向; 对于 $k_c J_s T$, 可以通过标准白板进行标定. 标定过程中, 标准白板可以认为是均匀的近似朗伯体漫反射平面, 且认为平面法向量是 $\mathbf{n}_{\text{calib}}=(0,0,1)^T$. 根据式(2)可以得到标准白板成像过程关系式为

$$I_{p_{\text{calib}}} = \frac{\rho_{\text{white}}}{\pi} k_c J_s T \mathbf{n}_{\text{calib}}^T \mathbf{l}. \quad (6)$$

标定过程中, 手动分割包含白板的感兴趣区域, 对该区域的像素记录值取均值, 随后求得相对的入射光线强度 $\tilde{k}_c \tilde{J}_s \tilde{T}$, 即

$$\tilde{k}_c \tilde{J}_s \tilde{T} = \frac{\text{mean}(I_{p_{\text{calib}}})}{\mathbf{n}_{\text{calib}}^T \mathbf{l}} = \frac{\rho_{\text{white}}}{\pi} k_c J_s T. \quad (7)$$

需要指出的是, 尽管标定出来的是相对的入射光线强度 $\tilde{k}_c \tilde{J}_s \tilde{T}$, 但由于 $\frac{\rho_{\text{white}}}{\pi}$ 是全局系数, 所以不影响法向量的求解精度. 完成了入射光线强度的标定, 我们就得到了像素记录值 $\tilde{I}_p = \frac{I_p}{\tilde{k}_c \tilde{J}_s \tilde{T}}$, $\tilde{f}_{\text{BRDF}} = \frac{\mathbf{n}^T \mathbf{l}}{\tilde{I}_p}$, 入射方向 $\mathbf{l}$ 和标准法向量 $\mathbf{n}$, 即真实实验的数据集.

2.4 训练与预测

训练和预测过程中, 为了消除阴影点与高光点的影响, 本文将像素记录值按从小到大的顺序排列, 去除部分低值点和部分高值点, 去除像素记录值部分大约占20%~40%. 训练过程中, 把渲染或真实图片中的像素记录值保存为一个矩阵 $\mathbf{O}_{p \times q}$, 其中 p 是掩模图中非零像素记录值的个数, q 是入射光线方向总数. 为了得到训练点, 本文先均匀选取像素点, 接着对每一个像素点选取足够数量的像素记录值, 得到一一对应的 $\mathbf{X}$ 与 $\tilde{f}_{\text{BRDF}}$, 输入该参数进行训练得到高斯模型, 随后基于该模型逐点预测 $\tilde{f}_{\text{BRDF}}^*$. 而对于 $\tilde{f}_{\text{BRDF}}$, 又

有 $\tilde{f}_{\text{BRDF}}^* = \frac{\mathbf{n}^T \mathbf{l}}{\tilde{f}_p}$, 我们取 $\tilde{\mathbf{L}} = \left(\frac{\mathbf{l}_1}{\tilde{f}_{p1}}, \frac{\mathbf{l}_2}{\tilde{f}_{p2}}, \dots, \frac{\mathbf{l}_{n^*}}{\tilde{f}_{pn^*}} \right)^T$, n^* 为该像素点有效像素记录值的数量, 那么由式(8)可得 $\tilde{\mathbf{n}}$, 其中 $\mathbf{f}_{\text{BRDF}}^* = (\tilde{f}_{\text{BRDF}1}^*, \tilde{f}_{\text{BRDF}2}^*, \dots, \tilde{f}_{\text{BRDF}n^*}^*)^T$.

$$\tilde{\mathbf{n}} = \tilde{\mathbf{L}}^+ \mathbf{f}_{\text{BRDF}}^* \quad (8)$$

由于预测过程中存在不确定度, 所以 $\tilde{\mathbf{n}}$ 基本不可能是单位法向量, 还要通过式(9)归一化后, 得到单位法向量 $\mathbf{n}$, 训练与预测算法如图4所示.

$$\mathbf{n} = \frac{\tilde{\mathbf{n}}}{\|\tilde{\mathbf{n}}\|}. \quad (9)$$

2.5 算法流程

上述方法本质上是一种基于高斯过程的全局优化算法, 算法总流程如图5所示. 高斯模型的训练点既包含了像素点的位置信息又包括了光照方向信息, 所以它可以预测任意光照条件下, 任意位置的法向量. 由于该方法是全局优化方法, 与像素级迭代方法相比, 该方法不需要逐点迭代, 因此计算效率较高.

3 实验结果

为了证明本方法的优越性, 本文进行了仿真实验和真实实验, 并与传统的最小二乘算法(LS算法)、先进的约束双变量回归算法(CBR算法)进行了比较. 三个算法均在Matlab 2016b软件中运行, 计算机配置为Intel-i5-7500型号的CPU, 主频为3.4 GHz, 内存为8 G.

