

# 一种鲁棒的基于光度立体视觉的表面重建方法

吴仑<sup>1</sup> 王涌天<sup>1</sup> 刘越<sup>1</sup>

**摘要** 提出一种基于先进的凸优化技术的光度立体视觉重建框架. 首先通过鲁棒的主成分分析 (Robust principle component analysis, RPCA) 祛除图像噪声, 得到低秩矩阵和物体表面向量场, 然后再通过表面重建算法从向量场来恢复物体形状. 相对于先前的一些使用最小二乘或者一些启发式鲁棒技术的方法, 该方法使用了所有可用的信息, 可以同时修复数据中的丢失和噪声数据, 显示出了较高的计算效率以及对于大的稀疏噪声的鲁棒性. 实验结果表明, 本文提出的框架大大提高了在噪声存在情况下物体表面的重建精度.

**关键词** 光度立体视觉, 鲁棒主成分分析, 表面重建, 稀疏噪声, 低秩矩阵

**引用格式** 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(8): 1339–1348

**DOI** 10.3724/SP.J.1004.2013.01339

## A Robust Approach Based on Photometric Stereo for Surface Reconstruction

WU Lun<sup>1</sup> WANG Yong-Tian<sup>1</sup> LIU Yue<sup>1</sup>

**Abstract** We present a new framework for surface reconstruction with technique of photometric stereo, which is based on advanced convex optimization technique. We firstly remove the errors in images by robust principle component analysis (RPCA), and then obtain low-rank matrix and surface normal field. Unlike previous approaches, this method uses all the available information to simultaneously fix missing and erroneous entries. The new technique is more computationally efficient and provides theoretical assurance for robustness to large errors. Experimental results demonstrate that this framework can improve the precision for surface reconstruction with noise.

**Key words** Photometric stereo, robust principle component analysis (RPCA), surface reconstruction, sparse error, low-rank matrix

**Citation** Wu Lun, Wang Yong-Tian, Liu Yue. A robust approach based on photometric stereo for surface reconstruction. *Acta Automatica Sinica*, 2013, 39(8): 1339–1348

三维表面重建一直是计算机视觉领域重要的研究方向. 使用三维激光扫描仪或接触式的测量等手段可以获得非常高的表面精度, 但是需要的硬件设备昂贵、自动化程度较低, 同时还受限于物体尺寸. 随着计算机视觉理论的逐渐成熟, 使用图像来获取物体表面信息的方法得到了广泛的应用.

光度立体视觉是 1980 年 Woodham<sup>[1]</sup> 提出的一种由不同光源下拍摄的多帧图像恢复物体表面法向量的方法, 作为一种便捷的非接触式测量方法被广泛应用于物体的三维表面重建中. 相对于其他基于三角化方式的方法来说, 光度立体视觉可以生成相对稠密的向量场. 物体表面深度则可以通过求解

向量场和物体深度构成的偏微分方程或者对向量场积分来获得, 因此物体表面向量场恢复的准确程度直接影响了物体表面恢复的精度.

在朗伯体表面反射模型下, 物体表面一点的法向量可以由三幅固定视点、不同光源情况下的图像对应的像素唯一的决定. 在不存在阴影的情况下, 对于凸的朗伯体, 由来自不同方向的光源照射产生的图像集合张成一个三维的子空间<sup>[2]</sup> 或者一个光锥<sup>[3]</sup>. 在存在投射阴影 (Cast shadow) 的情况下, 其图像所张成的空间仍然可以被一个低维度的子空间来近似<sup>[4–5]</sup>.

在光源已经标定, 也就是光源方向已知的情况下, 物体表面的法向量通常可以采用最小二乘法来获得. 通过求解由观察图像和光源构成的线性方程来求解物体表面的向量场.

在光源未知的时候, 这就是一个双线性问题. 针对这个问题, Hayakawa<sup>[6]</sup> 提出了一种矩阵分解的框架. 首先通过主成分分析 (Principal component analysis, PCA)<sup>[7]</sup> 分解得到和真实值相差一个可逆的线性变换的表面法向和光源方向. 这个变换可以通过一些附加的约束来消除, 例如可以假设至少 6

收稿日期 2011-12-12 录用日期 2012-07-15

Manuscript received December 12, 2011; accepted July 15, 2012  
国家重点基础研究发展计划 (973 计划) (2010CB732505), 国家自然科学基金 (61072096) 资助

Supported by National Basic Research Program of China (973 program) (2010CB732505) and National Natural Science Foundation of China (61072096)

本文责任编辑 戴琼海

Recommended by Associate Editor DAI Qiong-Hai

1. 北京理工大学光电学院 北京 100086

1. School of Optoelectronics, Beijing Institute of Technology, Beijing 100086

个光源的亮度相同或者至少 6 个表面法向的反射率是相等的. 在这种情况下这个变换可以被矫正到和真实值相差一个旋转变换的程度. 理论上, 使用积分约束, 可以进一步将这个变换矫正到和真实值相差一个通用浮雕 (Generalized bas-relief, GBR)<sup>[8]</sup> 变换的程度. 然后使用文献 [9] 中的算法消除 GBR 多义性, 从而得到最后法向量的估计结果.

当观察图像中的噪声是高斯的或者小尺度的, 使用最小二乘或者矩阵分解的方法可以得到比较好的结果. 但是在实际情况中, 由于一些比较复杂的因素, 如镜面反射以及阴影等的影响, 不同光源下的图像不再具备低秩的结构, 这样就导致上述方法不再有效. 如何有效地去除这些因素的影响, 恢复出物体表面的法向量成为人们研究的焦点. 在光照情况比较理想、噪声所占比例比较少的情况下, 可以通过丢弃被噪声影响的数据, 或者设置全局阈值来移除噪声的简单方法来消除噪声的影响. 但是当噪声数据比较多的情况下, 这些简单的处理将会对估计结果造成较大的误差.

