

# 双边法向积分

## 摘要

本课程研究报告针对三维表面重建中的法向积分问题进行了研究，重点关注如何有效地处理具有不连续性的表面。报告复现了 Cao 等人提出的一种基于双边加权策略的法向积分算法。该算法的核心思想是引入半光滑表面假设，即假设真实表面在每个像素的邻域内在水平和垂直方向上至多只有一个方向存在不连续。基于此假设，该论文构建了一个双边加权泛函来描述表面重建的能量函数，并利用迭代重加权最小二乘法 (IRLS) 进行优化求解。该方法能够自适应地调整不同区域的权重，从而在保持不连续性的同时抑制噪声的影响。此外，该论文还提出了一种统一的正交和透视投影法向积分框架，使得所提算法能够适用于不同的成像模型。本报告详细阐述了该算法的原理、实现细节以及实验结果，并在法向图数据集 DiLiGenT 上验证了该算法的有效性。

**关键词：**法向积分；保持不连续性；半光滑表面；单侧可微性

## 1 引言

三维表面重建是计算机视觉领域一项基础且重要的研究课题，其目标是从二维图像或其它形式的数据中恢复物体的三维几何信息。在虚拟现实、增强现实、机器人技术、工业检测、医学影像、文物保护等诸多领域，精确的三维重建技术都扮演着至关重要的角色。法向积分作为三维重建技术中的一个关键分支，通过对物体表面的法向矢量场进行积分来推导出其深度信息，从而实现三维形状的重建。这种方法广泛应用于光度立体视觉、结构光扫描等主动三维成像技术中。

然而，传统的法向积分方法通常假设待重建表面是光滑且连续的，即表面上任意一点的法向量在邻域内都是连续变化的。这一假设在许多实际应用场景中难以满足。例如，在工业零部件、人造物体以及具有复杂纹理的自然场景中，物体表面往往存在着尖锐的边缘、棱角或由于遮挡导致的深度不连续。在这些情况下，传统的基于光滑假设的法向积分方法会在不连续区域产生较大的重建误差，导致重建结果出现边缘模糊、细节丢失、几何结构扭曲等问题，严重影响了三维重建的精度和可靠性。

尽管近年来研究者们已经提出了一些改进的法向积分方法来处理不连续性，例如基于鲁棒估计的方法和基于加权的方法，但这些方法仍然存在一些局限性。基于鲁棒估计的方法通常难以准确地识别所有类型的不连续点，而基于加权的方法的性能又很大程度上依赖于权重函数的设计和参数的选择。

Cao 等人在其论文《Bilateral Normal Integration》[2] 中提出了一种新颖的基于双边加权策略的法向积分方法。该方法引入了半光滑表面假设，并构建了一个双边加权泛函来描述

表面重建的能量函数。这种方法在理论上能够更好地处理不连续性，并在实验中展现出优于传统方法的效果。

通过对 Cao 等人提出的双边加权法向积分方法的复现和研究，可以深入理解该方法的原理、优势和局限性，为进一步改进和发展法向积分技术提供参考和借鉴。同时，这项工作也有助于加深对半光滑表面模型以及双边滤波思想在三维重建中应用的理解。该方法能够更准确地重建具有不连续性的表面，这对于许多实际应用具有重要意义。例如，在工业检测中，可以更精确地测量零件的尺寸和形状；在逆向工程中，可以更忠实地复制原始物体的几何模型；在文物保护中，可以更完整地记录和还原文物的三维形态。

## 2 相关工作

本章将回顾与法向积分相关的研究工作，重点关注保持不连续性的法向积分方法。

### 2.1 统一法向积分方程

自法向积分问题出现以来，大多数方法都基于以下偏微分方程 (PDE) 估计深度图：

$$\partial_u z - p = 0 \quad \text{和} \quad \partial_v z - q = 0, \quad (1)$$

其中  $[p, q]^T = [-n_x/n_z, -n_y/n_z]^T$  是从法向图计算出的梯度场。通过引入对数深度图  $\tilde{z}$ ，可以将透视情况下的偏微分方程统一为与公式 (1) 相同的形式，只是梯度场  $[p, q]^T$  不同 [3, 4, 8]。因此，法向积分问题也被称为从梯度场恢复形状/高度/深度 [5, 6, 10]。

公式 (1) 是从法向量应平行于两个切向量的叉积这一约束推导出来的： $\mathbf{n} \parallel \frac{\partial z}{\partial u} \times \frac{\partial z}{\partial v}$ 。另外，Zhu 和 Smith [14] 从正交约束公式 (1) 推导出偏微分方程，并发现它有利于数值稳定性。然而，Zhu 和 Smith [14] 在正交和透视情况下推导出不一致的偏微分方程。在透视情况下，Zhu 和 Smith [14] 使用奇异值分解求解一个齐次系统，这比在正交情况下求解一个非齐次系统更耗时。

本文（指原论文）结合了两种推导的优点。像 Zhu 和 Smith [14] 一样，为了数值稳定性，从正交约束推导出偏微分方程。像 [3, 4, 8] 一样，引入对数深度图来统一公式。通过这种方式，可以在数值更稳定的同时，以相同的方式求解两种情况下的法向积分问题。

### 2.2 保持不连续性的表面恢复

现在讨论两种不连续性的保持策略：基于鲁棒估计器的方法和加权方法。

由于公式 (1) 的残差在不连续点处变大，基于鲁棒估计器的方法将鲁棒函数  $\rho$  应用于数据项：

$$\min \iint_{\Omega} \rho \left( \frac{\partial z}{\partial u} - p \right) + \rho \left( \frac{\partial z}{\partial v} - q \right) dudv. \quad (2)$$

