

基于各向异性全变分的GBR参数估计

魏昕钰, 杨奋林

吉首大学数学与统计学院, 湖南 吉首

收稿日期: 2024年4月29日; 录用日期: 2024年5月22日; 发布日期: 2024年5月31日

摘要

GBR参数估计是求解未标定光度立体技术的一般方法。本文针对GBR参数估计中模糊深度不连续的问题, 引入具有保持深度不连续的各向异性扩散全变分做正则建立新的变分模型, 并运用具有超线性收敛的拟牛顿法对模型进行求解。实验表明各向异性扩散全变分方法能有效降低法线误差, 提高深度图不连续, 且拟牛顿法求解该模型所用时间远低于EM方法和DM方法。

关键词

未标定光度立体, GBR参数, 各向异性全变分, 拟牛顿方法

Estimation of GBR Parameters Based on Anisotropic Total Variation

Xinyu Wei, Fenlin Yang

College of Mathematics and Statistics, Jishou University, Jishou Hunan

Received: Apr. 29th, 2024; accepted: May 22nd, 2024; published: May 31st, 2024

Abstract

GBR parameter estimation is a general method for solving uncalibrated photometric stereoscopic technique. In order to solve the problem of fuzzy depth discontinuity in GBR parameter estimation, a new variational model is established by introducing the anisotropic diffusion total variational model with depth discontinuity, and the quasi-Newton method with superlinear convergence is used to solve the model. Experiments show that the anisotropic diffusion total variational method can effectively reduce the normal error and improve the discontinuity of depth map, and the time of quasi-Newton method to solve the model is much lower than that of EM method and DM method.

Keywords

Uncalibrated Photometric Stereo, GBR Parameters, Anisotropic Total Variation, Quasi-Newton Method

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

三维重建是指由二维图像得到三维场景的非接触式的重建过程, 有很高的实用价值, 如应用于医学影像、文物数字化和虚拟现实等[1]。光度立体三维重建是利用同一视角具有不同光度信息的图像进行物体表面恢复的方法, 光照信息未知的光度立体问题称为未标定光度立体问题。若图像通过正交投影获得, Hayakawa [2]表明可以在这一问题中应用奇异值分解法(Singular value decomposition, SVD), 再增加表面可积性约束条件, 则模糊性减小为通用浅浮雕(Generalized bas-relief, GBR)模糊性[3]。在未知光源条件下, 使用正交投影会产生 GBR 问题, 使用透视投影不会产生相同问题[4], 因此透视投影模型可以解决 GBR 模糊性问题, 但是该模型比远光源的正交投影模型复杂得多, 且计算困难。因此, GBR 参数估计是未标定光度立体问题的一般解决方向。

Alldrin 等人[5]通过最小化反照率分布的熵值求解 GBR 参数。Favaro 和 Papadimitri [6]利用局部漫反射最大值解决 GBR 歧义。Yvain Quéau [7]通过法向量场的全变分(Total Variation, TV)最小化来估计 GBR 参数, 提高了重建精度。

GBR 参数估计是解决未标定光度立体模糊性问题的关键方向, 其仍是一个有待研究的问题。TV 最小化广泛应用于图像去噪领域[8], 已取得很大进展, 其特点是能很好的保持边缘特性, 所以可以应用到三维重建领域。本研究借鉴 Yvain Quéau 等人的思想, 在 GBR 模糊性的基础上引入深度函数 u 的 TV 最小化, 利用深度正则化估计 GBR 参数, 从而提出了基于各向异性全变分(Anisotropic Total Variation, ATV)的三维重建模型, 提高了重建精度。

2. GBR 变换

在正交投影下的未标定光度立体问题中, 光源信息未知, 此时需要同时估计法向量和光源信息。这是一个不适定问题, 这一不适定问题可以减小为 GBR 模糊性。

GBR 模糊性是指当从某一特定角度观察物体时, 不能对物体进行准确的深度估计。当使用正交投影观察一个具有朗伯反射率的物体时, 其深度的确定会存在歧义。对于物体和通过 GBR 变换后的物体, 选择不同的光照角度, 这两种不同的物体在某些视角下可以得到完全相同的图片。