研究者们一直致力于更鲁棒的算法的研究. Hayakawa<sup>[6]</sup> 提出了一种通过矩阵生长的策略来处理阴影, 该方法首先从一个没有被阴影影响的子矩阵开始, 然后通过逐行或者逐列的计算矩阵中没有被噪声影响的数据来达到祛除噪声影响的目的. Jacobs 的线性拟和理论<sup>[10]</sup> 被 Belhumeur 等<sup>[8]</sup> 用于填充被阴影遮挡的区域. Yuille 等<sup>[11]</sup> 提出了一种利用鲁棒统计<sup>[12]</sup> 的策略来祛除阴影的方法, 通过迭代算法来确定阴影区域. Julià 等将两步法 (Two-step) 应用到光度立体视觉中<sup>[13-14]</sup>, 使得计算效率大幅度提高. 针对矩阵中的噪声问题, Chen 等<sup>[15]</sup> 提出了基于 LM (Levenberg-Marquardt) 优化的算法, 有效地降低了参数的数目. 然而这些算法均需要有较好的初始值, 否则将会得到退化的运算结果. 上述算法的另外一个局限性就是需要观察矩阵中不能有大尺度的噪声, 也就是在图像中不能有镜面反射. 近年来, 一些研究者利用随机抽样一致性 (Random sample consensus, RANSAC)<sup>[12]</sup> 算法, 将这些噪声作为野点来进行处理. 如 Mukaigawa 等<sup>[16]</sup> 和 Miyazaki 等<sup>[17]</sup> 的工作. 但是这类算法需要调整较多的参数, 并且速度较慢. 在过去的几十年内, 针对噪声存在的情况下, 传统算法或者计算精度低, 或者需要较多的步骤. 并没有理论可以既精确同时又高效地获得物体表面的法向量.

受到近来在统计数据分析方面, 即鲁棒性的主成分分析 (Robust PCA, RPCA) 工作<sup>[18-19]</sup> 的进展的启发, 我们提出了一种简单的基于光度立体视觉的重建框架, 可以有效地祛除图像表面的噪声, 恢复物体表面深度. 首先采用鲁棒的主成分分析技

术, 通过对核范数 (Nuclear norm) 以及 1 范数 ( $l_1$  norm) 优化来获得对低维度空间的精确估计, 祛除图像的噪声影响, 然后再使用最小二乘法获得物体表面的向量场, 最后再通过基于泊松方程的表面重建算法得到物体表面. 相对于先前的一些基于光度立体视觉的物体表面重建方法, 该方法并不需要刻意地挑选图像, 或者忽略掉被噪声影响的图像数据, 却能获得较高的表面重建精度. 显示出了较传统方法更高的计算效率.

## 1 光度立体视觉表面重建

在介绍光度立体视觉概念的基础上, 本节将给出数学模型以及整体的解决方案.

### 1.1 光度立体视觉中的低秩结构

设图像  $I_j$  上第  $k$  点亮度为  $i_j^k$ , 其对应的表面单位法向量为  $n_k$ , 光源方向为  $l_j$ , 物体表面反射率为  $\rho_k$ , 则根据朗伯体表面反射模型有:

$$i_j^k = \rho_k n_k \cdot l_j \quad (1)$$

假设有  $m$  个像素点和  $n$  个光源, 将所有观察图像向量化并组合成矩阵  $D \in \mathbf{R}^{m \times n}$  有

$$D = [\text{vec}(I_1) | \cdots | \text{vec}(I_n)] \quad (2)$$

$D$  可以被分解成为

$$D = NL \quad (3)$$

其中,  $N = [\rho_1 n_1 | \cdots | \rho_m n_m]^T \in \mathbf{R}^{m \times 3}$ ,  $L = [l_1 | \cdots | l_n] \in \mathbf{R}^{3 \times n}$ , 在理想情况下, 矩阵  $D$  的秩应该为 3 或者接近于 3.

### 1.2 光度立体视觉中的噪声

在真实的数据中, 因为阴影以及镜面反射的存在, 上述的低秩的结构是很难被观察到的. 为了准确地恢复出低秩的结构, 有必要先来分析一下这些噪声的属性. 由于遮挡原因而形成的阴影被称为投射阴影 (Cast shadow), 由于物体表面法向背离光源而形成的阴影被称为附属阴影 (Attached shadow), 在理想情况下, 这两种阴影反映在图像上均为黑色区域. 这样图像像素亮度值则表示为

$$i_j^k = \max\{\rho_k n_k \cdot l_j, 0\} \quad (4)$$

如果观察的物体不是纯朗伯体, 就会产生阴影和镜面反射. 镜面反射的亮度和观察者的角度有关, 当观察角度和反射角度一致时, 就会观察到高亮度区域, 也就是说镜面反射部分的亮度值总会较大. 如果把上述噪声使用一个噪声矩阵来表示的话, 观察到的图像测量矩阵  $D$  被修改为

$$D = NL + E \quad (5)$$

假设观察到的图像仅仅有少部分的像素点被高光部分和阴影所覆盖, 则大多数的像素点都符合式 (5) 所给出的低秩的模型, 也就是说  $E$  中的大多数值都为 0,  $E$  是一个稀疏的矩阵. 令  $A = NL$ , 则  $A$  为一个秩为 3 的矩阵, 这样, 问题可以转化为: 给定一个观察矩阵  $D$ , 它是由一个低秩的矩阵和一个稀疏的矩阵组合而成, 求最接近低秩矩阵以及稀疏矩阵的分解.

### 1.3 稀疏噪声下的低秩矩阵恢复

使用拉格朗日 (Lagrangian) 公式, 上述问题可以被写成一个在稀疏噪声下的低秩矩阵恢复的优化问题, 即:

$$\begin{aligned} \min_{A,E} \text{rank}(A) + \lambda \|E\|_0 \\ \text{s.t. } D = A + E \end{aligned} \quad (6)$$

使用  $\hat{A}$ ,  $\hat{N}$  来表示上述优化问题的最优解. 在光源方向已知的情况下, 表面法向矩阵可以很容易地通过  $N = AN^{\text{inv}}$  得到, 其中  $N^{\text{inv}}$  为  $N$  的伪逆. 在光源未知的情况下, 通过文献 [9] 中介绍的方法来求解未标定的光源以及表面法向.