3.1 标准数据库实验比较

为了证明该方法的有效性和训练模型的扩展性,

输入: 像素记录值矩阵 $\mathbf{O}_{p \times q}$,
入射光线方向 $\mathbf{l}$
输出: 法向量 $\mathbf{n}$
1) 均匀采样, 得到一一对应的 $\mathbf{X}$ 与 $\tilde{f}_{\text{BRDF}}$;
2) 训练得到高斯模型;
3) 基于训练的高斯模型, 逐点预测, 得到 $\tilde{f}_{\text{BRDF}}$;
4) 计算 $\tilde{\mathbf{L}}$ 并求得 $\tilde{\mathbf{n}}$;
5) 归一化求得 $\mathbf{n}$.

图4 训练与预测算法

本文分别生成了球与兔子两种形状的渲染图片. 因为CBR算法需要300张图片才能得到比较好的实验结果, 所以为了更好地与CBR算法比较, 本文渲染了300张图片. 其中球渲染图片的分辨率为 128×128 , 兔子渲染图片的分辨率为 256×256 . 图6展示了5种材料的渲染图片, 为了清楚地展示渲染图片, 每种材料的渲染数据均进行了同倍数增大.

因为球包含了所有的法向量信息, 所以本文选择球的数据集作为高斯模型的训练数据. MERL库中记录的是 r, g, b 三个通道的BRDF, 但上文推导过程是针对黑白图片的, 因此实验过程中用Matlab的rgb2gray函数将彩色图片转化为黑白图片. 如2.4节所述, 选择非零像素记录值以消除阴影点的影响; 对像素记录值按从小到大的顺序排列, 取小于序列中60分位数的像素记录值以消除高光点的影响, 由此得到像素记录值矩阵 $\mathbf{O}_{p \times q}$. 准备训练点时, 先均匀选择10~40个像素点, 随后对于每个像素点, 选择前5~20非零像素记录值, 这样得到了50~800个训练点. 对该训练点训练得到高斯模型, 并根据该模型进行预测. 需要指出的是, 渲染过程中, 每种材料会生成大约 4×10^6 个数据, 50~800个训练点会显得代表性不够. 但一方面高斯过程的优势是训

图5 (网络版彩图)算法总流程

图 6 (网络版彩图)球与兔子部分材料渲染图

训练样本量小, 另一方面实际实验表明, 增加训练点数量不会进一步提高预测精度, 反而会增加预测时间, 因此该方法只需要较少的训练点就可以得到较好的法向量预测精度。

与大多数文献一样, 本文采用相同的方法表示法向量估计图, 即让彩色图片的3个通道 r , g , b 分别代表单位法向量的3个分量 n_x , n_y , n_z , 其中 n_x , n_y , n_z 的取值范围均为 $[-1, 1]$ 。为了方便显示, 将这3个分量归一化至 $[0, 1]$ 区间, 即 $I_{\text{show}} = \frac{I_{\text{true}} + 1}{2}$, 其中 I_{show} 是展示法向量估计图, I_{true} 是真实法向量估计图。为了清晰地表示法向量的估计误差, 实验中生成了相应的误差图, 该误差图是估计法向量与真值法向量夹角的分布图。图中颜色越偏蓝, 夹角越小, 误差越小, 颜色越偏红, 夹角越大, 误差越大。

图7展示的是真值法向量图以及3种方法的法向量估计图。与图6一样, 图7展示了相同的5种材料。我们可以直观地看出, DPS方法的法向量估计图与真值图非常接近, 而其他两种方法与真值图有较大的出入。图8展示的误差图更清楚地表明了这一点, DPS方法的误差图基本是蓝色, 表明误差小。此外, 100种材料的实验结果整理为图9(a), 图9(a)纵坐标表示的是估计法向量与真值法向量的夹角角度值, DPS算法曲线位于最下端, 预测精度最高。

为了验证训练模型的扩展性, 即该模型可以预测其他形状的物体, 本文又进行了兔子图片实验。实验过程中没有重新训练高斯模型, 而是直接将之前训练好的高斯模型应用到该场景。图9(b)展示的是100种材料的实验结果, 纵坐标意义与图9(a)相同。比较图9(a)

图 7 (网络版彩图)5种材料法向量恢复图(从上到下: alum-bronze, color-changing-paint2, ipswich-pine-221, pvc和ss440)

和(b)可以看出, DPS算法的预测精度保持稳定。

3.2 真实实验比较

本文搭建了一个光度立体视觉实验平台。该平台保证暗室环境, 并用46盏20 W的点光源照射拍摄距离

图 8 (网络版彩图)五种材料误差图(从上到下: alum-bronze, color-changing-paint2, ipswich-pine-221, pvc和ss440)

为1.8 m的物体, 拍摄物体是100 mm×100 mm的铝合金抛物面. 相机型号为大恒相机水星系列MER-502-79U3C工业相机, 图片分辨率为2448×2048. 本文假设该抛物面的加工误差很小, 抛物面的真值法向量可以通过相应数学模型计算得到. 在自动生成掩模图后, 通过插值就可以得到标准法向量图, 图10展示了标准法向量图和其中一张真实照片.

