

# 基于多视点编码光场的全景三维重建方法

王泽宇<sup>1,2</sup>, 向森<sup>1,2\*</sup>, 邓慧萍<sup>1,2</sup>, 吴谨<sup>1,2</sup>

<sup>1</sup>武汉科技大学信息科学与工程学院, 湖北 武汉 430081;

<sup>2</sup>武汉科技大学冶金自动化与检测技术教育部工程研究中心, 湖北 武汉 430081

**摘要** 在基于光场的一系列应用中,目标的三维重建是基础且关键的任务。普通光场只能重建单一视角而无法重建全景,并且在纹理特征匮乏的区域也无法生成准确的三维信息。针对以上问题,提出一种基于多视点编码光场的全景三维重建方法。首先,利用结构光编码向场景投射正弦模板生成像素级的相位码字以丰富场景特征,进而在多个角度独立地进行光场采集。然后,在各个光场采样视点的内部利用空间信息和角度信息求取深度值。最后,在不同的光场采样点之间由位姿引导实现三维信息融合。实验结果表明,所提方法能有效恢复物体的全景信息,而且对物体纹理特征匮乏区域的重建也更加准确。

**关键词** 图像处理; 三维重建; 光场; 结构光; 相位编码; 点云

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP221282

## Panoramic Three-Dimensional Reconstruction Method Based on Multi-View Encoded Light Field

Wang Zeyu<sup>1,2</sup>, Xiang Sen<sup>1,2\*</sup>, Deng Huiping<sup>1,2</sup>, Wu Jin<sup>1,2</sup>

<sup>1</sup>*School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China;*

<sup>2</sup>*Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China*

**Abstract** In a series of applications based on a light field, three-dimensional (3D) reconstruction of objects is a basic and key task. Ordinary light fields can only reconstruct a single perspective; however, they cannot reconstruct a panorama and generate accurate 3D information in areas with scarce texture features. To solve these problems, a panoramic 3D reconstruction method based on a multi-view coding light field is proposed. First, the structured light coding is used to project sine templates to the scene to generate pixel-level phase codewords to enrich the scene features. Then, the light field is collected independently from multiple angles. Furthermore, the depth value is obtained using spatial and angle information in the interior of each light field sampling viewpoint. Finally, 3D information fusion is realized between different light field sampling points under the pose guidance. The experimental results show that the proposed method can effectively restore the panorama information of the object. Moreover, the reconstruction of an area with insufficient texture features of the object is more accurate.

**Key words** image processing; three-dimensional reconstruction; light field; structured light; phase-coding; point cloud

## 1 引言

三维重建是计算机视觉中的一项重要任务,广泛应用于虚拟现实技术<sup>[1]</sup>、无人驾驶<sup>[2]</sup>、机器人视觉导航<sup>[3]</sup>等领域。基于多视图匹配的三维重建技术如 structure from motion (SFM)<sup>[4]</sup>和 simultaneous localization

and mapping (SLAM)<sup>[5]</sup>已经被用于大规模场景的稀疏重建。此类方法通过特征点匹配计算特征点的空间坐标,进而重建出物体的三维模型。然而它们面临着两方面不足:1)场景中的特征点分布稀疏,导致重建的三维模型也十分稀疏,无法描述几何细节;2)特征点的匹配计算复杂,制约三维重建的速度。

收稿日期: 2022-04-12; 修回日期: 2022-04-25; 录用日期: 2022-06-14; 网络首发日期: 2022-06-24

基金项目: 国家自然科学基金(61702384, 62001180, 61871437)

通信作者: \*xiangsen@wust.edu.cn

四维光场(LF)<sup>[6]</sup>同时记录光线的强度和方向,其数据高维度的优点催生了大量应用;并且随着 Lytro 和 Raytrix 等便携式光场摄像机的普及,基于光场的三维重建方法被陆续提出。光场根据数字重聚焦中的最佳剪切值<sup>[7]</sup>或极平面图像(EPI)中的最优斜率<sup>[8-9]</sup>来获取场景深度。Tao 等<sup>[10]</sup>综合度量散焦线索和相关性线索两种代价获取深度值。Williem 等<sup>[11]</sup>综合度量约束角熵代价和约束自适应离焦代价获取深度值。Peng 等<sup>[12]</sup>认为在同一视点不同视图中的深度图包含互补信息,并提出融合不同视图实现目标重建的方法。此类方法通过精确的深度值估计实现物体重建,但深度线索依赖光场的空间和角度信息,只能在单一视角重建目标,无法恢复目标完整的全景三维信息。

全景三维重建需要从不同的视角获得场景的几何信息再进行融合,常见的方法有 360°重建和 720°重建。360°重建中,视点沿水平环形布置,720°重建则同时沿着水平和垂直方向环形布局。这一类方法利用场景几何与光场结构之间的关系计算位姿<sup>[13]</sup>或者利用高精度的采集设备获取位姿<sup>[14]</sup>,并由位姿引导实现立体匹配。Chen 等<sup>[15]</sup>使用双边一致性度量进行可靠的光场立体匹配。Zhang 等<sup>[16]</sup>在此基础上提出 P-SFM,结合立体匹配技术和光场光束平差技术计算光场相机位姿和场景几何形状。Vianello 等<sup>[14]</sup>利用高精度采集设备获取 360°环形光场,利用 Hough 变换计算环形 EPI 获取单视点深度并使用采样过程中得到的位姿实现信息融合。宋征玺等<sup>[17]</sup>在此基础上构建三维霍夫空间提取空间点轨迹实现重建。此类方法可以实现物体的全景三维重建,然而深度计算和位姿估计均依赖于表面纹理,可靠性相对较差。因此通常需要较多的数据,利用数

据冗余来提高三维重建的可靠性。

针对以上问题,本文提出一种基于多视点编码光场融合的全景三维重建方法。使用光场相机环绕物体进行采样,利用光场内部的高维数据独立获取各个视点的深度值,进而将不同视点的深度值转换为点云并在全局坐标空间中融合,实现物体的全景重建。同时,在普通光场的基础上引入相位编码光场以丰富场景特征,从而提高深度值质量。所提方法具有以下优点:1)采用编码光场的相位编码代替自然场景,解决弱纹理或无纹理区域由于缺乏特征导致的深度估计不准确问题。2)采用多视点融合策略,解决普通光场只能实现单视角重建的问题,实现全景重建。同时在采样过程中获取单视点精确位姿,消除位姿计算导致的视点间匹配不准确问题。3)采用沿水平和垂直方向进行环形布局的信息采集方法,在获取物体全景信息的同时减少视点采集数。

## 2 多视点三维重建算法

所提基于多视点编码光场的精确三维全景重建方法的总体框架如图 1(a)所示,采用结构光编码技术赋予场景像素级码字特征,使用光场相机获取不同视点的编码光场,然后独立计算不同视点的深度值,最后将不同视点的信息在三维坐标空间中融合。相对于已有的三维重建方案,所提框架具有以下特点:1)在深度值估计阶段,采用一种符合结构光特性的重聚焦方法,在提高深度值精度的同时降低复杂度;2)在视点融合阶段,使用采样中的位姿进行点云融合,避免重新计算位姿引起的误差;3)提出一种基于统计的飞行像素去除方法,优化重建目标在深度不连续区域的精度。