### 1.4 从向量场恢复物体表面

在得到物体表面的法向量  $N$  后, 可以通过向量场来重建物体表面. 在将法向量的第 3 项归一化后, 可以将  $N$  表示为  $N = (p, q)$ , 其中  $p, q$  表示图像平面方向上的法向分量. 令  $(p_x, q_y)$  为图像平面方向上的方向导数,  $Z$  为表面深度,  $(Z_x, Z_y)$  为深度在图像平面上的方向导数, 则从向量场恢复物体表面深度转化为以下优化问题:

$$\min_Z \iint ((Z_x - p)^2 + (Z_y - q)^2) dx dy \quad (7)$$

令  $E = ((Z_x - p)^2 + (Z_y - q)^2)/2$ , 由 Euler-Lagrange 方程可以得到:

$$\frac{\partial E}{\partial Z} - \frac{d}{dx} \frac{\partial E}{\partial Z_x} - \frac{d}{dy} \frac{\partial E}{\partial Z_y} = 0 \quad (8)$$

其中,  $\frac{\partial E}{\partial Z} = 0$ ,  $\frac{\partial E}{\partial Z_x} = Z_x - p$ ,  $\frac{\partial E}{\partial Z_y} = Z_y - q$ , 带入式 (8) 得到由深度和表面向量场构成的泊松方程:

$$\nabla^2 Z = p_x + q_y \quad (9)$$

使用文献 [20] 中的算法可以快速对方程进行求解, 从而得到物体表面的深度.

## 2 算法

在本节中将介绍对式 (6) 的优化问题的求解方法, 首先对原问题进行松弛处理, 然后再采用增广的

拉格朗日 (Augmented Lagrangian multiplier) 法对松弛后的问题进行求解.

### 2.1 问题的松弛

式 (6) 所描述的优化问题是一个 NP-hard 问题. 对此, Candès 等<sup>[18]</sup> 和 Chandrasekaran 等<sup>[21]</sup> 提出一种松弛的方法: 用核范数 (矩阵奇异值的和) 替代矩阵的秩, 用 1 范数替代稀疏矩阵的 0 范数, 并证明在矩阵  $A$  秩不是很高、 $E$  中的非零输入不是很多的情况下式 (6) 和下式等价.

$$\begin{aligned} \min_{A,E} \|A\|_* + \lambda \|E\|_1 \\ \text{s.t. } D = F + E \end{aligned} \quad (10)$$

其中,  $\|\cdot\|_*$  和  $\|\cdot\|_1$  分别代表核范数和 1 范数,  $\lambda$  是权重系数. 文献 [18] 中建议  $\lambda$  取  $C/\sqrt{\max\{m, n\}}$  的形式,  $C$  一般设为 1,  $m$  和  $n$  为矩阵的行数和列数. 在这个框架下, 稀疏矩阵中的非 0 输入的位置是假设未知的, 但是如果一部分非 0 输入的位置是已知的, 问题就会变得相对简单. 在光度立体视觉中, 尽管阴影和镜面反射部分都对原来的低秩结构造成了破坏, 但是它们具备不同的属性. 相对于镜面反射部分来说, 阴影部分的位置更容易被定位到, 比如使用一些简单的阈值. 用  $\Omega$  代表矩阵  $A$  中不存在数据的点构成的子空间,  $\pi_\Omega$  代表在子空间上的正投影算子. 用  $\Omega^c$  代表在大小为  $m \times n$  的空间内的补空间.  $\pi_{\Omega^c}$  代表在子空间  $\Omega^c$  上的正投影算子. 此时, 可以将式 (10) 问题写成以下形式:

$$\begin{aligned} \min_{A,E} \|A\|_* + \lambda \|E\|_1 \\ \text{s.t. } \pi_{\Omega^c} W = \pi_{\Omega^c} (A + E) \end{aligned} \quad (11)$$

### 2.2 问题的求解

式 (11) 对应的问题可以被认为是一个半定规划的问题, 利用内点理论就可以求解. 内点理论结构较简单, 但其运算时间复杂度高, 当处理的矩阵较大时, 计算量大、收敛速度慢. 近年来, 人们针对高维矩阵核范数最小化问题做了一系列研究<sup>[22-27]</sup>. 文献 [22] 中提出可以采用增广的拉格朗日乘子法 (Augmented Lagrangian method, ALM)<sup>[28]</sup> 来求解式 (11) 所对应的优化问题. 相对其他算法来说, ALM 方法处理速度快、精度高、内存消耗少. ALM 方法的基本思想是通过最小化增广的拉格朗日函数来代替直接对原始约束优化问题的求解. 首先来看一下增广的拉格朗日函数. 对于一般的目标函数, 假设优化问题如下:

$$\begin{aligned} \min_X f(X) \\ \text{s.t. } A(X) = b \end{aligned} \quad (12)$$

其中,  $f(X)$  是凸函数.  $A$  是线性函数,  $b$  是矢量. 拉格朗日函数可以写成如下形式:

$$\mathcal{L}_\mu(X, Y) = f(X) + \langle Y, b - A(X) \rangle + \frac{\mu}{2} \|b - A(X)\|_f^2 \quad (13)$$

$Y \in \mathbf{R}^{m \times n}$  代表拉格朗日乘子矩阵,  $\mu$  是常数,  $\langle \cdot, \cdot \rangle$  代表矩阵的内积,  $\|\cdot\|_f$  代表矩阵的 Frobenius 范数. 对于适当的拉格朗日乘子矩阵  $Y$  和足够大的常数  $\mu$ , 可以证明, 原始的约束优化问题和增广拉格朗日函数的最小化问题是等价的<sup>[28]</sup>. ALM 方法通过迭代估计拉格朗日乘子  $Y$  和  $(A, E)$  得到最优解  $(\hat{A}, \hat{E})$ . 将该方法应用于式 (11) 所示的目标函数, 其增广拉格朗日函数可以表示如下:

$$\mathcal{L}_\mu(X, Y) = \|A\|_* + \lambda \|E\|_1 + \langle Y, \pi_{\Omega^c}(D - A - E) \rangle + \frac{\mu}{2} \|\pi_{\Omega^c}(D - A - E)\|_f^2 \quad (14)$$