在场景中放入镀铬钢球标定光源入射方向, 随后放入标准白板, 按2.3节所述方法进行光源强度标定.

由于本实验用的是彩色相机, 所以实验中得到的是 r , g , b 三个通道的光源强度值. 准备训练数据时, 每个通道的像素记录值除以该通道的光源强度值以消除光源强度的影响, 再对每个通道进行中值滤波, 随后将彩色图片转为黑白图片. 选择训练点时, 先均匀采10个像素点, 然后对每个像素点均匀选5个非零像素记录值. 按照上述数据训练并预测, 得到法向量估计图和误差图如图11所示. 从图中可以看出, DPS算法的误差比LS和CBR要小.

图9 (网络版彩图)球平均误差角度变化图(a)和兔子平均误差角度变化图(b)

图10 (网络版彩图)标准法向量图和实际照片

4 讨论与结论

本文提出一种基于高斯过程的数据驱动光度立体视觉全局优化算法,该方法利用高斯过程建立入射光

线方向 $\mathbf{l}$ 、观测方向 $\mathbf{v}$ 、像素记录值 I_p 与BRDF之间的映射关系,并根据该高斯模型预测法向量.高斯过程在处理高度非线性问题上有明显优势,因此法向量恢复精度得到了很大的提高.此外,本文用一种全局高斯模型描述非朗伯体材料的非线性反射性质,不需要逐点迭代得到预测值,并且较少的训练数据就可以得到该全局模型,所以预测时间短.仿真实验与真实实验验证了上述优越性.

表1给出了仿真实验的误差角度值,即估计法向量与真值法向量的夹角角度值.表中既包含了球实验的误差数据,又包含了兔子实验的误差数据.针对每种形状,3种方法的误差角度值均被记录下来.从表中可以看到,球实验中,DPS算法100种材料平均角度误差为 0.208° ,而LS算法和CBR算法的误差分别为 12.7° 和 2.13° ,DPS算法大大提高了法向量恢复精度.从兔子

图 11 (网络版彩图)法向量恢复图与误差图(从上往下: LS, CBR, DPS)

表 1 仿真实验角度误差表

角度误差	球			兔子		
	LS	CBR	DPS	LS	CBR	DPS
alum-bronze	24.6°	2.12°	0.155°	26.7°	2.29°	0.144°
color-changing-paint2	28.6°	3.93°	0.170°	31.9°	4.26°	0.183°
ipswich-pine-221	12.9°	1.85°	0.191°	14.5°	1.73°	0.187°
pvc	9.89°	1.50°	0.103°	11.6°	1.58°	0.0990°
ss440	30.8°	4.89°	0.644°	33.3°	5.23°	0.640°
平均误差	12.7°	2.13°	0.208°	13.8°	2.17°	0.211°

实验记录的误差角度值中可以看出, 本文训练的高斯模型扩展性较好, 即用该模型预测其他形状的物体, 法向量恢复精度保持稳定.

表2给出了实际实验3种方法的误差角度值与计算时间. 本文记录的计算时间是法向量估计程序在Matlab中运行的时间, 预测像素数量为700140. 从表中可

表 2 真实实验误差值与计算时间

方法	角度误差	计算时间 (s)
LS	36.5°	22.6
CBR	30.9°	397.64
DPS	12.1°	84.6

以看出, DPS在获得最高的精度的同时, 比CBR计算时间缩短了许多. 需要指出的是, 真实实验的误差与仿真数据相比高了很多, 其中很多因素导致这一现象的产生, 比如用点光源近似平行光源、加工误差、阴影点高光点的影响等.

综上两表所述, 该方法法向量估计精度高于LS算法与CBR算法, 计算效率比CBR算法高. 正是由于上述优势, 该方法在大规模单一材料表面缺陷检测方面有较好的应用前景.

本文所提出的算法目前还只适用于需暗室环境中的光度立体视觉, 尚无法处理自然光尤其是时变环境中自然光所造成的影响. 该问题是当前光度立体视觉最具挑战的难点之一, 也是本文后续研究的兴趣点. 此外, 本文所提出的方法对非朗伯体材料的光反射模型进行了全局建模, 避免了逐点迭代, 可通过CPU多线程或者GPU并行化计算, 进一步提升计算效率.