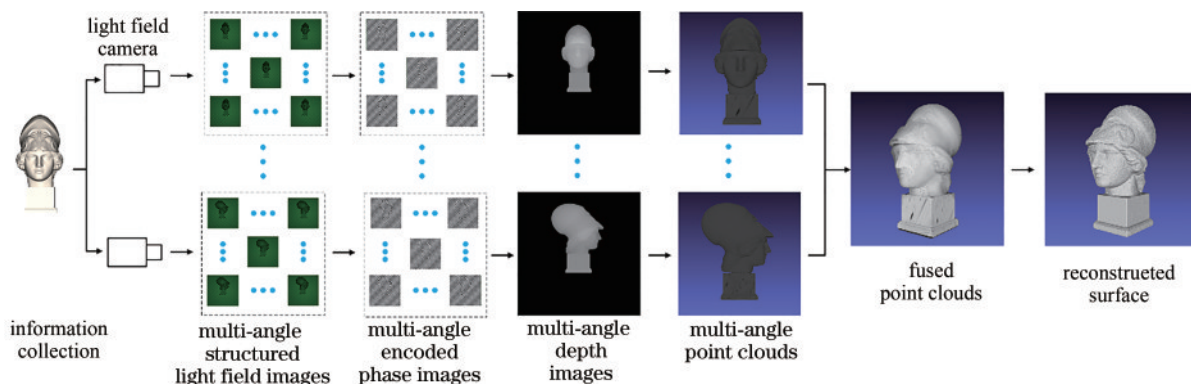


图 1 所提方法框架

Fig. 1 Framework of proposed method

### 2.1 相位编码光场的获取

在各个视点采用相位编码光场进行深度值估计,采用多步相移法进行像素级编码,在减小量化误差的同时产生准确连续的码字。具体来说,使用  $N$  张同时受水平和垂直坐标调制的正弦函数模板,其中,第  $i$  张模板被定义为

$$I_i(s, t) = A \cos \left[ \frac{2\pi(s+t)}{T} - i \frac{2\pi}{N} \right], \quad (1)$$

式中:  $A$  表示振幅;  $(s, t)$  表示像素的坐标;  $T$  表示正弦函数周期。这些图像由光场相机采集,通过  $N$  步相移法获得包裹相位:

$$\varphi(s, t) = \arctan \left[ \frac{\sum_{i=1}^N I_i(s, t) \sin\left(i \frac{2\pi}{N}\right)}{\sum_{i=1}^N I_i(s, t) \cos\left(i \frac{2\pi}{N}\right)} \right]. \quad (2)$$

包裹相位  $\varphi(x, y)$  通过  $\arctan$  函数求出, 相位值分布在  $(-\pi, \pi)$  中, 并在水平和垂直方向上周期性循环。通过多步相移法得到的图像, 像素值在其邻域内连续且唯一。

### 2.2 改进的重聚焦算法

通过第 2.1 节得到编码光场后, 使用重聚焦算法进行深度估计。对于编码光场, 通过四维剪切对其进行重聚焦:

$$\varphi_\alpha(s, t, u, v) = \varphi \left[ s + u \left( 1 - \frac{1}{\alpha} \right), t + v \left( 1 - \frac{1}{\alpha} \right), u, v \right], \quad (3)$$

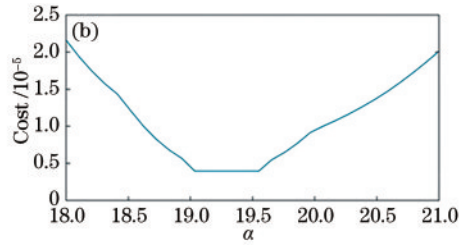
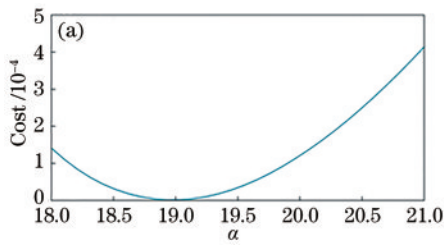


图 2 代价曲线。(a)所提方法的代价曲线;(b)普通光场的代价曲线

Fig. 2 Cost curves. (a) Cost curve of proposed method; (b) cost curve of conventional light field

在光场重采样阶段, 传统数字重聚焦的精度受光场分辨率的制约, 采样像素在多个的剪切值下均保持相同。当剪切值为  $\alpha_0$  和  $\alpha_1$  时, 最邻近的采样像素均为  $(s_0, t_1)$ , 产生量化误差降低视差的精度, 如图 3(a) 所示。为解决这个问题, 采用双线性插值法重新计算亚像素相位值, 如图 3(b) 所示, 将剪切值  $\alpha$  采样下的亚像素相位值记为

$$\varphi_\alpha = (1-m)(1-n)\varphi(s_0, t_0) + (1-m)n\varphi(s_0, t_1) + m(1-n)\varphi(s_1, t_0) + mn\varphi(s_1, t_1), \quad (5)$$

式中:  $\varphi_\alpha(s, t)$  表示空间坐标  $(s, t)$  处的相位值;  $n$  和  $m$

式中:  $\varphi$  表示光场图像;  $\alpha$  表示剪切值;  $\varphi_\alpha$  表示重聚焦之后的光场图像;  $(u, v)$  和  $(s, t)$  表示角度坐标与空间坐标。当重聚焦图像聚焦于一个特定深度值时, 聚焦平面上的点所处境块的强度一致, 成像最为清晰; 而聚焦平面以外的散焦点则变得模糊。剪切值  $\alpha$  反映场景的深度值, 通过代价函数  $\text{Cost}(\varphi_\alpha)$  评估深度线索, 代价最小的  $\alpha$  对应最佳深度值。

$$\alpha_{\text{right}} = \underset{\alpha}{\operatorname{argmin}} \text{Cost}(\varphi_\alpha). \quad (4)$$

传统光场的代价曲线如图 2(b) 所示, 受均匀纹理、复杂光照和遮挡等因素影响, 场景的强度值相似或相同, 代价函数出现多个最小值使  $\alpha$  无法唯一地收敛, 导致深度值估计不准确。针对上述问题, 对传统重聚焦算法在光场采集和代价计算两方面进行改进, 充分利用编码光场特性, 在减小计算量的同时提升深度估计的精度。

为亚像素与其左上方整像素点间的水平和垂直距离。普通光场使用强度作为编码信息, 像素间不具有连续性, 使用插值法破坏了图像的空间结构; 编码光场的像素在水平和垂直方向均连续, 基于双线性插值的亚像素采样方法充分利用这一特性, 在不破坏光场空间结构的同时提升数字重聚焦的采样精度。

得到重采样的相位后, 数字重聚焦通过代价最小化计算深度值。Tao 等<sup>[10]</sup>将散焦线索和相关性线索作为代价计算场景深度。当  $\alpha$  重新聚焦到正确的深度时, 相关性线索要求光场图像中所有与  $(s, t)$  对应的像