利用 ALM 迭代更新  $Y, A, E$ :

$$\begin{cases} (A_{k+1}, E_{k+1}) = \arg \min_{A, E} \mathcal{L}_{\mu_k}(A, E, Y_k) \\ Y_{k+1} = Y_k + \mu_k \pi_{\Omega^c}(D - A_{k+1} - E_{k+1}) \\ \mu_{k+1} = \rho \cdot \mu_k \end{cases} \quad (15)$$

其中,  $\mu_k$  是一个单调递增的正数序列  $\rho > 1$ . 对于上述迭代中的第一步, 很难通过最小化  $\mathcal{L}_{\mu_k}(\cdot)$  同时获得  $A$  和  $E$  的值. 为此, 采用如下的交替最小化策略:

$$\begin{cases} E_{j+1} = \arg \min_E \lambda \|E\|_1 - \langle Y_k, \pi_{\Omega^c}(E) \rangle + \frac{\mu_k}{2} \|\pi_{\Omega^c}(D - A_j - E)\|_f^2 \\ A_{j+1} = \arg \min_A \|A\|_* - \langle Y_k, \pi_{\Omega^c}(A) \rangle + \frac{\mu_k}{2} \|\pi_{\Omega^c}(D - A - E_{j+1})\|_f^2 \end{cases} \quad (16)$$

式 (16) 中的第一步可以通过下述方法进行求解:

$$E_{j+1} = \text{shrink} \left( \pi_{\Omega^c}(D) + \frac{1}{\mu_k} Y_k - \pi_{\Omega^c}(A_j) \frac{\lambda}{\mu_k} \right) \quad (17)$$

其中, 收缩算子 (Shrink operator) 定义如下:

$$\text{shrink}(x, \alpha) = \text{sgn}(x) \cdot \max\{|x| - \alpha, 0\} \quad (18)$$

对于式 (16) 中  $A$  的求解, 使用快速渐近梯度法 (An accelerated proximal gradient)<sup>[27-28]</sup> 进行迭代的方法. 具体迭代过程由下式给出:

$$\begin{cases} (U_i, \Sigma_i, V_i) = \text{svd} \left( \frac{1}{\mu_k} Y_k + \pi_{\Omega^c}(D) - E_{j+1} + \pi_{\Omega^c}(Z_i) \right) \\ A_{i+1} = U_i \text{shrink} \left( \Sigma_i, \frac{1}{\mu_k} \right) V_i^T \\ Z_{i+1} = A_{i+1} + \frac{t_i - 1}{t_{i+1}} (A_{i+1} - A_i) \end{cases} \quad (19)$$

其中,  $\text{svd}(\cdot)$  代表奇异值分解,  $t_i$  是一个  $t_1 = 1, t_{i+1} = 0.5(1 + \sqrt{1 + 4t_i^2})$  的正序列.

### 3 实验

在本节中, 将分别在仿真数据以及真实数据上验证算法的有效性, 并与常用的几种算法: 最小二乘法 (Least square, LS)、基于 RANSAC 的算法 (RANSAC)<sup>[16]</sup>、两步法 (Two step, TS)<sup>[13-14]</sup>、基于 LM 优化的方法 (LM)<sup>[15]</sup> 进行了对比. 本文所有实验均在 Matlab 2010 a, i7 4 核处理器, 3 G 主频, 16 G 内存, 64 位 Windows7 系统环境下进行. 为了保证算法比较的公平性, 本文方法和其他算法均使用相同的阈值来对阴影区域进行标记.

#### 3.1 仿真数据

首先我们使用仿真数据来展示本文算法性能.

##### 3.1.1 初步实验

在本节中使用了 40 幅不同光源下的图像, 如图 1 所示, 图像大小为 288 像素  $\times$  213 像素. 光源方向随机分布在以物体为中心的半球上. 镜面反射模型采用了 Cook-Torrance<sup>[29]</sup> 模型, 阴影阈值设置为 0. 在图像中, 平均 10% 为阴影区域, 30% 为镜面反射区域.

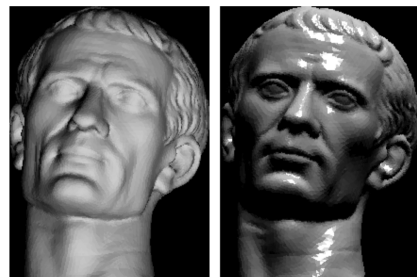
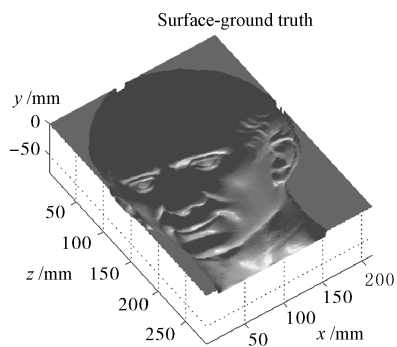


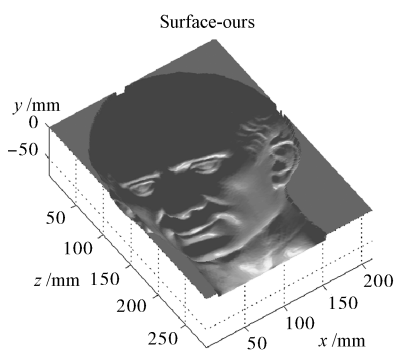
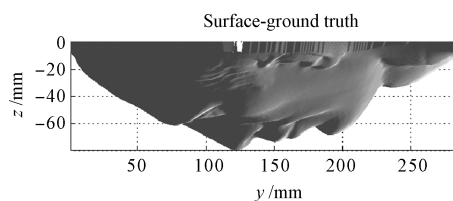
图 1 样本图像