合理设计的、能够抑制大残差影响的  $\rho$  函数有望保持不连续性。已经研究了 Lorentzian 函数 [3]、总变差和三重稀疏 [1]。

加权方法不是应用鲁棒函数，而是为偏微分方程的二次残差分配权重，以消除不连续性的影响：

$$\min_z \iint_{\Omega} w_u \cdot \left( \frac{\partial z}{\partial u} - p \right)^2 + w_v \cdot \left( \frac{\partial z}{\partial v} - q \right)^2 dudv. \quad (3)$$

如果为不连续点处的数据项分配合适的较小权重，则有望保持不连续性。

一类加权方法在优化之前将检测不连续点作为预处理步骤。这些工作之间的主要区别在于用于不连续性检测的线索。Karacali 和 Snyder [7] 基于公式 (1) 的残差检测不连续性。Wu 和 Tang [12] 使用期望最大化算法从法向图估计不连续图。Wang 等人 [11] 使用光度立体图像和法向图检测二元不连续图。Xie 等人 [13] 从法向图手工设计特征。一次性检测可能很脆弱，因为没有方案可以在优化之后纠正可能错误的检测。

一种更有效的加权方法类型是迭代地更新权重。Alpha-surface 方法首先从积分域创建最小生成树，然后迭代地将视为连续的边添加到生成树中。各向异性扩散将扩散张量应用于梯度场。Quéau 等人 [9] 将扩散张量设计为深度的函数，并在优化期间迭代更新扩散张量。Mumford-Shah 积分器通过假设不连续性是域中的短曲线来绕过检测，从而联合优化权重和深度。

本文（指原论文）的方法可以归类为加权方法。与以前的方法不同，该方法假设了一个半光滑的目标表面，并相对地加权每个点两侧的单侧可微性。在优化期间迭代更新权重，这与在优化之前确定权重的方法 [11–13] 不同。此外，与大多数在没有深度信息的情况下确定权重的方法不同，该方法的权重是未知深度的函数，因此在优化期间自适应地确定。

### 3 本文方法

本节将详细介绍所提出的基于双边加权策略的法向积分方法。该方法的目标是在给定法向图和相机投影参数的情况下，估计表面上每个点的深度或高度。

#### 3.1 统一法向积分方程

设  $\mathbf{p} = [x, y, z]^T \in \mathbb{R}^3$  为三维空间中的一个表面点， $\mathbf{n}(\mathbf{p}) = [n_x, n_y, n_z]^T \in \mathbb{S}^2 \subset \mathbb{R}^3$  为该点处的单位表面法向量。当通过相机观察表面时，表面点及其法向量被投影到图像平面上，图像平面坐标为  $\mathbf{u} = [u, v]^T \in \mathbb{R}^2$ 。因此，可以将表面及其法向图参数化为向量值函数  $\mathbf{p}(\mathbf{u}) = [x(\mathbf{u}), y(\mathbf{u}), z(\mathbf{u})]^T$  和  $\mathbf{n}(\mathbf{u}) = [n_x(\mathbf{u}), n_y(\mathbf{u}), n_z(\mathbf{u})]^T$ 。根据定义，法向量  $\mathbf{n}(\mathbf{u})$  与点  $\mathbf{p}(\mathbf{u})$  处的表面切平面正交。因此， $\mathbf{n}(\mathbf{u})$  与切平面上的两个切向量正交：

$$\mathbf{n}^T \frac{\partial \mathbf{p}}{\partial u} = 0 \quad \text{和} \quad \mathbf{n}^T \frac{\partial \mathbf{p}}{\partial v} = 0. \quad (4)$$

这里， $\frac{\partial}{\partial u}$  和  $\frac{\partial}{\partial v}$  分别表示关于  $u$  和  $v$  的偏导数，为了简洁起见，省略了  $\mathbf{p}$  和  $\mathbf{n}$  对  $\mathbf{u}$  的依赖关系。

考虑在图像平面上的一个闭合连通子集  $\Omega$  (即  $\mathbf{u} \in \Omega \subset \mathbb{R}^2$ ) 上观察到的法向图，分别在正交投影和透视投影下进行研究。在正交投影下，

$$\mathbf{p}(\mathbf{u}) = \begin{bmatrix} u \\ v \\ z(\mathbf{u}) \end{bmatrix}, \quad \frac{\partial \mathbf{p}}{\partial u} = \begin{bmatrix} 1 \\ 0 \\ \frac{\partial z}{\partial u} \end{bmatrix}, \quad \frac{\partial \mathbf{p}}{\partial v} = \begin{bmatrix} 0 \\ 1 \\ \frac{\partial z}{\partial v} \end{bmatrix}. \quad (5)$$

将公式 (5) 代入公式 (4)，得到一对偏微分方程：

$$n_z \frac{\partial z}{\partial u} + n_x = 0 \quad \text{和} \quad n_z \frac{\partial z}{\partial v} + n_y = 0. \quad (6)$$