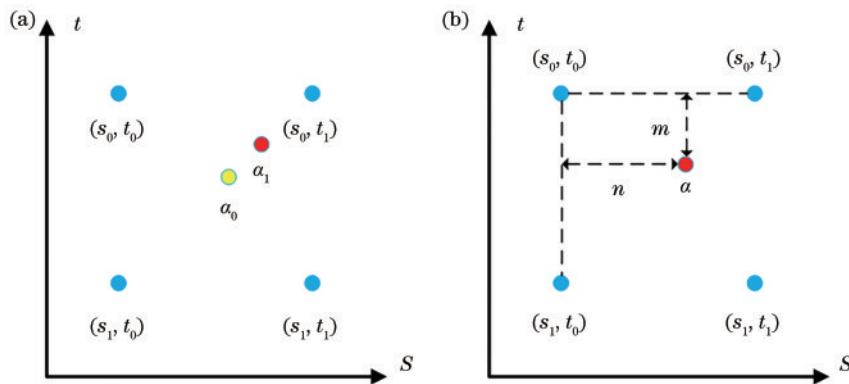


图 3 亚像素重采样示意图。(a)不同剪切值下整像素重采样;(b)亚像素的采样方法

Fig. 3 Schematic diagrams of subpixel resampling. (a) Subpixels at different shear values; (b) subpixel sampling method



素构成一个具有最小方差的角块,同时散焦线索要求该角块的像素值与中心视图相应的像素值的差异最小。根据 Cai 等<sup>[18]</sup>的研究,相位编码光场对散焦线索不敏感而对相关性线索敏感。因此仅使用相关性线索定义代价函数:

$$\text{Cost}(\varphi_\alpha) = \frac{1}{N^2} \sum_{u,v=1}^N [\varphi_\alpha(s, t, u, v) - \bar{\varphi}_\alpha(s, t, u, v)]^2, \quad (6)$$

$$\bar{\varphi}_\alpha(s, t, u, v) = \frac{1}{N^2} \sum_{u,v=1}^N \varphi_\alpha(s, t, u, v), \quad (7)$$

式中: $N$ 表示光场的角度分辨率; $\varphi_\alpha(s, t, u, v)$ 表示剪切值 $\alpha$ 下采样像素的相位值。仅使用相关性线索定义代价函数,在保持光场深度估计准确性的同时减小了运算量,加快了运算速度,如式(4)所示,通过计算求取最小化代价 $\text{Cost}(\varphi_\alpha)$ 的剪切值 $\alpha$ ,进一步通过文献[10]中的方法即可得到场景的深度值。

### 2.3 多视点三维信息融合

为实现物体全景三维重建,在相位编码重聚焦求得单视点深度图的基础上,还需要融合多视点的三维信息。对测量目标在多个水平和垂直的视点进行采样得到不同视点信息,使用数字重聚焦计算各视点的深度图,将各视点深度图中的像素投射到三维坐标空间,并使用位姿匹配实现信息融合。

对于任意一张深度图,通过下式将像素 $(s, t)$ 转换

为空间坐标 $(x, y, z)$ 进而转换为点云:

$$\begin{cases} x = \frac{z_0(s - s_0)dx}{f} - d \\ y = \frac{z_0(t - t_0)dy}{f} - d \\ z = z_0 - d \end{cases}, \quad (8)$$

$$\mathbf{P} = \bigcup_{i \in N_{\text{pix}}} (x_i, y_i, z_i), \quad (9)$$

式中: $z_0$ 表示像素 $(s, t)$ 处的深度值; $f$ 表示相机的焦距; $d$ 表示焦点与旋转中心的距离; $(s_0, t_0)$ 表示深度图中心像素的坐标; $dx$ 和 $dy$ 表示单一像素点在水平和垂直方向上的物理尺寸; $\mathbf{P}$ 表示单视点的点云; $\cup$ 表示空间坐标的集合; $N_{\text{pix}}$ 表示深度图像素个数。

图4(a)中,将物体首次采样的视点作为参考系,沿水平和垂直两个方向进行采样。将采样过程中水平和垂直方向的旋转角度定义为 $\theta_1, \theta_2$ ,则水平和垂直视点的位姿为

$$\mathbf{H}_\theta = \begin{pmatrix} \cos \theta_1 & 0 & \sin \theta_1 \\ 0 & 1 & 0 \\ -\sin \theta_1 & 0 & \cos \theta_1 \end{pmatrix}, \quad (10)$$

$$\mathbf{V}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_2 & -\sin \theta_2 \\ 0 & \sin \theta_2 & \cos \theta_2 \end{pmatrix}, \quad (11)$$

式中: $\mathbf{H}_\theta$ 表示水平视点的位姿; $\mathbf{V}_\theta$ 表示垂重视点的位姿。

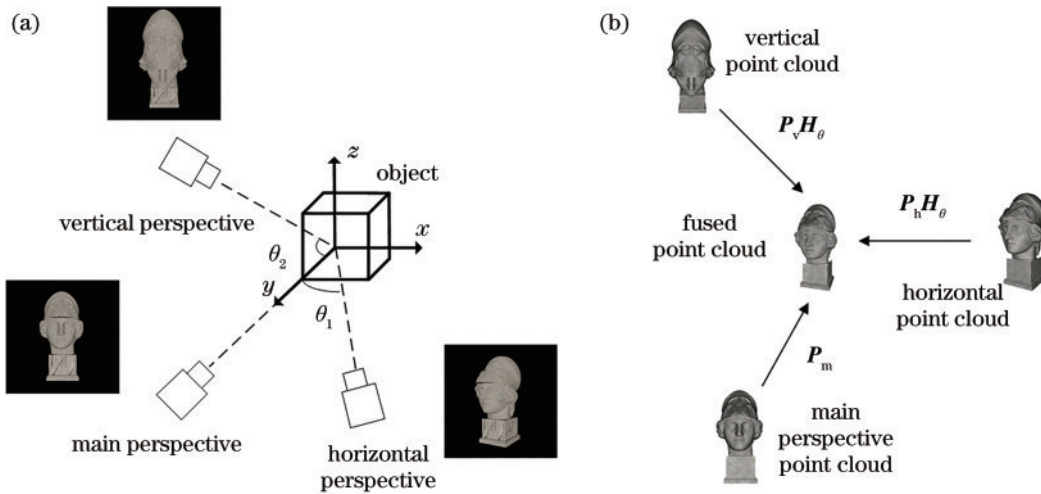


图4 多视点融合。(a)采样示意图;(b)点云融合过程

Fig. 4 Multi-angle fusion. (a) Sampling schematic; (b) point cloud fusion

在实际应用中,由于光场相机的旋转角度 $\theta_1$ 难以测量,通常固定光场相机并使用步进电机控制的高精度旋转台对物体进行指定角度的旋转<sup>[17]</sup>实现水平视点采集。由于水平方向进行 $360^\circ$ 采集已经覆盖了目标的大部分区域,垂直方向的视点只需覆盖顶部和底部,即 $\theta_2$ 为 $90^\circ$ 和 $270^\circ$ 两个特殊视点即可。

在获得不同视角点云位姿后,通过下式将多视点

的点云融合,如图4(b)所示:

$$\mathbf{P} = \mathbf{P}_v \mathbf{V}_\theta \cup \mathbf{P}_h \mathbf{H}_\theta \cup \mathbf{P}_m, \quad (12)$$

式中: $\mathbf{P}_v$ 表示垂重视点的点云; $\mathbf{P}_h$ 表示水平视点的点云; $\mathbf{P}_m$ 表示主视点的点云; $\cup$ 表示单视点下点云的集合; $\mathbf{P}$ 表示融合后的点云。