Fig. 1 Sample images

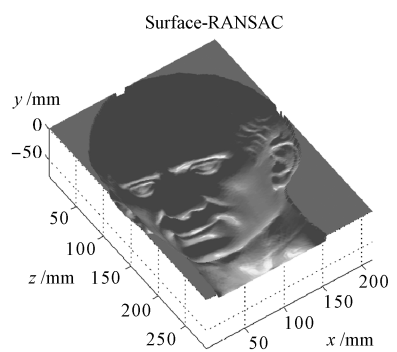
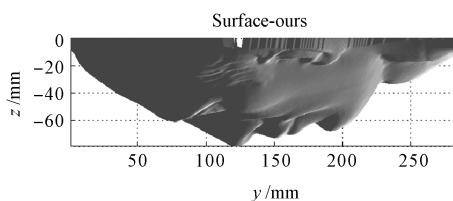
图 2 分别显示了恢复后的物体表面的整体图以及侧面图. 图 2(a) 为真实物体表面, 图 2(b)~2(f) 分别为使用本文方法、RANSAC、LM、TS、LS 方法后得到的表面重建结果, 表 1 定量地比较了这几种算法的性能.



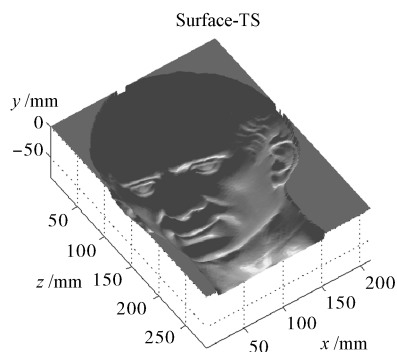
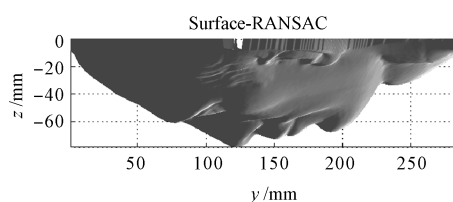
(a) 标准值  
(a) Ground truth



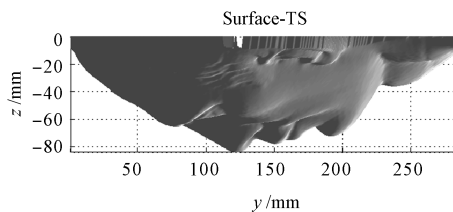
(b) 本文算法  
(b) Results with the proposed method



(c) RANSAC算法  
(c) Results with RANSAC



(d) TS算法  
(d) Results with TS



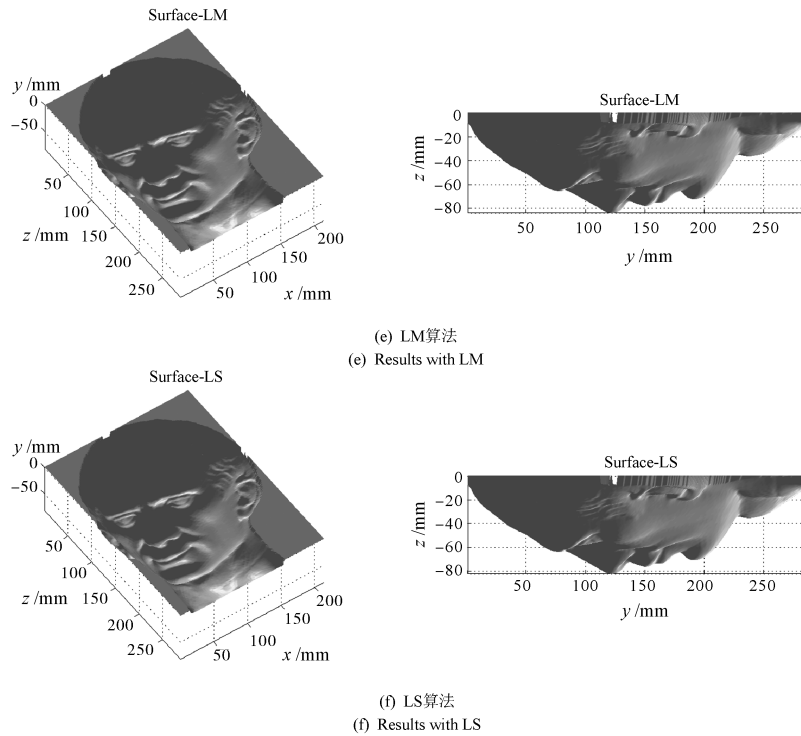


图 2 表面重建结果

Fig. 2 Reconstruction results

表 1 重建误差

Table 1 Reconstruction errors

算法	最大投影误差 (MDE)	平均投影误差 (MSE)	运算时间 (s)	表面深度范围 (mm)
本文算法	2.4238	0.3038	59	(0, -79.3197)
RANSAC 算法	2.4643	0.3290	1 275	(0, -79.2700)
TS 算法	4.9299	2.1723	25	(0, -84.4318)
LM 算法	4.9099	2.1572	226	(0, -84.4104)
LS 算法	3.0530	1.1586	13	(0, -82.5682)

从图 2 中可以看出,使用本文算法以及 RANSAC 算法得到的表面重建精度较高,而其他几种方法的结果标准值有较为明显的差距.表 1 显示了本文方法和其他几种算法在精度上以及运算时间上的对比结果,标准深度范围为 (0, -79.5971) mm.在此使用了均方根误差 (Mean square errors, MSE) 以及最大深度误差 (Maximum depth errors, MDE) 来衡量重建精度.从表 1 中可以看出,无论是误差的平均值,还是误差最大值,本文所提方法均优于其他几个算法, RANSAC 算法的结果虽然精度上和本文算法差别不大,但是其运算速度大大低于本文算法.