记 I 为图像灰度矩阵, 矩阵中每一列为在不同光照信息下得到的图像灰度值, $n(x, y) = (n_x, n_y, n_z)^T$ 是像素点 (x, y) 的表面法向量, ρ 是将像素点的反照率按列排列的向量生成的对角矩阵, N 是像素点的归一化法向量排成的矩阵, L^i 是第 i 个光源向量, $i = 1, 2, \dots, m$, 其中 $m \geq 3$ 是光源数量, L 是这些光源排列成的光照矩阵, 令 $M = \rho N$, 则 I 可表示为

$$I = ML. \quad (1)$$

将 I 应用 SVD 方法进行主成分分析, 仅取前三个最大的奇异值及其对应的特征向量, 这样可以保留

矩阵的主特征, 则图像矩阵 I 可表示为

$$I = UDV^T, \quad (2)$$

其中, $U \in R^{R \times 3}$, $D \in R^{3 \times 3}$, $V \in R^{m \times 3}$, R 为像素点总数量。记 P 和 Q 是满足 $PQ^T = D$ 的两个 3×3 矩阵, 则存在无穷多个这样的 (P, Q) 矩阵, 此时得到法向量值和实际值相差的一个 3×3 参数的可逆线性变换 A , 即:

$$I = (UP^T)(QV^T) = \hat{M}\hat{L} = \hat{M}^1(AA^{-1})\hat{L}^1 = (\hat{M}^1A)(A^{-1}\hat{L}^1), \quad (3)$$

其中 (\hat{M}^1, \hat{L}^1) 是 \hat{M} 和 \hat{L} 的一组解。假设表面连续, 即在此基础上增加表面可积性约束, 则有 $A = G$, $A^{-1} = G^{-1}$, 即可以将此线性变换模糊性减少为 GBR 模糊性。GBR 矩阵及其逆矩阵为:

$$A = G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \varrho & \lambda \end{pmatrix}, A^{-1} = G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\mu}{\lambda} & -\frac{\varrho}{\lambda} & \frac{1}{\lambda} \end{pmatrix}. \quad (4)$$

3. 基于各向异性 TV 的 GBR 参数估计

3.1. 基于各向异性 TV 的 GBR 参数估计模型

基于 L1 范数的各向异性全变分定义为:

$$ATV(u) = \int_{\Omega} (|\nabla_x u| + |\nabla_y u|) dX. \quad (5)$$

Yvain Quéau 分别观察了 μ, ϱ 和 λ 对 $TV(u)$ 的影响, 当观察其中一个参数的影响时, 固定其他两个参数为实际值, 发现 μ 和 ϱ 越接近实际值时, $TV(u)$ 的值越小。ATV 的特性是可以更好地保留物体的边缘结构信息, 因此本文考虑通过深度场的 ATV 最小化估计 μ 和 ϱ 。

GBR 变换影响深度及其梯度的变化, 假设深度函数 $u(x, y)$ 分段两次可微, 则几乎处处满足可积性约束 $u_{xy} = u_{yx}$ 。Yuille 和 Snow 提出了在最小二乘的意义上 M 场满足可积性约束且 M 和 L 满足朗伯体定律的一个封闭解[9], 称为 (M^0, L^0) , 取其作为初始解, 用 $u^0(x, y)$ 表示这一方法估计的法向量场 N^0 积分得到的深度场, 其中, $M^0 = (M_1^0, M_2^0, M_3^0)$, $N^0 = (N_1^0, N_2^0, N_3^0)$ 。

在 GBR 变换的作用下, M^0 经变换后得到的 M 场为:

$$\bar{M} = (M_1^0 + \mu M_3^0, M_2^0 + \varrho M_3^0, \lambda M_3^0). \quad (6)$$