在多视点点云融合过程中,不同视点的点云存在重叠部分,如图5(a)所示。这些点被视为物体表面的

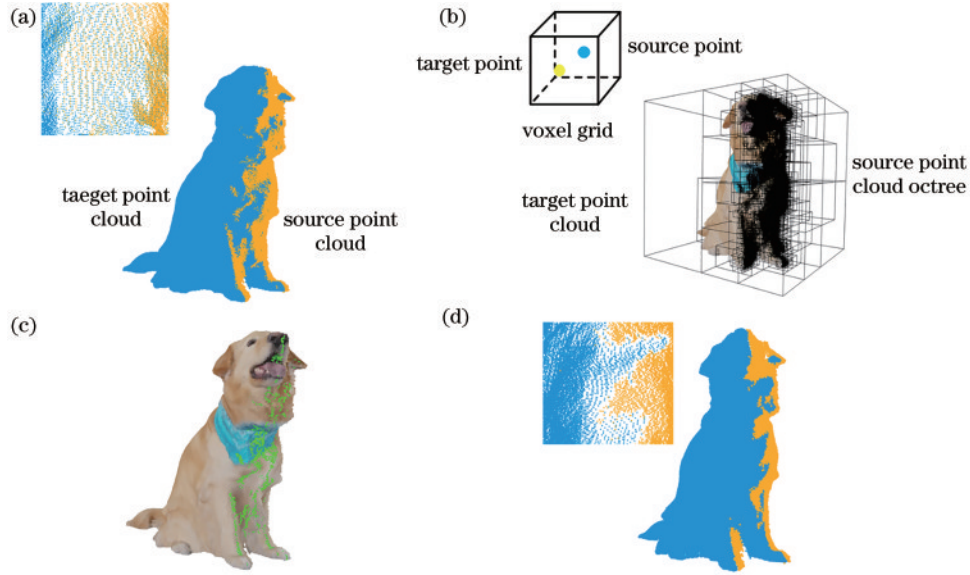


图 5 重叠部分点云处理示意图。(a)重叠点云图;(b)建立点云八叉树进行体素化;(c)重复点;(d)重复点处理效果

Fig. 5 Schematic diagrams of point cloud processing for overlapping parts. (a) Overlapping point clouds; (b) point cloud octree for voxelization; (c) overlapping points; (d) overlapping point processing effect

重复采样点,使点云重叠区域的密度大于单视点云密度,给物体重建带来了误差。使用一种基于体素的重复点云处理方法来解决此问题,如图 5(b)所示,对于源点云,建立其对应的八叉树,并将目标点云和源点云中存在于同一体素的点定义为重叠点。此时两个点云的重叠点在同一体素中,使用体素下采样或者 Laplace 平滑<sup>[19]</sup>均可有效融合重复点云。处理后的点云,重叠区域和不重叠区域的密度保持一致,提升了重建物体的质量。

### 2.4 飞行像素移除

为进一步提升三维重建质量,提高点云在深度不连续区域的精度,提出一种基于统计式的离群点移除策略消除飞行像素。飞行像素的产生原理如图 6(a)所示,像素  $p_3$  同时记录来自多条路径的光,当像素看到前景物体的边缘和非遮挡的背景物体时,采集的强度/颜色将是前景和背景强度/颜色的混合。飞行像素被视为错误的采样点,从而带来了误差。



图 6 飞行像素示意图。(a)飞行像素的产生原理;(b)飞行像素引起的误差;(c)去除飞行像素后的效果  
Fig. 6 Schematics of flying pixel. (a) Generation principle of flying pixels; (b) errors brought by flying pixels; (c) effect after removing flying pixels

飞行像素的强度是前景和背景的加权和,将深度图转换为点云后,飞行像素点与其他点的距离往往更远。利用这一特性,提出一种统计式离群点移除策略,这些点满足

$$\sum_{i=1}^k \sqrt{(x_p - x_k)^2 + (y_p - y_k)^2 + (z_p - z_k)^2} \geq \delta, \quad (13)$$

式中:  $(x_p, y_p, z_p)$  表示为离点  $p$  最近  $k$  个点的空间坐标;  $\delta$  为所设的阈值。式(13)可以有效寻找所有与邻域的

其他点距离较大的点,通过调整适合的阈值  $\delta$  索引飞行像素点并移除。具体效果如图 6(c)所示,该方法计算量小,且在去除离群点的同时提升了物体的边缘质量。

## 3 实验结果与分析

### 3.1 实验设置

将所提算法与 3 种三维重建算法进行对比,分别

为普通光场三维重建<sup>[10]</sup>、clustering views for multi-view stereo (CMVS)<sup>[20]</sup>、multi-view stereo (MVS)<sup>[21]</sup>。相应地,构建了 3 种数据,具体的测试方法和对应的实验数据如表 1 所示。1) 针对 CMVS 和 MVS, 构建水平 360° 环形光场, 如图 7(d) 所示, 相机环绕旋转轴 1° 等间隔均匀采样以获取场景的 360° 信息, 一个数据集包含 360 张普通纹理图像。2) 针对纹理光场的三维重建<sup>[10]</sup>, 构建 720° 普通光场, 如图 7(c) 所示, 该数据集包含 6 个视点, 在水平和垂直方向分别构成 360° 环形采集, 每个视点为图 7(a) 所示的普通光场采集的纹理信息。光场相机焦距为 43.456 mm, 每个视点包含 7×7 共 49 张图像, 整个数据集包含 294 张视图。3) 针对所提编码光场的三维重建, 构建 720° 编码光场,

其采样方式仍如图 7(c) 所示, 但各视点为图 7(b) 所示的编码光场<sup>[22]</sup>。其光场相机参数和普通光场相同, 模板使用正弦函数的周期为 42, 整个数据集包含 294 张视图。

值得注意的是, CMVS 和 MVS 需要密集的采样视点建立视点间相关性, 当视点间距大时无法进行匹配, 故使用图 7(c) 所示的采样方法无法得到重建结果。根据 Vianello 等<sup>[14]</sup> 和宋征玺等<sup>[17]</sup> 的研究, 使用图 7(d) 所示的环形光场并采用 1° 间隔采样时可取得可接受的重建效果, 此时需 360 张采样图像, 已经高于普通光场和编码光场 294 张视图的数据量。为在尽可能接近的数据规模上比较性能, 对环形光场仅进行水平 360° 重建。

表 1 不同数据集对比

Table 1 Comparison of different datasets

Dataset	Panoramic information	Number of images	Applicable method
Conventional light field	720° (horizontal 360°+vertical 360°)	294	Method of reference [10]
Phase-encoded light field	720° (horizontal 360°+vertical 360°)	294	Proposed method
Circular light field	horizontal 360°	360	CMVS/MVS