在透视投影情况下，设  $f$  为相机的焦距， $[c_u, c_v]^T$  为图像平面上主点的坐标，则表面可以表示为  $\mathbf{p}(\mathbf{u}) = z(\mathbf{u})[(u - c_u)/f, (v - c_v)/f, 1]^T$ 。两个切向量为：

$$\frac{\partial \mathbf{p}}{\partial u} = \begin{bmatrix} \frac{1}{f}((u - c_u)\frac{\partial z}{\partial u} + z) \\ \frac{1}{f}((v - c_v)\frac{\partial z}{\partial u}) \\ \frac{\partial z}{\partial u} \end{bmatrix}, \quad \frac{\partial \mathbf{p}}{\partial v} = \begin{bmatrix} \frac{1}{f}((u - c_u)\frac{\partial z}{\partial v}) \\ \frac{1}{f}((v - c_v)\frac{\partial z}{\partial v} + z) \\ \frac{\partial z}{\partial v} \end{bmatrix}. \quad (7)$$

类似于 [3,4,8]，引入对数深度图  $\tilde{z}(\mathbf{u})$  满足  $z(\mathbf{u}) = \exp(\tilde{z}(\mathbf{u}))$  以统一公式。根据链式法则，有：

$$\frac{\partial z}{\partial u} = z \frac{\partial \tilde{z}}{\partial u} \quad \text{和} \quad \frac{\partial z}{\partial v} = z \frac{\partial \tilde{z}}{\partial v}. \quad (8)$$

将公式 (8) 代入公式 (7)，得到：

$$\frac{\partial \mathbf{p}}{\partial u} = z \begin{bmatrix} \frac{1}{f}((u - c_u)\frac{\partial \tilde{z}}{\partial u} + 1) \\ \frac{1}{f}((v - c_v)\frac{\partial \tilde{z}}{\partial u}) \\ \frac{\partial \tilde{z}}{\partial u} \end{bmatrix}, \quad \frac{\partial \mathbf{p}}{\partial v} = z \begin{bmatrix} \frac{1}{f}((u - c_u)\frac{\partial \tilde{z}}{\partial v}) \\ \frac{1}{f}((v - c_v)\frac{\partial \tilde{z}}{\partial v} + 1) \\ \frac{\partial \tilde{z}}{\partial v} \end{bmatrix}. \quad (9)$$

将公式 (9) 代入公式 (4)，消去  $z$ ，整理后得到：

$$\begin{cases} (n_x(u - c_u) + n_y(v - c_v) + n_z f) \frac{\partial \tilde{z}}{\partial u} + n_x = 0 \\ (n_x(u - c_u) + n_y(v - c_v) + n_z f) \frac{\partial \tilde{z}}{\partial v} + n_y = 0 \end{cases}. \quad (10)$$

记  $\tilde{n}_z = n_x(u - c_u) + n_y(v - c_v) + n_z f$ ，公式 (10) 可以简化为：

$$\tilde{n}_z \frac{\partial \tilde{z}}{\partial u} + n_x = 0 \quad \text{和} \quad \tilde{n}_z \frac{\partial \tilde{z}}{\partial v} + n_y = 0, \quad (11)$$

公式 (11) 与正交投影情况下的公式 (6) 形式相同。可以从法向图和相机参数预先计算  $\tilde{n}_z$ 。一旦估计出对数深度图  $\tilde{z}$ ，就可以通过指数运算得到深度图  $z$ 。

在公式 (6) 和 (11) 中，已经将正交和透视情况下的偏微分方程统一起来。在接下来的章节 (2.2 和 2.3) 中，将不再区分正交和透视情况。

### 3.2 双边加权泛函

现在描述用于保持不连续性法向积分的双边加权泛函。所提出的泛函可以应用于公式 (6) 或 (11)，具体取决于相机投影模型。不失一般性，下文将使用公式 (6) 中的符号。为了引出在半光滑表面假设下的双边加权泛函，首先讨论在光滑表面假设下的传统二次泛函。

当假设表面光滑时，目标表面处处可微，即偏导数  $\frac{\partial z}{\partial u}$  和  $\frac{\partial z}{\partial v}$  处处存在。因此，可以通过最小化二次泛函来找到深度图：

$$\min \iint_{\Omega} (n_z \frac{\partial z}{\partial u} + n_x)^2 + (n_z \frac{\partial z}{\partial v} + n_y)^2 dudv. \quad (12)$$

当一个函数在某一点可微时，它在该点两侧也是单侧可微的。因此，水平方向（ $\frac{\partial^+ z}{\partial u}$  和  $\frac{\partial^- z}{\partial u}$ ）和垂直方向（ $\frac{\partial^+ z}{\partial v}$  和  $\frac{\partial^- z}{\partial v}$ ）的单侧偏导数在该点两侧都存在，并且等于偏导数，即  $\frac{\partial^+ z}{\partial u} = \frac{\partial^- z}{\partial u} = \frac{\partial z}{\partial u}$  且  $\frac{\partial^+ z}{\partial v} = \frac{\partial^- z}{\partial v} = \frac{\partial z}{\partial v}$ 。因此，在光滑表面假设下，二次泛函 (12) 等价于：

$$\min \iint_{\Omega} 0.5(n_z \frac{\partial^+ z}{\partial u} + n_x)^2 + 0.5(n_z \frac{\partial^- z}{\partial u} + n_x)^2 + 0.5(n_z \frac{\partial^+ z}{\partial v} + n_y)^2 + 0.5(n_z \frac{\partial^- z}{\partial v} + n_y)^2 dudv, \quad (13)$$