对应的深度 \bar{u} 的梯度场为:

$$\nabla \bar{u} = -\frac{1}{\lambda} \left(\frac{M_1^0}{M_3^0} + \mu, \frac{M_2^0}{M_3^0} + \varrho \right)^T. \quad (7)$$

令 $U^0 = (x, y, u^0(x, y))^T$, $g(\mu, \varrho, \lambda) = \frac{1}{\lambda}(-\mu, -\varrho, 1)$, $g(\mu, \varrho, \lambda)$ 对应 G^{-1} 矩阵的第三行元素, 则通过简单的微积分得到 GBR 变换下的深度函数:

$$\bar{u} = g(\mu, \varrho, \lambda)U^0. \quad (8)$$

GBR 参数的估计可以描述为(9)式:

$$\hat{u} = \arg \min ATV(\bar{u}),$$

$$\|\bar{M}\bar{L} - I\|^2 = \left\| \rho L \frac{(-\partial_x \bar{u}, -\partial_y \bar{u}, 1)}{\sqrt{\nabla \bar{u}^2 + 1}} \bar{L} - I \right\|^2 = \min. \quad (9)$$

由于 $\|M^0 L^0 - I\|^2 = \|(M^0 G(\hat{\mu}, \hat{\vartheta}, \hat{\lambda}))(G^{-1}(\hat{\mu}, \hat{\vartheta}, \hat{\lambda})L^0) - I\|^2 = \|\overline{ML} - I\|^2$, 因此, 已知 $\left\| \rho^0 \frac{(-\partial_x u^0, -\partial_y u^0, 1)}{\sqrt{\nabla u^0{}^2 + 1}} L^0 - I \right\|^2 = \min$, 则有 $\left\| \overline{\rho} \frac{(-\partial_x \bar{u}, -\partial_y \bar{u}, 1)}{\sqrt{\nabla \bar{u}^2 + 1}} \bar{L} - I \right\|^2 = \min$ 。

由于 λ 的符号与物体的凸性直接相关, 而凹凸模糊性不能在没有物体附加信息的情况下求解, 因此通常假设 $\lambda > 0$ 。由此, (9)的解可由式(10)和(11)给出:

$$\hat{u} = g(\hat{\mu}, \hat{\vartheta}, \hat{\lambda})U^0, \tag{10}$$

其中,

$$(\hat{\mu}, \hat{\vartheta}, \hat{\lambda}) = \arg \min_{\substack{\mu, \vartheta \\ \lambda > 0}} ATV(g(\mu, \vartheta, \lambda)U^0). \tag{11}$$

为了准确估计 GBR 参数, 关注的是式(11)。由(7)式可知, μ, ϑ 的估计独立于 λ 的估计, 同时可知, 最小化 $ATV(u)$ 导致 λ 趋于无穷大, 因此, 仅利用深度函数的 ATV 最小化估计 μ 和 ϑ 的值是可行的, 此时, 先取 λ 为固定值 1。通过下式估计 μ 和 ϑ :

$$(\hat{\mu}, \hat{\vartheta}) = \arg \min_{\mu, \vartheta} ATV(g(\mu, \vartheta, 1)U^0). \tag{12}$$

利用反照率熵最小化的方法估计 GBR 变换中的参数 λ , 对 λ 可能值进行由粗到精的估计, 将式(12)和熵最小化模型[5]相结合, 则有以下模型:

$$\begin{aligned} (\hat{\mu}, \hat{\vartheta}) &= \arg \min_{\mu, \vartheta} ATV(g(\mu, \vartheta, 1)U^0), \\ \hat{\lambda} &= \arg \min_{\lambda} f\left(\|M^0 G(\hat{\mu}, \hat{\vartheta}, \lambda)\|\right), \end{aligned} \tag{13}$$

其中, $f\left(\|M^0 G(\hat{\mu}, \hat{\vartheta}, \lambda)\|\right)$ 表示变换后反照率分布的熵值。