首先来分析一下 RANSAC 算法.该算法分为两个阶段: 1) 系数估计阶段; 2) 图像线性化阶段.每个阶段均采用 RANSAC 策略<sup>[12]</sup>. RANSAC 算

法精度是建立在大样本基础上的,因此需要较大样本数才可以保证结果以较高的概率落在正确的解空间内,这就导致了运算量的增加,速度慢是该算法的一个首要的缺点;设置最佳采样数可以有效地降低计算量,但是最佳采样数很难获取,这是因为采样数与样本的分布是有关系的,而在多数情况下,样本分布是未知的,通常的处理方法是设置一个最大迭代次数,这就会导致得到次优的解,这是 RANSAC 算法的第二个缺点; RANSAC 算法需要手工调节两个阈值参数,分别为阴影阈值以及像素亮度值的相对比例系数,并且需要首先选择三幅朗伯反射较多的图像作为基图像初始化,这给算法带来了相当大的不便性,这是 RANSAC 算法的第三个缺点.相对于 RANSAC 算法,本文提出的算法不仅在精度上毫不

逊色, 并且在计算速度以及方便性上具有明显的优势.

TS 算法和 LM 算法都是基于矩阵分解的算法, 前者具备较好的收敛速度, 而后者在丢失数据较多的情况下具有较好的收敛性能. 在只存在阴影的情况下, 这两种算法都可以得到较好的结果. 但是这两种算法都只考虑了丢失数据而没有考虑大尺度的噪声, 因此对于镜面反射没有很好的抑制作用. 本文算法考虑了大尺度噪声的影响, 得到了更高精确的结果, 如表 1 所示. 另外, TS 算法和 LM 算法都需要一个较好的初始值才能保证收敛. 本文算法是基于凸函数优化的, 因此对于初始值的好坏不敏感, 初始值均设置为 0 即可. 可以看出 TS 算法具有速度上的优势, 但是和本文算法差别并不大.

LS 算法是光度立体视觉中常用的算法. 通过选择最佳的三个光源方向的图像来计算物体表面法向量以及表面深度, 具备方便快速的性能, 但是对于拥有噪声的数据鲁棒性很差. 从表 1 中可以看出, LS 算法虽然具备速度上的优势, 但是精度却明显低于本文算法. 另外, 本文算法以及 RANSAC 算法、LM 算法由于可以直接恢复整个观察矩阵, 因此可以方便地应用于光源未标定的情况, 而 LS 算法则只能用在光源已知的情况下.

### 3.1.2 镜面反射的影响

从上面的实验可以看出, 相对于其他算法来说, 本文所提出的算法对于镜面反射具有较好的抑制作用. 那么到底多大数量的镜面反射可以被有效地抑制呢? 本节实验将展示本文算法在这方面的性能. 本实验中, 使用了 40 幅不同光照情况下的图像, 通过控制镜面反射参数来调整被镜面反射影响到的像素点. 我们对比了本文算法以及其他算法相对于标准值的重建精度. 由于 RANSAC 算法参数调节繁琐以及计算速度过慢, 本次实验不对该算法进行比较. 图 3 和图 4 分别显示了在镜面反射区域增大的情况下的平均重建精度误差以及最大重建精度误差. 为了保证数据的可靠性, 对于每一个镜面反射参数, 均在 10 组随机生成的 40 个光源上进行了测试, 误差是这 10 组不同光照情况下的结果的平均值. 图中显示了统计的平均结果.

可以观察到本文方法在镜面反射比例小于 25% 的情况下, 均保证了误差的稳定性, 而其他几个算法的结果呈现出线性增长的趋势. 虽然在 25% 以后本文算法误差逐渐升高, 但是仍然低于其他几个算法的结果.

### 3.1.3 真实数据实验

在本实验中, 将采用真实数据对本文所提方法进行测试. 使用了 40 幅不同光源下的玩具图像. 光

源标定通过一个放置在场景中心的反射球来计算, 使用 Canon5D 摄像机进行拍摄. 示例图像如图 5 所示.

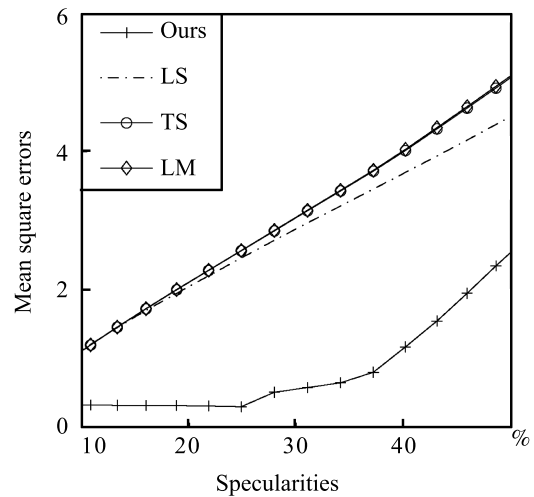


图 3 镜面反射区域的影响 (均方根误差)  
Fig. 3 Effect of increasing size of specular area (Mean square errors)

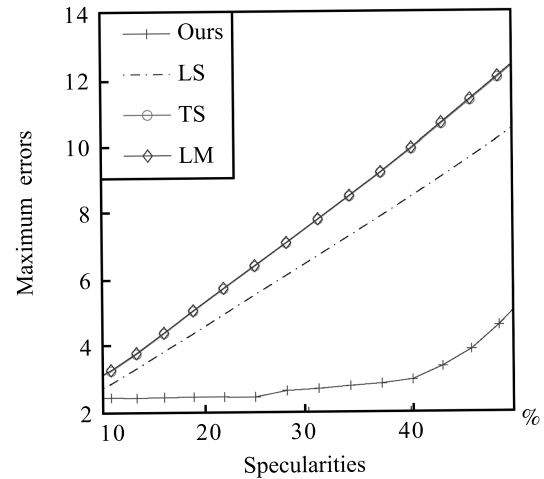


图 4 镜面反射区域的影响 (最大误差)  
Fig. 4 Effect of increasing size of specular area (Maximum errors)

真实数据中除了由镜面反射以及阴影带来的噪声外, 还有一些和采集设备相关的未知的噪声可以导致理想的朗伯模型被破坏. 由于没有标准值, 很难衡量算法表面重建的精度, 本实验将通过观察表面重建结果来定性的评价本文方法和其他的几种算法. 由于初始值的不理想, LM 方法运行失败, 没有得到输出结果.