其中单侧偏导数定义为：

$$\frac{\partial^+ z}{\partial u} = \lim_{h \rightarrow 0^+} \frac{z(u+h, v) - z(u, v)}{h}, \quad \frac{\partial^- z}{\partial u} = \lim_{h \rightarrow 0^-} \frac{z(u+h, v) - z(u, v)}{h}, \quad (14)$$

$$\frac{\partial^+ z}{\partial v} = \lim_{h \rightarrow 0^+} \frac{z(u, v+h) - z(u, v)}{h}, \quad \frac{\partial^- z}{\partial v} = \lim_{h \rightarrow 0^-} \frac{z(u, v+h) - z(u, v)}{h}. \quad (15)$$

考虑一个点两侧的单侧可微性，可以引出半光滑表面假设。假设一个半光滑表面在水平和垂直方向上，至多在一个点的一侧是不可微的（因此是不连续的）。如图 2 所示，这个假设包含三种情况。在可微点，半光滑表面也保证是双侧可微的（图 2 (a)， $\frac{\partial^+ z}{\partial u}$  和  $\frac{\partial^- z}{\partial u}$  都存在）。与光滑表面不同，半光滑表面在单侧不连续点可以是单侧不可微的（图 2 (b) 和 (c)， $\frac{\partial^+ z}{\partial u}$  或  $\frac{\partial^- z}{\partial u}$  存在）。另一方面，半光滑表面不包含在水平或垂直方向上两侧都不可微的点（图 2 (d) 和 (e)，不允许  $\frac{\partial^+ z}{\partial u}$  和  $\frac{\partial^- z}{\partial u}$  都不存在的情况）。

在半光滑表面假设下，提出双边加权泛函：

$$\min \iint_{\Omega} w_u(n_z \frac{\partial^+ z}{\partial u} + n_x)^2 + (1-w_u)(n_z \frac{\partial^- z}{\partial u} + n_x)^2 + w_v(n_z \frac{\partial^+ z}{\partial v} + n_y)^2 + (1-w_v)(n_z \frac{\partial^- z}{\partial v} + n_y)^2 dudv, \quad (16)$$

其中  $w_u$  和  $w_v$  表示每个点两侧的单侧可微性：

$$w_u = \begin{cases} 1 & \text{(仅右侧可微)} \\ 0.5 & \text{(左侧和右侧都可微)} \\ 0 & \text{(仅左侧可微)} \end{cases} \quad \text{和} \quad w_v = \begin{cases} 1 & \text{(仅上侧可微)} \\ 0.5 & \text{(上侧和下侧都可微)} \\ 0 & \text{(仅下侧可微)} \end{cases}. \quad (17)$$

双边加权泛函 (16) 表示，例如，当深度图在某一点左侧可微但右侧不可微时，数据项在该点左侧被保留，但在右侧被忽略。当深度图在某一点可微时，两侧的数据项被平等加权。这种相对加权因此涵盖了半光滑表面上每个点的所有可能情况。

现在，问题是如何确定表面在每个点处的单侧可微性。为此，可以使用单侧可微性需要单侧连续性这一事实。例如，一个函数在某一点右侧（不）可微，则它在该点必须右侧（不）连续。一个函数在某一点单侧连续，要求该点的函数值等于从相应一侧逼近的极限。形式上，将函数值与某一点处的单侧极限之间的差表示为：

$$\Delta_u^+ z = z(u, v) - \lim_{h \rightarrow 0^+} z(u+h, v), \quad \Delta_u^- z = z(u, v) - \lim_{h \rightarrow 0^-} z(u+h, v), \quad (18)$$

$$\Delta_v^+ z = z(u, v) - \lim_{h \rightarrow 0^+} z(u, v+h), \quad \Delta_v^- z = z(u, v) - \lim_{h \rightarrow 0^-} z(u, v+h). \quad (19)$$



如果  $\Delta_u^+ z = 0$ ，则函数右连续；如果  $\Delta_u^+ z \neq 0$ ，则函数右不连续；左侧、上侧和下侧连续性同理。为了进一步判断一个点在哪一侧不连续，可以通过一个函数比较该点两侧的单侧连续性。如果  $(\Delta_u^- z)^2 - (\Delta_u^+ z)^2 = 0$ ，则表面在该点处连续；如果  $(\Delta_u^- z)^2 - (\Delta_u^+ z)^2 < 0$ ，则仅左侧连续；如果  $(\Delta_u^- z)^2 - (\Delta_u^+ z)^2 > 0$ ，则仅右侧连续。将此差值与 Heaviside 阶跃函数  $H(x) = \{1 \text{ if } x > 0; 0.5 \text{ if } x = 0; 0 \text{ if } x < 0\}$  结合，得到水平和垂直方向上的权重函数：

$$w_u = H((\Delta_u^- z)^2 - (\Delta_u^+ z)^2) \quad \text{和} \quad w_v = H((\Delta_v^- z)^2 - (\Delta_v^+ z)^2). \quad (20)$$

可以验证权重函数公式 (20) 取值为公式 (17)。

### 3.3 求解方法

本节介绍双边加权泛函 (16) 和权重函数公式 (20) 的离散化，并推导出在离散像素域中观察到的法向图的求解方法，即  $\Omega \subset \mathbb{Z}^2$ 。在下文中，不会区分连续和离散情况下的符号。

首先通过前向或后向差分将单侧偏导数公式 (14) 和 (15) 离散化，即在公式 (14) 和 (15) 中分别取  $h = 1$  或  $-1$ ：

$$\frac{\partial^+ z}{\partial u} \approx z(u+1, v) - z(u, v), \quad \frac{\partial^- z}{\partial u} \approx z(u, v) - z(u-1, v), \quad (21)$$

$$\frac{\partial^+ z}{\partial v} \approx z(u, v+1) - z(u, v), \quad \frac{\partial^- z}{\partial v} \approx z(u, v) - z(u, v-1). \quad (22)$$