3.2. 拟牛顿法求解模型

深度函数 $u(x, y)$ 的 ATV 最小化可以减少未标定光度立体的模糊性。在这一过程中, 深度函数的 ATV 最小化只依赖于 GBR 参数, 由此产生的问题是凸的, 可以使用拟牛顿的凸优化方法求解式(13), 目标函数为:

$$E(\mu, \vartheta) = ATV(\bar{u}) = \int_{\Omega} (|\nabla_x \bar{u}| + |\nabla_y \bar{u}|) dX. \tag{14}$$

$E(\mu, \vartheta)$ 分别对 μ 和 ϑ 求偏导, 得到偏微分方程:

$$\begin{aligned} \partial_{\mu} E &= \int_{\Omega} \frac{M_1^0 + \mu M_3^0}{|M_1^0 + \mu M_3^0|} dX, \\ \partial_{\vartheta} E &= \int_{\Omega} \frac{M_2^0 + \vartheta M_3^0}{|M_2^0 + \vartheta M_3^0|} dX. \end{aligned} \tag{15}$$

已知目标函数及其偏微分方程, 可用拟牛顿法快速估计 μ 和 ϑ , 为加快估计速度, 选择 Tikhonov 正则化[10]的解 $\hat{\mu}_0$ 和 $\hat{\vartheta}_0$ 作为一组初始解, 其中, $\hat{\mu}_0$ 和 $\hat{\vartheta}_0$ 分别代表 $-\frac{M_1^0}{M_3^0}$ 和 $-\frac{M_2^0}{M_3^0}$ 的平均值。

以下为具体算法: $[\mu, \vartheta] \leftarrow (M^0, L^0)$

- 1: 给定初始值 $x_0 = [\mu_0, \vartheta_0] \leftarrow \left[-\text{mean}(M_1^0 ./ M_3^0), -\text{mean}(M_2^0 ./ M_3^0) \right]$ 和容许误差 $\varepsilon > 0$, $k = 0$, $H_0 = I$, 计算梯度 g_k ;
- 2: 若 $\|g_k\| < \varepsilon$, 则迭代停止;
- 3: 计算迭代方向 $d_k = -H_k g_k$;
- 4: 从点 x_0 出发, 沿着 d_k 做一维搜索求步长 α_k , 令 $x_{k+1} = x_k + \alpha_k d_k$;
- 5: 计算梯度 g_{k+1} , 令 $s_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$;
- 6: 若 $s_k^T y_k > 0$, 则令 $H_{k+1} = \left(I - \frac{\Delta x \Delta g^T}{\Delta g^T \Delta x} \right) H_k \left(I - \frac{\Delta g \Delta x^T}{\Delta g^T \Delta x} \right) + \frac{\Delta x \Delta x^T}{\Delta g^T \Delta x}$; 否则, 令 $H_{k+1} = H_k$;
- 7: 令 $k = k + 1$, 转步骤 2。

4. 实验结果与分析

漫反射极大值(Diffuse Maxima, DM)方法求解 GBR 参数是把最亮的像素看作朝向光源的像素, 这是已知的有效方法; 熵最小化(Entropy Minimization, EM)方法通过最小化反照率的熵来估计 GBR 参数, 考虑的是全局属性, 与本文方法有相同之处。因此本文所提算法的结果将与 EM、DM 方法进行对比。

为定量、定性评估构造方法的有效性和鲁棒性, 选取 4 个图像数据集进行了实验, 其中, Owl 数据集由 12 张分辨率为 512×340 的图像组成; Frog、Pig 和 Scholar 数据集均由 20 张灰度图像组成, 分辨率分别为 690×590 、 760×550 和 1070×790 。这些物体几乎是朗伯体物体。下面给出每一个数据集在不同光照条件下的三幅图像, 如图 1 所示。