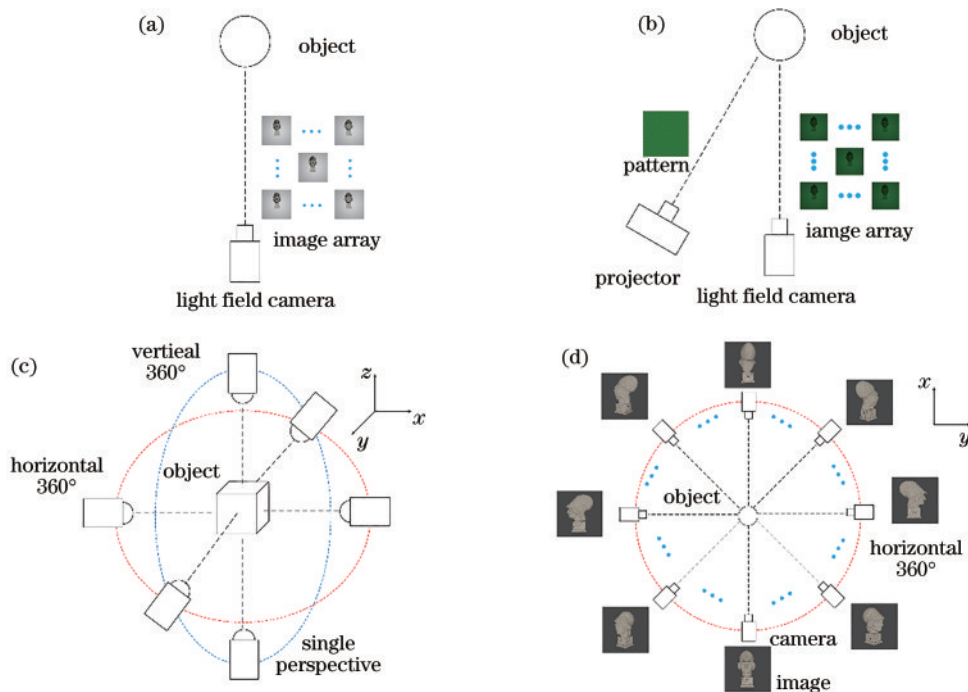


图 7 实验图。(a) 普通光场单视点光路图; (b) 编码光场单视点光路图; (c) 720° 全景重建光路图, 适用于普通光场和编码光场; (d) 360° 水平方向环形光场光路图

Fig. 7 Experimental images. (a) Conventional light field single viewpoint light path; (b) phase-encoded light field single viewpoint light path; (c) 720° multi-view light path, both applicable to conventional light field and phase-encoded light field; (d) 360° horizontal circular light field light path

### 3.2 与现有方法的精度比较

从表 1 可以看出, 采用 4 种方法在 3 组数据集上测试重建效果及生成 mesh 的误差, 结果如图 8~10 所示, 其中, 误差图颜色越深代表误差越大。MVS 和 CMVS 在环形光场数据集的实验结果如图 8~10(a)、

(b) 所示, 可见 MVS 重建结果在边缘处具有较多噪声, 而 CMVS 重建点云密度较稀疏。使用泊松表面重建<sup>[23]</sup> 将点云重建为网格后, 所得到的重建结果误差较大。同时, 使用环形光场数据集进行三维重建会丢失顶部和底部信息, 只能完成水平 360° 重建。传统光场



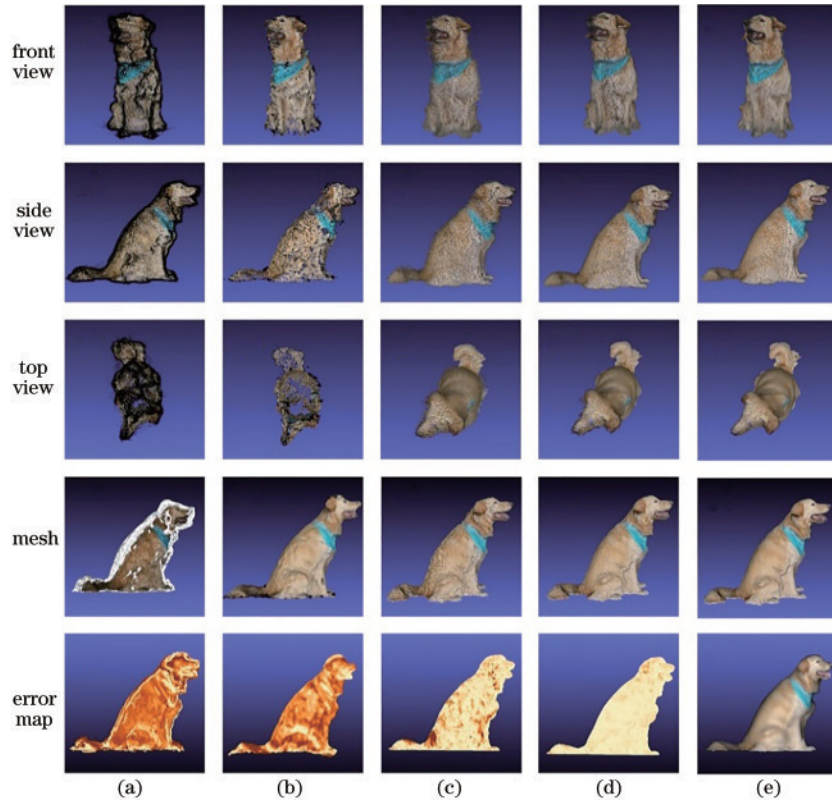


图 8 狗数据集实验数据对比。(a) MVS; (b) CMVS; (c) 普通光场的三维重建<sup>[10]</sup>; (d) 所提方法; (e) 真实值  
 Fig. 8 Experimental data comparison of dog dataset. (a) MVS; (b) CMVS; (c) conventional light field method<sup>[10]</sup>; (d) proposed method; (e) ground truth

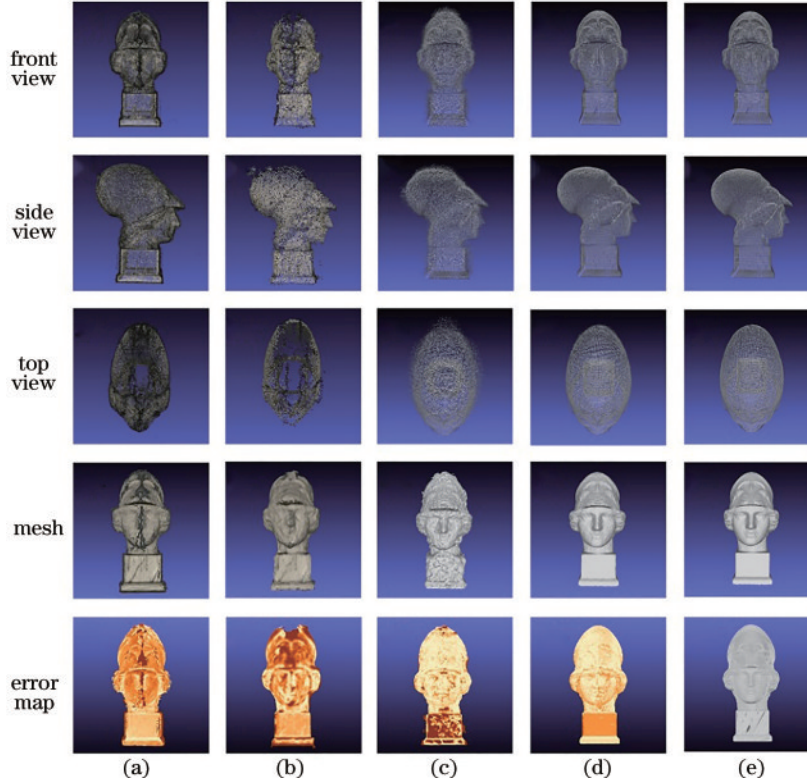


图 9 头像数据集实验数据对比。(a) MVS; (b) CMVS; (c) 普通光场的三维重建<sup>[10]</sup>; (d) 所提方法; (e) 真实值  
 Fig. 9 Experimental data comparison of head dataset. (a) MVS; (b) CMVS; (c) conventional light field method<sup>[10]</sup>; (d) proposed method; (e) ground truth