然后用相邻像素处的深度值来近似单侧极限，公式 (18) 和 (19) 变为：

$$\Delta_u^+ z \approx n_z(z(u, v) - z(u+1, v)), \quad \Delta_u^- z \approx n_z(z(u, v) - z(u-1, v)), \quad (23)$$

$$\Delta_v^+ z \approx n_z(z(u, v) - z(u, v+1)), \quad \Delta_v^- z \approx n_z(z(u, v) - z(u, v-1)). \quad (24)$$

这里，深度差由  $n_z$  缩放，以衡量沿该点法线方向的差异。为了避免阶跃函数在离散域中始终取二元值（即将每个像素都视为单侧不连续），将阶跃函数近似为 sigmoid 函数：

$$H(x) \approx \sigma_k(x) = \frac{1}{1 + e^{-kx}}, \quad (25)$$

其中参数  $k$  控制 sigmoid 函数的锐度。将公式 (21) 到 (25) 结合起来，得到了离散化的双边加权泛函：

$$\begin{aligned} \min_{z(u,v)} \sum_{\Omega} & \sigma_k((\Delta^- z)^2 - (\Delta^+ z)^2) (n_z \frac{\partial^+ z}{\partial u} + n_x)^2 \\ & + \sigma_k((\Delta^+ z)^2 - (\Delta^- z)^2) (n_z \frac{\partial^- z}{\partial u} + n_x)^2 \\ & + \sigma_k((\Delta^- z)^2 - (\Delta^+ z)^2) (n_z \frac{\partial^+ z}{\partial v} + n_y)^2 \\ & + \sigma_k((\Delta^+ z)^2 - (\Delta^- z)^2) (n_z \frac{\partial^- z}{\partial v} + n_y)^2. \end{aligned} \quad (26)$$

在公式 (26) 中, 使用了 sigmoid 函数的性质  $1 - \sigma_k(x) = \sigma_k(-x)$  使其更紧凑。直观地说, 优化问题 (26) 表明, 如果一个像素一侧的深度差远大于另一侧, 则较大的一侧更可能是不连续的, 并且相应地对二次数据项的加权较小。

由于权重是未知深度的非线性 sigmoid 函数, 优化问题 (26) 是非凸的。为了求解 (26), 使用迭代重加权最小二乘法 (IRLS)。为了描述迭代过程, 首先准备 (26) 的矩阵形式。

令  $\mathbf{z}, \mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z \in \mathbb{R}^m$  分别是由所有  $|\Omega| = m$  个像素的  $z(\mathbf{u}), n_x(\mathbf{u}), n_y(\mathbf{u}), n_z(\mathbf{u})$  以相同顺序序列化后的向量。令  $\text{diag}(\mathbf{x})$  为对角矩阵, 其第  $i$  个对角元素是  $\mathbf{x}$  的第  $i$  个元素, 并记  $\mathbf{N}_z = \text{diag}(\mathbf{n}_z)$ 。可以将优化问题 (26) 写成矩阵形式:

$$\min_{\mathbf{z}} (\mathbf{A}\mathbf{z} - \mathbf{b})^T \mathbf{W}(\mathbf{z})(\mathbf{A}\mathbf{z} - \mathbf{b}), \quad (27)$$

其中

$$\mathbf{A} = \begin{bmatrix} \mathbf{N}_z \mathbf{D}_u^+ \\ \mathbf{N}_z \mathbf{D}_u^- \\ \mathbf{N}_z \mathbf{D}_v^+ \\ \mathbf{N}_z \mathbf{D}_v^- \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -\mathbf{n}_x \\ -\mathbf{n}_x \\ -\mathbf{n}_y \\ -\mathbf{n}_y \end{bmatrix}, \quad \mathbf{W}(\mathbf{z}) = \text{diag} \begin{bmatrix} w_u(\mathbf{z}) \\ 1 - w_u(\mathbf{z}) \\ w_v(\mathbf{z}) \\ 1 - w_v(\mathbf{z}) \end{bmatrix}, \quad (28)$$

其中

$$w_u(\mathbf{z}) = \sigma_k \left[ (\mathbf{N}_z \mathbf{D}_u^- \mathbf{z})^{\circ 2} - (\mathbf{N}_z \mathbf{D}_u^+ \mathbf{z})^{\circ 2} \right], \quad (29)$$

$$w_v(\mathbf{z}) = \sigma_k \left[ (\mathbf{N}_z \mathbf{D}_v^- \mathbf{z})^{\circ 2} - (\mathbf{N}_z \mathbf{D}_v^+ \mathbf{z})^{\circ 2} \right]. \quad (30)$$

这里, 四个矩阵  $\mathbf{D}_u^+, \mathbf{D}_u^-, \mathbf{D}_v^+, \mathbf{D}_v^- \in \mathbb{R}^{m \times m}$  是离散偏导数矩阵。第  $i$  行要么只包含两个非零元素  $-1$  和  $1$ , 要么是一个零向量, 如果第  $i$  个像素的相邻像素在域  $\Omega$  之外; 更多细节可以在 [9] 中找到。此外,  $\mathbf{1} \in \mathbb{R}^m$  是一个全一向量,  $\sigma_k(\cdot)$  现在是 sigmoid 函数公式 (25) 的逐元素版本,  $(\cdot)^{\circ 2}$  是向量的逐元素平方函数。