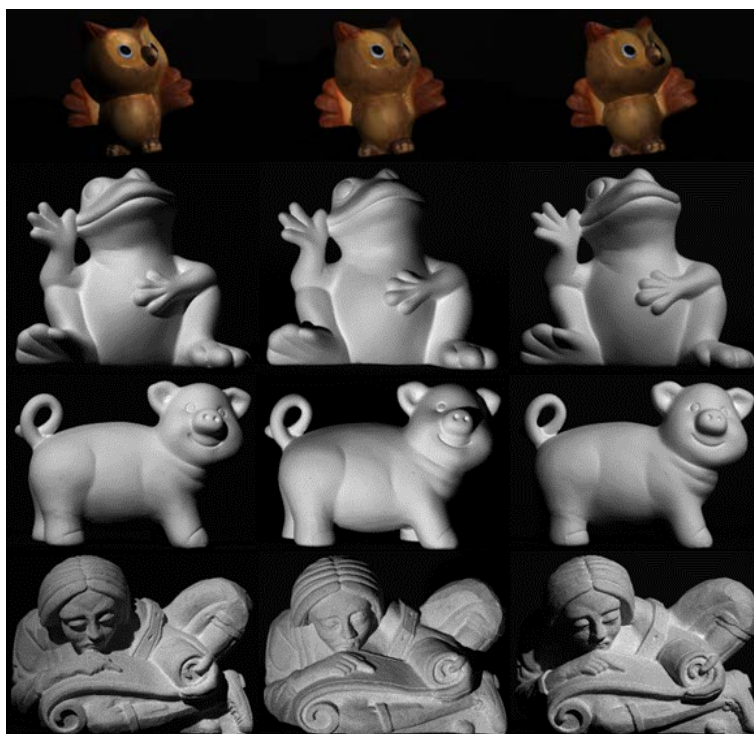


Figure 1. Images of objects under different lightsource directions

图 1. 不同光源方向下的物体图像

为定量分析本文方法的实验结果, 将标定光度立体三维重建的结果作为真实值, 将实验结果与真实值相比较, 引入法线的平均角度误差(mean angular error, MAE)作为一个评价指标, 定义如下:

$$\text{MAE} = \frac{1}{R} \cos^{-1} \left(N_r \cdot \hat{N}_r \right), \quad (16)$$

其中 R 是像素点总数量, r 是图片上的像素点。使用法线角度误差小于 5° 的像素点百分比和运行时间作为另两个评价指标。

为了客观比较, 计算了 EM、DM 和本文方法估计的法线的 MAE、误差小于 5° 的像素点百分比和 CPU 运行时间, 如表 1 所示。

Table 1. Reconstruction error and calculation time of each method

表 1. 各方法的重建误差和计算时间

数据集	重建方法	MAE(°)	5°(%)	CPU(s)	数据集	重建方法	MAE(°)	5°(%)	CPU(s)
Owl	EM	15.04	3.7	18.51	Pig	EM	8.73	5.48	57.24
	DM	11.35	0.96	3.10		DM	9.99	3.06	18.42
	ATV-u/EM	8.27	21.98	2.55		ATV-u/EM	5.82	27.74	8.79
Frog	EM	4.91	47.68	56.65	Scholar	EM	5.01	38.01	147.30
	DM	3.00	85.04	14.94		DM	4.69	47.00	89.88
	ATV-u/EM	3.83	64.86	8.61		ATV-u/EM	4.77	51.70	23.53

表 1 列出了三种方法估计的法线的误差数据和 CPU 运行时间。可以看到, Owl 数据集和 Pig 数据集通过本文方法得到的法线的平均角度误差低于 EM 方法和 DM 方法的误差, 本文方法的结果取得了更高的精度; Owl 数据集中, DM 方法得到的误差小于 5° 的像素点百分比为 0.96%, 本文方法得到的误差小于 5° 的像素点百分比为 21.98%, 是前者的近 23 倍; Scholar 数据集中, EM 方法的运行时间为 147.30s, 本文方法的运行时间为 23.53s, 是前者的近六分之一。因此, 无论从法线的 MAE 还是误差小于 5° 的像素点百分比上看, 本文方法与其他两种方法相比都有较好的估计结果, 且计算时间更短。ATV 具有比熵更好的微分性质, 故本文构造的方法的运行时间大大短于 EM 方法。EM 方法利用全局优化来求解三个 GBR 参数, DM 方法通过确定图像局部亮度最大值区域求解 GBR 参数, 误差均较小, 但当图像分辨率高时, 都会导致时间较长, 而本文提出的方法可以缩短运算时间。