本实验中统一使用了 0.01 作为阈值来检测阴影. 图 6 (a)~6 (f) 分别显示了使用本文方法和使用几种算法得到的表面重建结果. 本文算法和 LS 算法

都得到了较为理想的重建表面 (图 6 (a) 和图 6 (b)), 但是从颈部的放大视图 (图 6 (e) 和图 6 (f)) 来看, 由于有效地祛除了噪声的影响, 本文方法得到了较为光滑的重建表面, 而 LS 算法由于受到阴影以及镜面反射的影响, 重建表面则比较粗糙. TS 算法得到的结果拥有较大的噪声 (图 6 (c)). 对于 RANSAC 算法, 很难得到比较合适的阴影阈值来控制迭代次数, 在对系数计算以及图像线性化阶段分别使用 800 以及 200 个采样点的情况下, RANSAC 算法得到了误差非常大的表面 (图 6 (d)), 提高采样点数目以及使用适当的阴影阈值和朗伯反射区域的比例系数可以改善 RANSAC 算法的性能, 但是这需要比较多次数的手工调整. 相对来说, 本文算法以及 TS 算法、LS 算法都不需要太多的参数的调整. 在运算时

间上, LS 算法最优, RANSAC 算法耗时最长, 如表 2 所示.

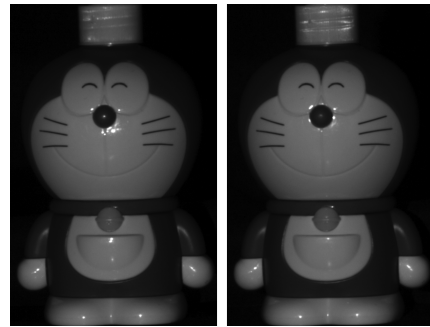


图 5 样本图像  
Fig. 5 Sample images

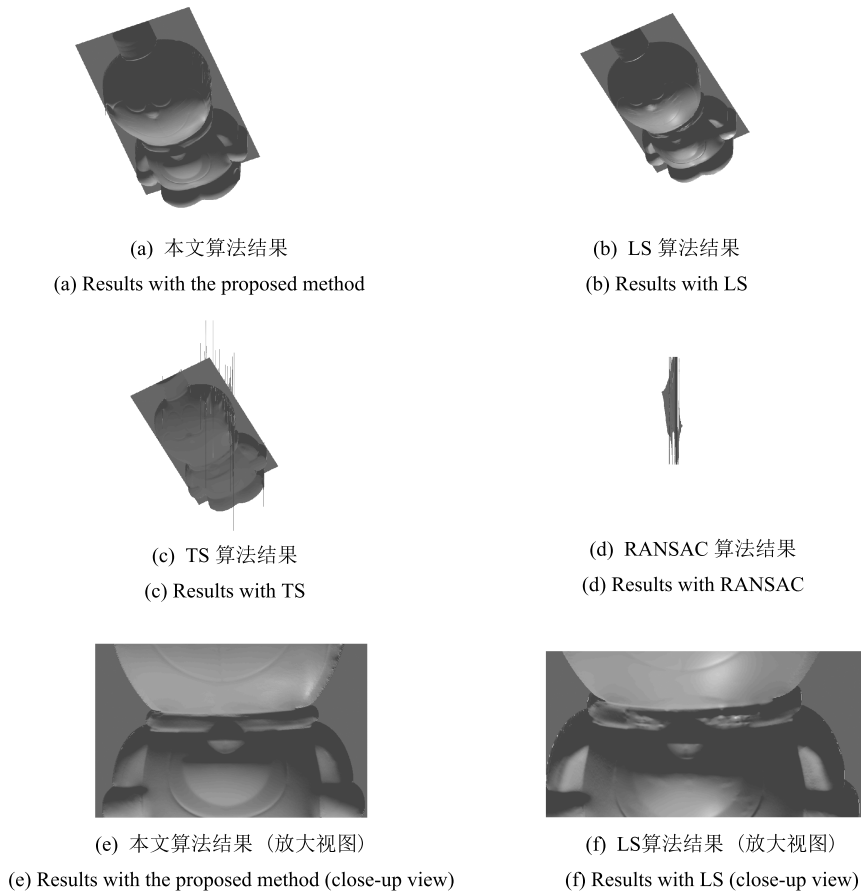


图 6 表面重建结果  
Fig. 6 Reconstruction results

表 2 计算时间  
Table 2 Computational time

算法	本文算法	LS 算法	TS 算法	RANSAC 算法
计算时间 (s)	694.7	90	173	8 689



## 4 结论

本文提出了一种基于先进的凸优化技术的光度立体视觉重建框架。相对于先前的一些光度立体视觉重建方法,该方法利用所有可用的信息,可以同时修复数据中的丢失和噪声数据,显示出了超高的计算效率以及对于大尺度稀疏噪声的鲁棒性。相对于一些常用的方法,本文提出的方法表现出了良好的综合性能,使用较少运算时间,可以保证较高的表面重建精度。不需要过多复杂的设置以及人工干预。另一方面,尽管本文只是针对光源已经标定的情况下进行的测试,但是该方法也可以很方便地扩展到光源未标定的情况下。目前的模型只是考虑了镜面反射以及阴影,图像中还会存在一些小的噪声。文献[30]显示低秩矩阵恢复的框架对于附加的小的高斯噪声也具有较强的稳定性。通过对于这项工作的分析,我们相信本文所提出的算法框架也可以处理固定光源以及环境光同时存在的情况。