在优化过程的每一步  $t$ , 首先固定权重矩阵  $\mathbf{W}(\mathbf{z}^{(t)})$ , 然后求解深度  $\mathbf{z}$ :

$$\mathbf{z}^{(t+1)} = \arg \min_{\mathbf{z}} (\mathbf{A}\mathbf{z} - \mathbf{b})^T \mathbf{W}(\mathbf{z}^{(t)})(\mathbf{A}\mathbf{z} - \mathbf{b}). \quad (31)$$

当  $\mathbf{W}(\mathbf{z}^{(t)})$  固定时, 公式 (31) 退化为一个凸加权最小二乘问题。可以通过求解公式 (31) 的正规方程来找到  $\mathbf{z}^{(t+1)}$ :

$$\mathbf{A}^T \mathbf{W}(\mathbf{z}^{(t)}) \mathbf{A} \mathbf{z} = \mathbf{A}^T \mathbf{W}(\mathbf{z}^{(t)}) \mathbf{b}. \quad (32)$$

矩阵  $\mathbf{A}$  是秩亏 1 的; 一维零空间基是一个全一向量, 对应于结果中的偏移歧义。在透视情况下, 偏移歧义在对结果进行指数运算后变为尺度歧义。在实现中, 使用共轭梯度方法来求解公式 (32)。

将  $\mathbf{z}^{(0)}$  初始化为一个平面, 或者等效地, 将所有权重初始化为 0.5。第一步  $\mathbf{z}^{(1)}$  估计的深度图与在光滑表面假设下估计的深度图相同, 即使用泛函 (12)。将第  $t$  步目标函数 (27) 的能量表示为  $E_t = (\mathbf{A}\mathbf{z}^{(t)} - \mathbf{b})^T \mathbf{W}(\mathbf{z}^{(t)})(\mathbf{A}\mathbf{z}^{(t)} - \mathbf{b})$ , 一旦相对能量  $|E_t - E_{t-1}|/E_{t-1}$  小于用户提供的容差或超过最大迭代次数, 就终止迭代。

## 4 复现细节

### 4.1 开源代码

在本报告的复现工作中，主要参考了以下开源代码：

代码名称: `bilateral_normal_integration` 代码来源: [https://github.com/xucaoy42/bilateral\\_normal\\_integration](https://github.com/xucaoy42/bilateral_normal_integration) 本报告复现并验证了该开源代码中实现的双边法向积分算法。代码结构和核心算法均基于该开源代码，并根据论文描述进行了细节上的核对和调整，以确保复现结果的准确性。该开源代码实现了 Cao 等人提出的双边加权法向积分方法，包括数据预处理、双边加权泛函构建、优化求解以及结果评估等功能。本报告在此基础上进行了实验验证和分析。

### 4.2 实验环境搭建

本报告的实验环境如下：

硬件环境:

CPU: AMD EPYC 7282 16-Core Processor @ 2.80GHz

内存: 128 GB DDR4

GPU: NVIDIA GeForce RTX 4090

软件环境 (参考 requirements.txt 文件):

操作系统: Ubuntu 18.04.3 LTS

编程语言: Python 3.7

NumPy: 1.20.2

SciPy: 1.6.2

OpenCV: 4.5.1.48 (opencv-python)

PyVista: 0.37.0

### 4.3 界面分析与使用说明

本程序基于命令行界面运行，入口文件为 ‘`bilateral_normal_integration_numpy.py`’。运行方式如下：

```
1 python bilateral_normal_integration_numpy.py --path <法线贴图相对路径> -k  
   <sigmoid函数的锐度> --iter <IRLS 的最大迭代次数> --tol <IRLS 的停止容  
   差>
```

参数说明:



参数	含义	类型	取值范围	默认值
path	法线贴图相对路径	字符串	-	-
k	sigmoid 函数的锐度	浮点数	大于 0 的实数	2
iter	IRLS 的最大迭代次数	无符号整数	大于 0 的整数	150
tol	IRLS 的停止容差	浮点数	小的正数	1e-4

表 1. 参数说明

运行示例:

```
1 python bilateral_normal_integration_numpy.py --path data/supp_vase -k 4
   --iter 100 --tol 1e-5
```

输出结果:

程序运行后, 会在 path 指定的目录下生成深度图重建的三维表面网格, 文件格式为 ply。

```
(bilateral) root@rt-res-public22-7d5df6b5b9-xts.jp:/public22_data/ocx/bilateral_normal_integration# python bilateral_normal_integration_numpy.py --path data/supp_vase -k 4
--iter 100 --tol 1e-5
Running bilateral normal integration with k=4.0 in the orthographic case.
The number of normal vectors is 28224.
step 48/100 energy: 111.515 relative energy: 0.000e+00: 47% | 47/100 [03:31<03:58, 4.51s/it]
Total time: 211.829 sec
saved data/supp_vase
```

图 1. 操作界面示意

## 5 实验结果与分析

### 5.1 实验设置

为复现 Cao 等人的工作, 本实验尽可能遵循了原论文中描述的实验设置。

### 5.2 数据集

实验使用了与原论文相同的数据集 DiLiGenT, 它是一个基准测试数据集, 使用数据集中提供的真实透视法线贴图和对应的真实深度图进行测试。该数据集包含以下物体: "Bear", "Buddha", "Cat", "Cow", "Harvest", "Pot1", "Pot2", "Reading", "Goblet"。原论文使用该数据集评估算法在透视投影下的性能。