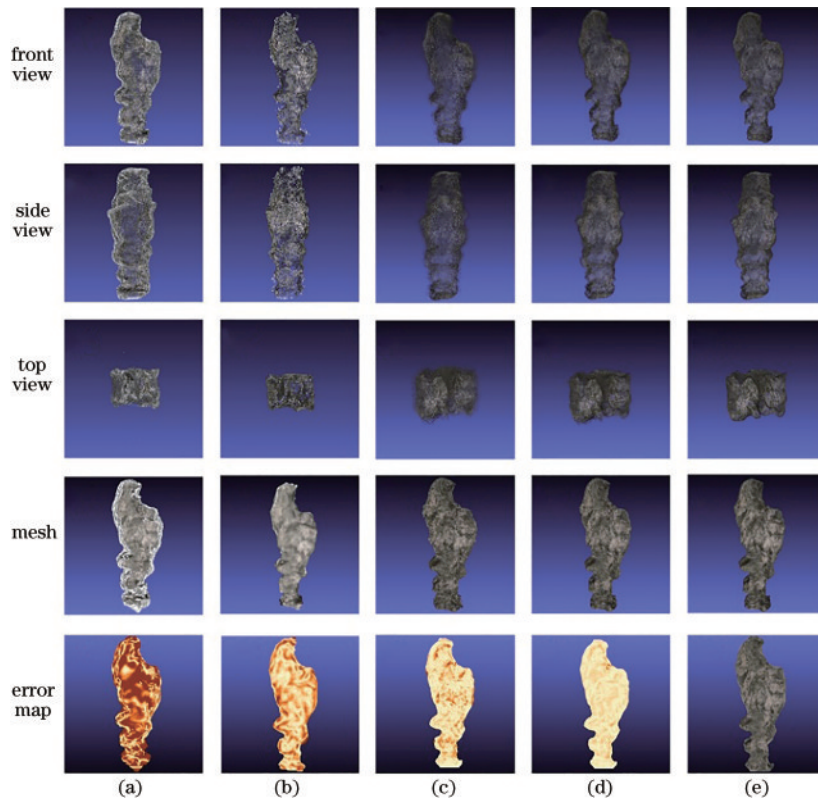


图 10 石头数据集实验数据对比。(a) MVS; (b) CMVS; (c)普通光场的三维重建<sup>[10]</sup>; (d)所提方法; (e)真实值

Fig. 10 Experimental data comparison of stone dataset. (a) MVS; (b) CMVS; (c) conventional light field method<sup>[10]</sup>; (d) proposed method; (e) ground truth

三维重建方法和所提方法的测试结果如图 8~10(c)、(d)所示,结果表明,通过不同视点求取深度值再进行融合,所得重建效果整体优于 CMVS 和 MVS,利用更少的输入图像生成了更精确的三维信息。同时从图 8~10(c)可以看出,普通光场数据集难以对弱纹理或无纹理区域进行有效重建,而从图 8~10(d)可以看出,使用

编码光场可以得到更好的点云,网格的重建结果也更加平滑和准确,实现了精确的全景三维重建。

为更好呈现全景三维重建效果,选取 5 个随机角度对生成的网格与真实值进行对比,结果如图 11 所示,所提方法可以获取物体的全景信息并融合,在各个角度均能观测到物体的三维信息。

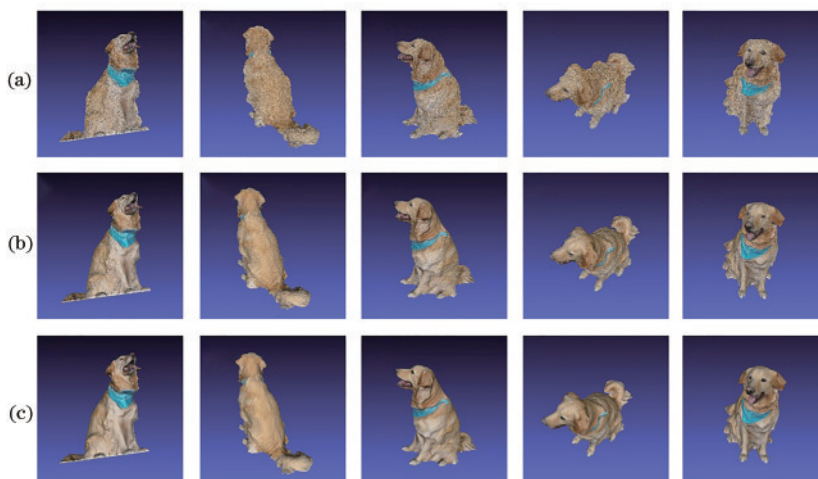


图 11 所提方法不同角度结果。(a)普通光场三维重建<sup>[10]</sup>; (b)所提方法; (c)真实值

Fig. 11 Results from different angles of proposed method. (a) conventional light field method<sup>[10]</sup>; (b) proposed method; (c) ground truth

### 3.3 光场角度分辨率对结果的影响

为分析编码光场的角度分辨率对结果的影响,即重建质量随输入数据量的变化,测试角度分辨率为

$3 \times 3$ 、 $5 \times 5$ 、 $7 \times 7$  的普通光场与编码光场所得点云的均方根误差(RMSE)<sup>[24]</sup>,结果如表 2 所示。从表 2 可以看出,随着单视点光场采样角度分辨率的增加,普通光



表 2 不同角度分辨率在不同数据集上的 RMSE

Table 2 RMSE on different datasets with different angular resolutions

RMSE for each dataset	dog	head	stone
3×3 encoded	0.309	0.182	0.327
5×5 encoded	0.123	0.102	0.155
7×7 encoded	<b>0.0669</b>	<b>0.0815</b>	<b>0.0838</b>
3×3 conventional	0.687	0.786	1.272
5×5 conventional	0.361	0.426	0.602
7×7 conventional	0.187	0.296	0.450

场和编码光场的 RMSE 不断减小,表明视点数目的增

加提供了更强的约束,从而生成更准确的三维信息。同时,在相同角度分辨率的情况下,编码光场得到的点云质量优于普通光场,即光编码提供了更准确的深度线索,生成了更准确的深度值。

除了客观指标,还比较了重建 mesh 和真实值的误差,结果如图 12~14 所示。从图中可以看出:输入光场的角度分辨率改变时,普通算法受角度分辨率的影响更大;而所提算法在输入不同尺度的光场情况下,均能得到完整的信息和精确的轮廓。此外,其结果也表明所提算法的 mesh 质量优于普通算法,重建结果更加准确且平滑,从而验证了其有效性。