参考文献

- 1 Horn B K P. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. Dissertation for Doctoral Degree. Cambridge: Massachusetts Institute of Technology, 1970
- 2 Woodham R J. Photometric method for determining surface orientation from multiple images. [Opt Eng](#), 1980, 19: 191139
- 3 徐科, 周鹏, 杨朝霖. 基于光度立体视觉的金属板带表面微小缺陷在线检测. [机械工程学报](#), 2013, 49: 25–29
- 4 程耀瑜, 常国立, 田晓杰. 基于光度立体技术的地板疵点检测方法. [科学技术与工程](#), 2017, 30: 263–267
- 5 刘瑞鑫, 李立轻, 汪军, 等. 基于光度立体视觉的织物褶裥等级评定. [东华大学学报\(自然科学版\)](#), 2013, 39: 48–52
- 6 李健, 马泳潮, 张玉杰, 等. 利用改进的彩色光度立体法实现非刚性体三维数字化. [计算机辅助设计与图形学学报](#), 2015, 27: 1750–1758
- 7 Coleman Jr. E N, Jain R. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. [Comput Graphics Image Processing](#), 1982, 18: 309–328
- 8 Solomon F, Ikeuchi K. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. [IEEE Trans Pattern Anal Machine Intell](#), 1996, 18: 449–454
- 9 Barsky S, Petrou M. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. [IEEE Trans Pattern Anal Machine Intell](#), 2003, 25: 1239–1252
- 10 Mukaigawa Y, Ishii Y, Shakunaga T. Analysis of photometric factors based on photometric linearization. [J Opt Soc Am A](#), 2007, 24: 3326
- 11 Miyazaki D, Hara K, Ikeuchi K. Median photometric stereo as applied to the Segonko Tumulus and museum objects. [Int J Comput Vis](#), 2010, 86: 229–242
- 12 吴仑, 王涌天, 刘越. 一种鲁棒的基于光度立体视觉的表面重建方法. [自动化学报](#), 2013, 39: 1339–1348
- 13 Ikehata S, Wipf D, Matsushita Y, et al. Robust photometric stereo using sparse regression. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, 2012
- 14 Goldman D B, Curless B, Hertzmann A, et al. Shape and spatially-varying BRDFs from Photometric Stereo. [IEEE Trans Pattern Anal Mach Intell](#), 2010, 32: 1060–1071
- 15 Ward G. Measuring and modeling anisotropic reflection. [ACM SIGGRAPH](#). 1992, 26: 265–272
- 16 Higo T, Matsushita Y, Ikeuchi K. Consensus photometric stereo. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, 2010
- 17 Shi B X, Tan P, Matsushita Y, et al. A biquadratic reflectance model for radiometric image analysis. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, 2012
- 18 Ikehata S, Aizawa K. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In: The IEEE Conference on Computer Vision and Pattern Recognition. Columbus, 2014
- 19 Chandraker M, Ramamoorthi R. What an image reveals about material reflectance. In: 2011 International Conference on Computer Vision. Barcelona, 2011

- 20 Han T Q, Shen H L. Photometric stereo for general BRDFs via reflection sparsity modeling. *IEEE Trans Image Process*, 2015, 24: 4888–4903
- 21 Shen H L, Han T Q, Li C. Efficient photometric stereo using kernel regression. *IEEE Trans Image Process*, 2016, 26: 439–451
- 22 Matusik W, Pfister H, Brand M, et al. A data-driven reflectance model. Dissertation for Doctoral Degree. Cambridge: Massachusetts Institute of Technology, 2003

A data-driven photometric stereo algorithm

WANG Xi, REN MingJun, XIAO GaoBo, CHEN MingHan & FU Lin

State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Representing the complicated reflectance of non-Lambertian material is one of the difficulty in photometric stereo. This paper presents a Gaussian process based data-driven photometric stereo method to address this problem. The method makes use of Gaussian process, a non-parametric Bayesian inference model, to establish a global plausible model to represent non-linear reflectance property of non-Lambertian material based on measured bidirectional reflectance distribution functions. The statistic nature of the Gaussian process makes the method has much better flexibility and accuracy in describing the complicated reflectance of various non-Lambertian materials so that a better estimation accuracy can be achieved for a photometric stereo system. The validity of the proposed method is verified by both computer simulation and real experiment. The results show that the proposed method has much better accuracy and efficiency in estimating surface normal than other existing methods.

data-driven, photometric stereo, Gaussian process

doi: [10.1360/N092018-00035](https://doi.org/10.1360/N092018-00035)