以下展示 Owl 数据集的真实反照率图和在三方法下生成的反照率图, 见图 2, 由图可知本文方法生成的反照率图更接近反照率实际值, 与其他两种方法相比误差更小。

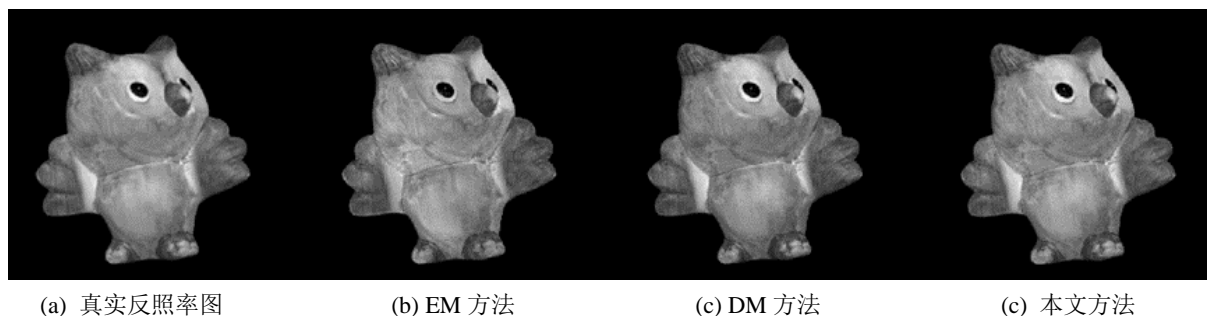


Figure 2. The real albedo map and the albedo map generated by the three methods

图 2. 真实反照率图以及三种方法下生成的反照率图

将本文方法与 EM、DM 方法进行定性比较, 结果如图 3 所示。图 3 展示了数据集中四个物体的真实法向图、三种方法得到的法向图以及生成法向图和真实法向图之间的误差图。法线角度误差低的地方表