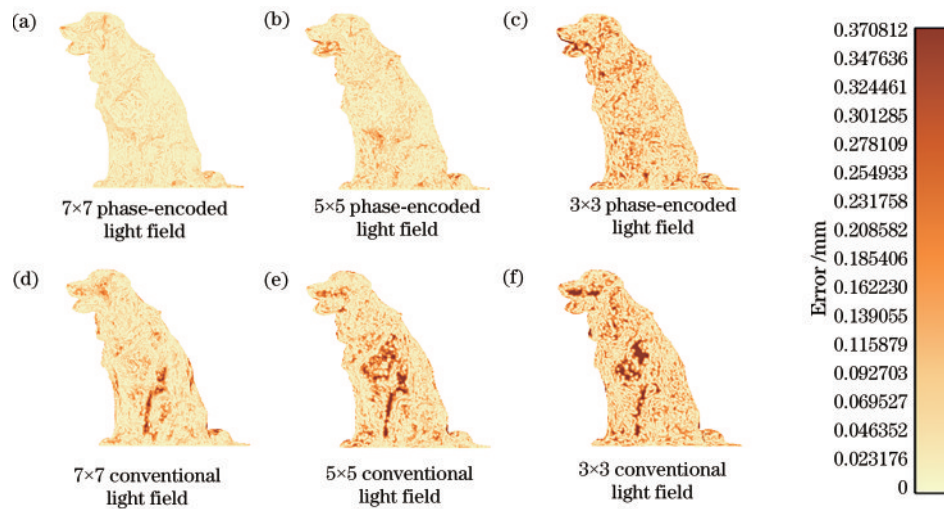


图 12 狗数据集重建效果对比。(a) 7×7 编码光场;(b) 5×5 编码光场;(c) 3×3 编码光场;(d) 7×7 普通光场;(e) 5×5 普通光场;(f) 3×3 普通光场

Fig. 12 Comparison of reconstruction effect of dog dataset. (a) 7×7 phase-encoded light field; (b) 5×5 phase-encoded light field; (c) 3×3 phase-encoded light field; (d) 7×7 conventional light field; (e) 5×5 conventional light field; (f) 3×3 conventional light field

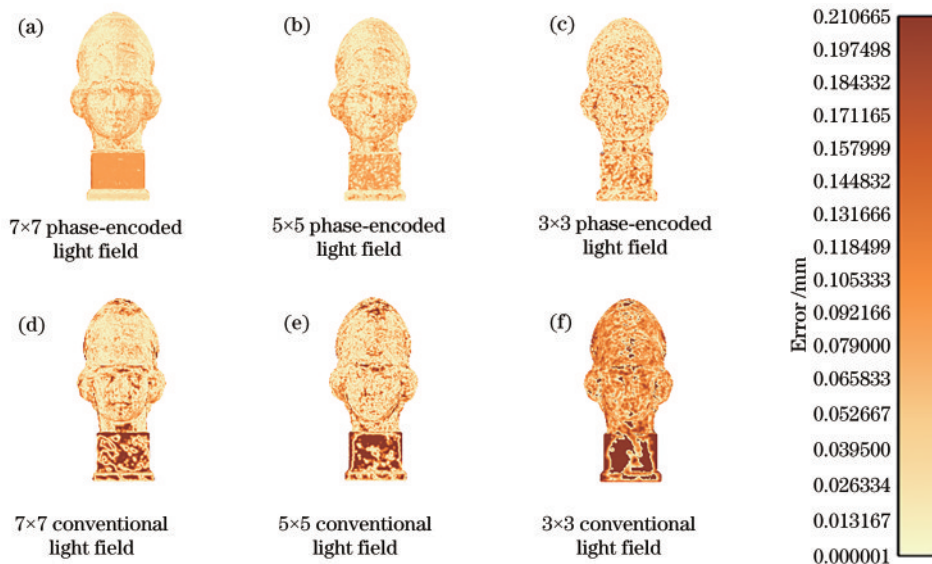


图 13 头像数据集重建效果对比。(a) 7×7 编码光场;(b) 5×5 编码光场;(c) 3×3 编码光场;(d) 7×7 普通光场;(e) 5×5 普通光场;(f) 3×3 普通光场

Fig. 13 Comparison of reconstruction effect of head dataset. (a) 7×7 phase-encoded light field; (b) 5×5 phase-encoded light field; (c) 3×3 phase-encoded light field; (d) 7×7 conventional light field; (e) 5×5 conventional light field; (f) 3×3 conventional light field

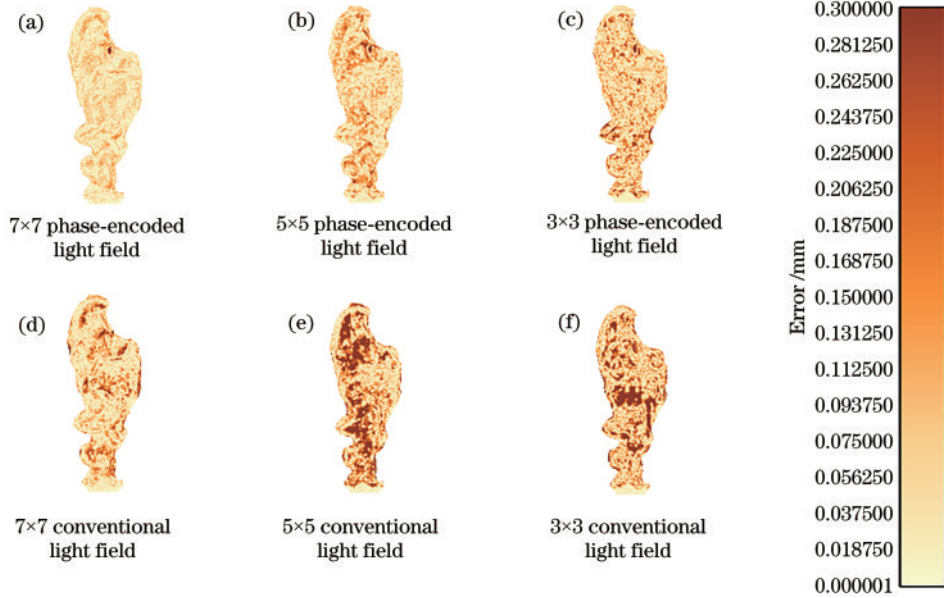


图 14 石头数据集重建效果对比。(a) 7×7 编码光场; (b) 5×5 编码光场; (c) 3×3 编码光场; (d) 7×7 普通光场; (e) 5×5 普通光场; (f) 3×3 普通光场

Fig. 14 Comparison of reconstruction effect of stone dataset. (a) 7×7 phase-encoded light field; (b) 5×5 phase-encoded light field; (c) 3×3 phase-encoded light field; (d) 7×7 conventional light field; (e) 5×5 conventional light field; (f) 3×3 conventional light field

### 3.4 重建精度影响因素分析

还探索了具体实施中不同因素对重建精度的影响, 主要包括输入图片空间分辨率、旋转角精度及模板的噪声。重建误差以 RMSE 衡量, 结果如表 3 所示, 重建 mesh 的误差如图 15 所示。输入图像的空间分辨率影响深度值的精确性, 进而影响重建结果, 测试并分析了在输入分辨率为 512×512、448×448、384×384 时的重建效果。随着空间分辨率降低, 视差和深度值精

度降低, 生成点云的误差增大。视点的旋转角  $\theta_1$  和  $\theta_2$  将影响点云的融合, 进而降低重建质量, 测试并分析了旋转角出现 1°~5° 误差时的重建效果。随着角度误差的增加, 重建结果的误差也迅速增大。模板噪声降低深度值的准确性, 降低重建质量, 对相位模板加入方差为 0.01、0.05、0.1 的高斯噪声并测试重建效果。随着噪声增强, 重建结果的 RMSE 略微增加。