### 5.3 参数配置

为确保复现的准确性, 本实验尽可能遵循了原论文中提供的参数配置。对于 IPF, 使用了其五点版本, 该方法没有超参数。对于 BiNI 算法, 原论文中唯一的超参数是  $k$ , 设置为  $k = 2$ 。迭代最大次数设置为 100, 停止迭代的误差阈值设置为  $1 \times 10^{-5}$ 。

### 5.4 评价指标

对于提供真实深度图 (GT) 的数据集, 计算了重建表面和真实深度图之间的平均绝对深度误差 (Mean Absolute Depth Error, MADE), 并展示了绝对深度误差图。该指标与原论文

中使用的指标相同。为了消除正交投影情况下的偏移歧义，对重建表面进行了平移，使得平移后的表面与真实深度图之间的 L1 范数最小。类似地，在透视投影情况下，通过缩放重建表面来消除尺度歧义。

## 5.5 定量分析



图 2. 三维网格图 reading

图2为算法生成 reading 三维网格图，位于目录 data/Fig4\_reading/，可用 3D 查看器查看。

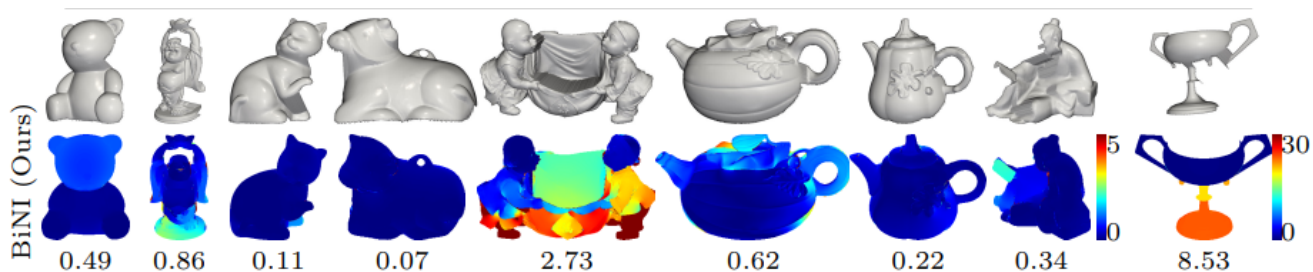


图 3. 原论文结果

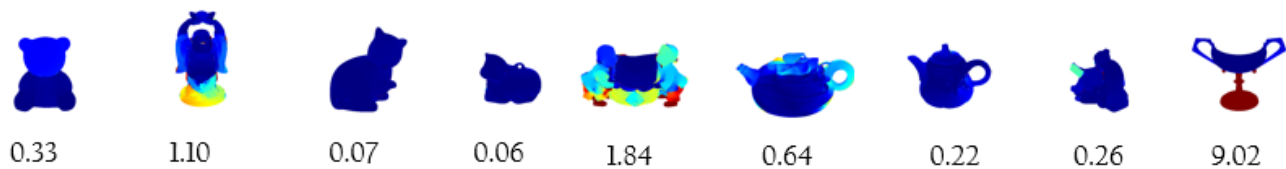


图 4. 复现结果

图3为原论文实验结果。图 4 展示了在合成正交法线贴图上的定量比较结果。第一行为得到的法线贴图的实验结果，第二行为绝对深度误差图的实验结果。

整体上看, 两张图的颜色分布趋势是比较一致的, 大部分区域都是蓝色, 表示误差较小; 部分边缘和复杂结构处呈现黄色到红色, 表示误差较大。图3 和图 4 中每个物体下方都标注了平均绝对深度误差 (MADE) 值。

Bear, Cat, Cow, Pot1, Pot2, Reading 这五个物体的复现结果与原论文结果非常接近, 无论是颜色分布还是 MADE 值都高度一致。这说明双边法向积分算法对于这些形状相对简单的物体具有较好的稳定性和可复现性。

Buddha 复现结果中"Buddha" 佛像的面部和身体部分出现了更多黄色和红色, MADE 值也更大 ( $1.10 > 0.86$ )。Harvest 复现结果中"Harvest" 的 MADE 值比原论文更小 ( $1.84 < 2.73$ ), 从颜色上看, 复现结果中黄色区域更少。两者差异较大。Goblet 复现结果中"Goblet" 的 MADE 值比原论文更大 ( $9.02 > 8.53$ )。观察颜色分布可以发现, 两者主要差别在于杯子底部。原论文中杯子底部呈现出非常明显的红色到蓝色的过渡, 而复现结果中杯子底部几乎全为红色。表明该区域的重建误差较大, 可能是由于 BiNI 对这种特定类型的几何形状(快速且连续的深度变化) 的敏感性。

有些结果比论文中的结果略好, 这可能是因为自从作者在论文中报告这些指标以来, 代码实现上可能有一些改进。不同的硬件、操作系统、软件包版本等都可能数值计算结果的微小差异。

总的来说, 复现结果与原论文结果基本一致, 表明双边法向积分算法能够有效地重建深度图, 并在大多数情况下保持较低的误差。但也存在一些局部差异, 特别是在细节丰富、深度变化剧烈或具有薄而尖锐结构的区域。这些差异可能是由于代码实现细节或运行环境的不同造成的。