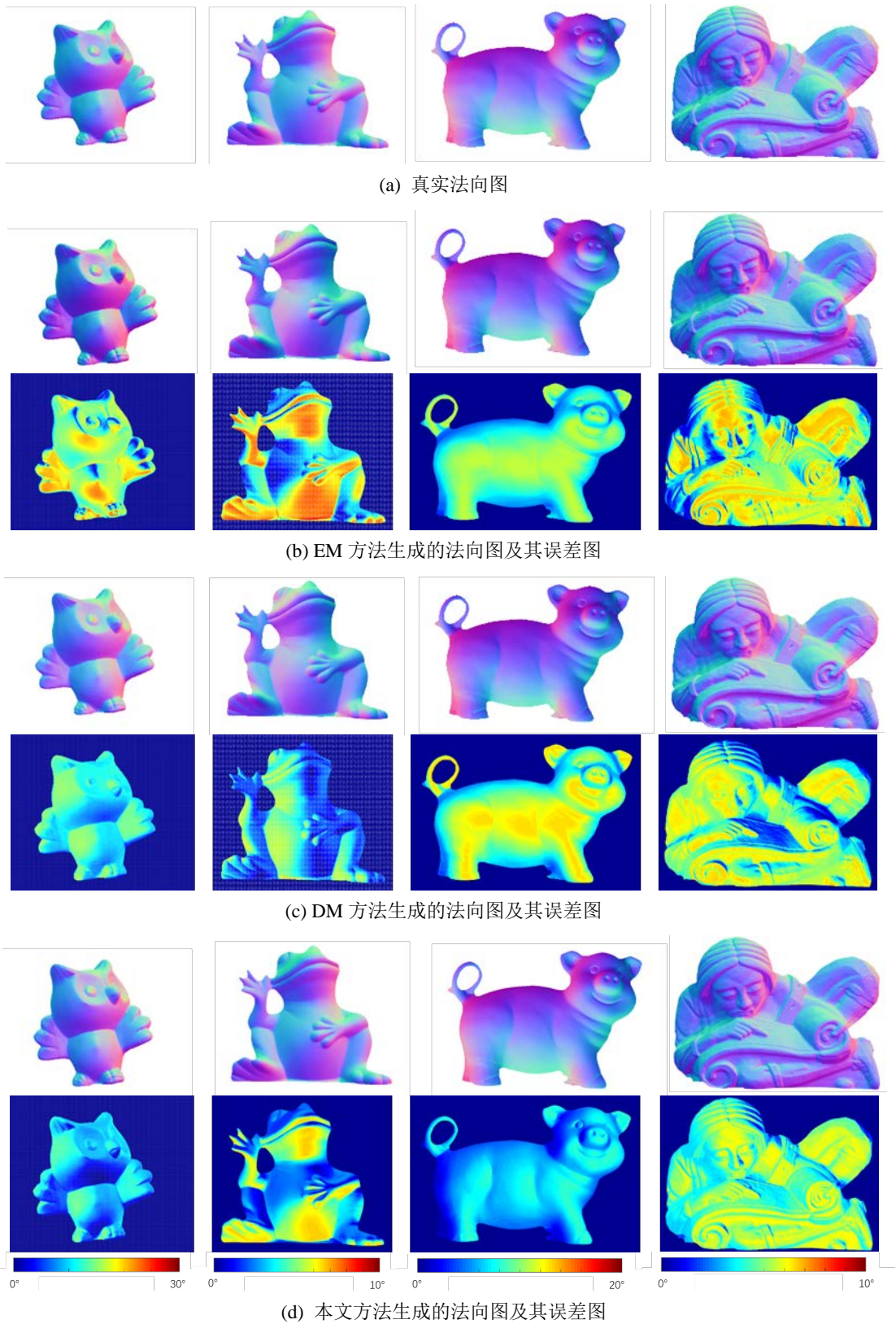


Figure 3. The real normal graph, the normal graph generated by three methods and the error graph
图 3. 真实法向图、三种方法生成的法向图及误差图

现为蓝色, 误差高的地方表现为红色。通过定性比较可知, 本文方法估计的法向量误差较小, 能有效提高表面重建的精度。

在图 4 中, 展示了 Pig 数据集的三维重建结果在三个不同视角下观察到的图片, 其中深度是由本文方法估计的法线积分后得到的。

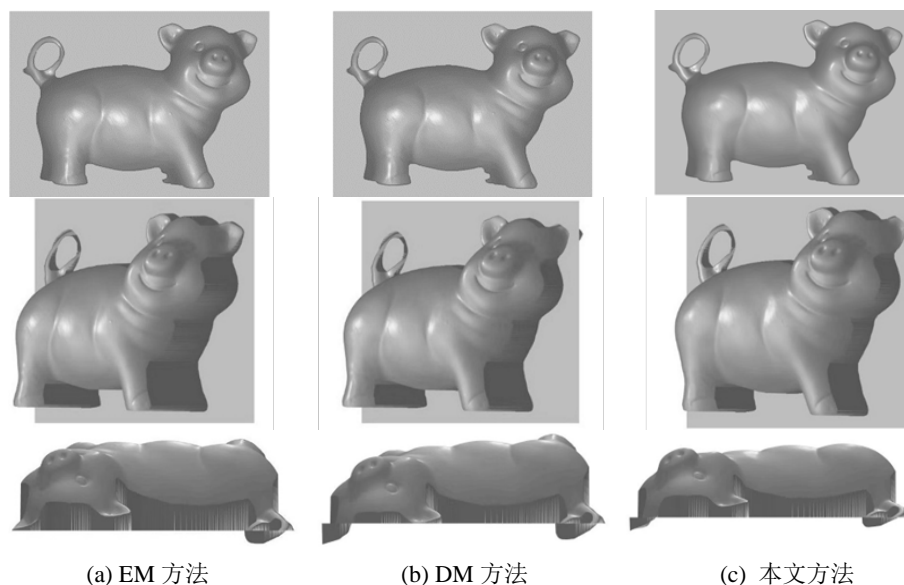


Figure 4. 3D reconstruction results of each method are in the front, right front and right top view of the picture