## References

- Woodham R J. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 1980, **19**(1): 139–144
- Shashua A. Geometry and Photometry in 3D Visual Recognition [Ph. D. dissertation], Massachusetts Institute of Technology Cambridge, USA, 1992
- Belhumeur P N, Kriegman D J. What is the set of images of an object under all possible lighting conditions? In: Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 1996. 270–227
- Basri R, Jacobs D W. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(2): 218–233
- Georghiades A S, Kriegman D J, Belhumeur P N. From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(6): 643–660
- Hayakawa H. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America*, 1994, **11**(11): 3079–3089
- Jolliffe I T. *Principal Component Analysis*. New York: Springer-Verlag, 1986
- Belhumeur P N, Kriegman D J, Yuille A L. The bas-relief ambiguity. In: Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Juan, Puerto Rico: IEEE, 1997. 1060–1066
- Shi B X, Matsushita Y, Wei Y C, Xu C, Tan P. Self-calibrating photometric stereo. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 1118–1125
- Jacobs D. Linear fitting with missing data: applications to structure-from-motion and to characterizing intensity images. In: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition. San Juan, Puerto Rico: IEEE, 1997. 206–212
- Yuille A L, Snow D, Epstein R, Belhumeur P N. Determining generative models of objects under varying illumination: shape and albedo from multiple images using SVD and integrability. *International Journal of Computer Vision*, 1999, **35**(3): 203–222
- Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981, **24**(6): 381–395
- Julià C, Lumbreiras F, Sappa A D. A factorization-based approach to photometric stereo. *International Journal of Imaging Systems and Technology*, 2011, **21**(1): 115–119
- Guerreiro R F C, Aguiar P M Q. Estimation of rank deficient matrices from partial observations: two-step iterative algorithms. *EMMCVPR*. Lisbon, Portugal: Springer, 2003. 450–466
- Chen P. Optimization algorithms on subspaces: revisiting missing data problem in low-rank matrix. *International Journal of Computer Vision*, 2008, **80**(1): 125–142
- Mukaigawa Y, Ishii Y, Shakunaga T. Analysis of photometric factors based on photometric linearization. *Journal of the Optical Society of America*, 2007, **24**(10): 3326–3334
- Miyazaki D, Hara K, Ikeuchi K. Median photometric stereo as applied to the segonko tumulus and museum objects. *International Journal of Computer Vision*, 2010, **86**(2–3): 229–242
- Candès E, Li X D, Ma Y, Wright J. Robust principal component analysis? *Journal of the ACM*, 2011, **58**(3): 1–37
- Wu L, Ganesh A, Shi B X, Matsushita Y, Wang Y T, Ma Y. Robust photometric stereo via low-rank matrix completion and recovery. In: Proceedings of the 10th Asian Conference on Computer Vision. Queenstown New Zealand: Springer, 2010. 703–717
- Simchony T, Chellappa R, Shao M. Direct analytical methods for solving Poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, **12**(5): 435–446
- Chandrasekaran V, Sanghavi S, Parrilo P A, Willsky A S. Sparse and low-rank matrix decompositions. In: Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing. Monticello, IL: IEEE, 2009. 962–967
- Lin Z C, Chen M M, Wu L Q, Ma Y. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-rank Matrices. Technical Report UILU-ENG-09-2215, UIUC, USA, 2009
- Recht B, Fazel M, Parrilo P A. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 2010, **52**(3): 471–501

- 24 Candès E, Recht B. Exact matrix completion via convex optimization. *Foundation of Computational Mathematics*, 2008, **9**(6): 717–772
- 25 Candès E, Tao T. The power of convex relaxation: near-optimal matrix completion. *IEEE Transactions on Information Theory*, 2010, **56**(5): 2053–2080
- 26 Ganesh A, Lin Z C, Wright J, Wu L Q, Chen M M, Ma Y. Fast algorithms for recovering a corrupted low-rank matrix. In: Proceedings of the 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing. Aruba, Dutch Antilles: IEEE, 2009. 213–216
- 27 Toh K C, Yun S. An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems. *Pacific Journal of Optimization*, 2010, **6**: 615–640
- 28 Bertsekas D P. *Nonlinear Programming*. Belmont, MA: Athena Scientific, 2004
- 29 Cook R L, Torrance K E. A reflectance model for computer graphics. *ACM SIGGRAPH Computer Graphics*, 1981, **15**(3): 307–316
- 30 Zhou Z H, Li X D, Wright J, Candès E, Ma Y. Stable principal component pursuit. In: Proceedings of the 2010 IEEE International Symposium on Information Theory. Austin, USA: IEEE, 2010. 1518–1522



**吴 仑** 北京理工大学光电学院博士研究生。2006 年获北京理工大学光学工程专业硕士学位。主要研究方向为运动结构重建与数学优化方法。

E-mail: lun.wu@hotmail.com

(**WU Lun** Ph.D. candidate at the School of Optoelectronics, Beijing Institute of Technology. He received his

master degree in optical engineering from Beijing Institute of Technology in 2006. His research interest covers structure from motion and optimization method.)



**王涌天** 1982 年获天津大学精密仪器专业学士学位, 1986 年获英国 Reading 大学光学专业博士学位。目前为北京理工大学光学专业以及计算机应用专业教授, 教育部长江学者特聘教授。主要研究方向为光学设计, 光学仪器, 3D 显示, 图像处理, 虚拟现实和增强现实技术。本文通信作者。E-mail: wyt@bit.edu.cn

(**WANG Yong-Tian** Received his bachelor degree in precision instrumentation from Tianjin University, China, in 1982, and his Ph.D. degree in optics from University of Reading, England, in 1986. He is currently a professor of optics and applied computer science in Beijing Institute of Technology, and a Yangtze River Scholar appointed by Chinese Ministry of Education. His research interest covers optical design and CAD, optical instrumentation, 3D display, image processing, virtual reality (VR) and augmented reality (AR) technologies and applications. Corresponding author of this paper.)



**刘 越** 北京理工大学光电学院教授。1996 年获吉林工业大学通讯和信息系统专业硕士学位, 2000 年获吉林大学通讯和信息系统专业博士学位。主要研究方向为人际交互, 虚拟现实和增强现实, 摄像机跟踪定位, 3D 显示和摄像机标定。

E-mail: liuyue@bit.edu.cn

(**LIU Yue** Professor of optical engineering at the School of Optoelectronics, Beijing Institute of Technology. He received his Ph.D. degree in telecommunication and information system from Jilin University in 2000 and his master degree in telecommunication and electronic system from Jilin University of Technology in 1996. His research interest covers human computer interaction, virtual and augmented reality, accurate tracking of the pose of camera, 3D display system, and camera calibration.)