## 6 总结与展望

### 6.1 总结

本报告针对三维重建中的法向积分问题, 复现并验证了 Cao 等人提出的基于双边加权的法向积分算法 (BiNI) [2]。该算法引入“半光滑表面”假设, 即真实表面在每个像素邻域内, 水平与垂直方向上至多仅一侧存在不连续。基于此, 构建了双边加权泛函来描述表面重建的能量函数, 并采用迭代重加权最小二乘法 (IRLS) 进行求解, 从而自适应调整各区域权重, 在保留不连续性的同时抑制噪声。此外, 原文提出了统一的正交及透视投影法向积分框架, 使得该算法能适配不同的成像模型。

本报告详细阐述了 BiNI 算法原理与实现细节, 并基于开源代码在 DiLiGenT 数据集 [?] 上进行了实验验证。复现结果与原论文结果的对比及定量分析表明, 复现工作较为成功, 验证了该算法的有效性和可复现性。实验结果显示, 相较于传统的基于光滑假设的法向积分方法, BiNI 能更准确地重建复杂几何结构, 特别是在不连续区域优势明显。虽然在某些特定对象上存在些许差异, 可能源于代码细节、运行环境或原论文代码的后续改进, 但总体上复现结果印证了 BiNI 算法的有效性。

## 6.2 不足与展望

尽管复现工作较为成功，但仍存在一些不足，主要体现在算法对特定几何形状（如快速且连续的深度变化区域）的敏感性，以及参数  $k$ （控制 sigmoid 函数锐度）的选择问题。此外，基于 IRLS 的求解方法在计算效率上仍有提升空间，特别是处理高分辨率法向图时。最后，本报告仅在 DiLiGenT 数据集上进行了验证，该数据集主要包含常见物体，在更广泛、更具挑战的数据集上评估算法性能仍有必要。

未来的研究方向可从以下几个方面展开：首先，探索更鲁棒、更自适应的权重计算方法，例如结合其他线索（如图像强度）辅助权重估计，或利用更精细的邻域信息改进单侧连续性的判断，并研究如何根据法向图的局部特征自适应地选择最优  $k$  值。其次，在更广泛、更具挑战的数据集上评估 BiNI 的性能，并探索将深度学习与 BiNI 结合的可能性，例如使用深度神经网络学习更准确的权重函数，或直接从法向图预测深度图。最后，将 BiNI 应用于实际的三维重建任务，如机器人导航、文物数字化、医学图像分析等，并根据具体应用场景对算法进行改进和优化。

总而言之，基于双边加权的法向积分算法为解决三维重建中的不连续性问题提供了一种有效途径。未来的研究将致力于进一步提升该算法的鲁棒性、准确性和效率，并将其推广至更广阔的应用领域。

## 参考文献

- [1] Hicham Badri, Hussein Yahia, and Driss Aboutajdine. Robust surface reconstruction via triple sparsity. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2283–2290, 2014.
- [2] Xu Cao, Hiroaki Santo, Boxin Shi, Fumio Okura, and Yasuyuki Matsushita. Bilateral normal integration. In *Proc. Euro. Conf. on Computer Vision*, page 552–567, 2022.
- [3] Jean-Denis Durou, Jean-François Aujol, and Frédéric Courteille. Integrating the normal field of a surface in the presence of discontinuities. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 261–273, 2009.
- [4] Jean-Denis Durou and Frédéric Courteille. Integration of a normal field without boundary condition. In *Proceedings of the First International Workshop on Photometric Analysis For Computer Vision-PACV 2007*, pages 8–p, 2007.
- [5] Matthew Harker and Paul O’Leary. Least squares surface reconstruction from measured gradient fields. In *2008 IEEE conference on computer vision and pattern recognition*, pages 1–7. IEEE, 2008.
- [6] Itsik Horowitz and Nahum Kiryati. Depth from gradient fields and control points: Bias correction in photometric stereo. *Image and Vision Computing*, 22(9):681–694, 2004.

- [7] Bilge Karaçalı and Wesley Snyder. Reconstructing discontinuous surfaces from a given gradient field using partial integrability. *Computer Vision and Image Understanding*, 92(1):78–111, 2003.
- [8] Yvain Quéau, Jean-Denis Durou, and Jean-François Aujol. Normal integration: a survey. *Journal of Mathematical Imaging and Vision*, 60:576–593, 2018.
- [9] Yvain Quéau, Jean-Denis Durou, and Jean-François Aujol. Variational methods for normal integration. *Journal of Mathematical Imaging and Vision*, 60:609–632, 2018.
- [10] Rafael FV Saracchini, Jorge Stolfi, Helena CG Leitão, Gary A Atkinson, and Melvyn L Smith. A robust multi-scale integration method to obtain the depth from gradient maps. *Computer Vision and Image Understanding*, 116(8):882–895, 2012.
- [11] Yinting Wang, Jiajun Bu, Na Li, Mingli Song, and Ping Tan. Detecting discontinuities for surface reconstruction. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pages 2108–2111. IEEE, 2012.
- [12] Tai-Pang Wu and Chi-Keung Tang. Visible surface reconstruction from normals with discontinuity consideration. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, volume 2, pages 1793–1800. IEEE, 2006.
- [13] Wuyuan Xie, Miaohui Wang, Mingqiang Wei, Jianmin Jiang, and Jing Qin. Surface reconstruction from normals: A robust dgp-based discontinuity preservation approach. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5328–5336, 2019.
- [14] Dizhong Zhu and William AP Smith. Least squares surface reconstruction on arbitrary domains. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXII 16*, pages 530–545. Springer, 2020.