图 4. 各方法的三维重建结果在正面、右前、正上方视角下的图片

从视觉效果上看, 相较于 EM、DM 方法, 本文方法的三维重建实验结果较好地恢复了物体实际的三维形状, 深度的凸出效果明显, 更好地恢复了物体的表面细节, 重建精度高, 表现出了良好的综合性能。

5. 结论

本文提出了一种基于各向异性全变分的三维重建方法, 该方法在 GBR 模糊性的基础上, 利用 ATV 最小化估计 GBR 参数 μ 和 ϱ , 通过实验证明, 提出的方法取得了较高的重建精度, 当图像分辨率更高时, 在保证重建精度的同时有效缩短了计算时间, 提高了计算效率。这表明 ATV 在解决未标定光度立体三维重建问题中是适用的, 其可以有效降低 GBR 模糊性, 提高计算效率。

基金项目

吉首大学科研项目(Jdy23029)。

参考文献

- [1] Yang, S., Hou, M., Shaker, A. and Li, S. (2021) Modeling and Processing of Smart Point Clouds of Cultural Relics with Complex Geometries. *ISPRS International Journal of Geo-Information*, **10**, Article 617. <https://doi.org/10.3390/ijgi10090617>
- [2] Hayakawa, H. (1994) Photometric Stereo under a Light Source with Arbitrary Motion. *Journal of the Optical Society of America A*, **11**, 3079-3089. <https://doi.org/10.1364/JOSAA.11.003079>
- [3] Belhumeur, P.N., Kriegman, D.J., and Yuille, A.L. (1999) The Bas-Relief Ambiguity. *International Journal of Computer Vision*, **35**, 33-44. <https://doi.org/10.1023/A:1008154927611>
- [4] Papadhimitri, T., and Favaro, P. (2013) A New Perspective on Uncalibrated Photometric Stereo. *Proceedings of the*

-
- IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 23-28 June 2013, 1474-1481.
<https://doi.org/10.1109/CVPR.2013.194>
- [5] Alldrin, N.G., Mallick, S.P., and Kriegman, D.J. (2007) Resolving the Generalized Bas-Relief Ambiguity by Entropy Minimization. 2007 *IEEE Conference on Computer Vision and Pattern Recognition*, 1-7.
<https://doi.org/10.1109/CVPR.2007.383208>
- [6] Favaro, P., and Papadimitri, T. (2012) A Closed-Form Solution to Uncalibrated Photometric Stereo via Diffuse Maxima. 2012 *IEEE Conference on Computer Vision and Pattern Recognition*, 821-828.
<https://doi.org/10.1109/CVPR.2012.6247754>
- [7] Quéau, Y., Lauze, F., and Durou, J.D. (2013) Solving the Uncalibrated Photometric Stereo Problem Using Total Variation. *Journal of Mathematical Imaging and Vision*, **52**, 87-107. <https://doi.org/10.1007/s10851-014-0512-5>
- [8] Guo, J., and Chen, Q. (2021) Image Denoising Based on Nonconvex Anisotropic Total-Variation Regularization. *Signal Processing*, **186**, Article ID: 108124. <https://doi.org/10.1016/j.sigpro.2021.108124>
- [9] Yuille, A., and Snow, D. (1997) Shape and Albedo from Multiple Images Using Integrability. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 158-164.
<https://doi.org/10.1109/CVPR.1997.609314>
- [10] Gerth, D. (2021) A New Interpretation of (Tikhonov) Regularization. *Inverse Problems*, **37**, Article ID: 064002.
<https://doi.org/10.1088/1361-6420/abfb4d>