表 3 不同影响因素在狗数据集上的 RMSE

Table 3 RMSE on dog datasets with different influencing factors

Influencing factor	Different spatial resolution			Rotation angle accuracy					Pattern with Gaussian noise		
	512 (proposed)	448	384	1°	2°	3°	4°	5°	0.01	0.05	0.1
RMSE	<b>0.0669</b>	0.104	0.181	0.142	0.176	0.299	0.573	0.818	0.132	0.141	0.157

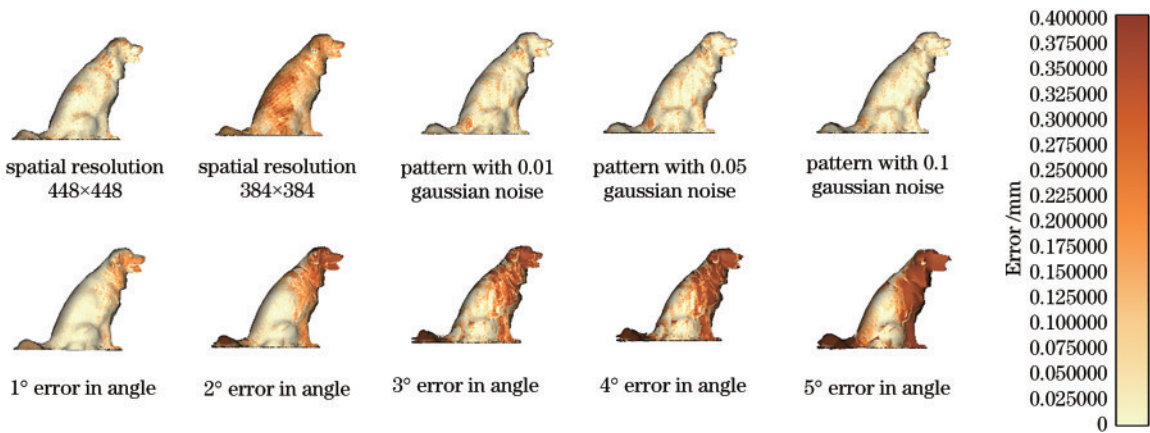


图 15 狗数据集在不同影响因素下的重建效果

Fig. 15 Reconstruction effect of dog dataset under different influencing factors

## 4 结 论

针对传统光场重建信息不全,在纹理匮乏和重复区域重建结果不精确等问题,提出一种基于多视点的编码光场三维重建的方法。通过相位解码、深度估计、多视点融合、飞行像素去除等一系列步骤进行优化,实现了物体的精确全景重建。实验结果表明,与现有方法相比,该方法可以在更少的输入情况下得到更加精确的重建结果,同时其在弱纹理或无纹理处重建的结果更加平滑。同时,还分析了角度分辨率、空间分辨率、模板噪声等因素对重建质量的影响。

### 参 考 文 献

- [1] Xiong J H, Hsiang E L, He Z Q, et al. Augmented reality and virtual reality displays: emerging technologies and future perspectives[J]. *Light: Science & Applications*, 2021, 10: 216.
- [2] 熊璐, 杨兴, 卓桂荣, 等. 无人驾驶车辆的运动控制发展现状综述[J]. *机械工程学报*, 2020, 56(10): 127-143.  
Xiong L, Yang X, Zhuo G R, et al. Review on motion control of autonomous vehicles[J]. *Journal of Mechanical Engineering*, 2020, 56(10): 127-143.
- [3] Yasuda Y D V, Martins L E G, Cappabianco F A M. Autonomous visual navigation for mobile robots: a systematic literature review[J]. *ACM Computing Surveys*, 2021, 53(1): 13.
- [4] Schönberger J L, Frahm J M. Structure-from-motion revisited[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4104-4113.
- [5] Engel J, Schöps T, Cremers D. LSD-SLAM: large-scale direct monocular SLAM[M]//Fleet D, Pajdla T, Schiele B, et al. *Computer vision-ECCV 2014. Lecture notes in computer science*. Cham: Springer, 2014, 8690: 834-849.
- [6] Levoy M, Hanrahan P. Light field rendering[C]//SIGGRAPH '96: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, August 4-9, 1996, New Orleans, LA, USA. New York: ACM Press, 1996: 31-42.
- [7] 刘晓旻, 杜梦珠, 马治邦, 等. 基于遮挡场景的光场图像深度估计方法[J]. *光学学报*, 2020, 40(5): 0510002  
Liu X M, Du M Z, Ma Z B, et al. Depth estimation method of light field image based on occlusion scene[J]. *Acta Optica Sinica*, 2020, 40(5): 0510002.
- [8] Wanner S, Goldluecke B. Variational light field analysis for disparity estimation and super-resolution[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(3): 606-619.
- [9] Zhang Y B, Lü H J, Liu Y B, et al. Light-field depth estimation via epipolar plane image analysis and locally linear embedding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(4): 739-747.
- [10] Tao M W, Hadap S, Malik J, et al. Depth from combining defocus and correspondence using light-field cameras[C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 673-680.
- [11] Williem, Park I K, Lee K M. Robust light field depth estimation using occlusion-noise aware data costs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(10): 2484-2497.
- [12] Peng J Y, Xiong Z W, Zhang Y Y, et al. LF-fusion: dense and accurate 3D reconstruction from light field images[C]//2017 IEEE Visual Communications and Image Processing, December 10-13, 2017, St. Petersburg, FL, USA. New York: IEEE Press, 2017.
- [13] Johannsen O, Sulc A, Goldluecke B. On linear structure from motion for light field cameras[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 720-728.
- [14] Vianello A, Ackermann J, Diebold M, et al. Robust Hough transform based 3D reconstruction from circular light fields[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7327-7335.
- [15] Chen C, Lin H T, Yu Z, et al. Light field stereo matching using bilateral statistics of surface cameras[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 1518-1525.
- [16] Zhang Y L, Yu P H, Yang W, et al. Ray space features for plenoptic structure-from-motion[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4641-4649.
- [17] 宋征玺, 吴淇, 王雪, 等. 基于三维霍夫变换的环形光场三维重建[J]. *西北工业大学学报*, 2021, 39(1): 135-140.  
Song Z X, Wu Q, Wang X, et al. 3D reconstruction with circular light field by using 3D Hough transformation [J]. *Journal of Northwestern Polytechnical University*, 2021, 39(1): 135-140.
- [18] Cai Z W, Liu X L, Pedrini G, et al. Accurate depth estimation in structured light fields[J]. *Optics Express*, 2019, 27(9): 13532-13546.
- [19] Vollmer J, Mencl R, Müller H. Improved Laplacian smoothing of noisy surface meshes[J]. *Computer Graphics Forum*, 1999, 18(3): 131-138.
- [20] Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(8): 1362-1376.
- [21] Schönberger J L, Zheng E L, Frahm J M, et al. Pixelwise view selection for unstructured multi-view stereo[M]//Leibe B, Matas J, Sbebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9907: 501-518.
- [22] Cai Z W, Liu X L, Peng X, et al. Structured light field 3D imaging[J]. *Optics Express*, 2016, 24(18): 20324-20334.
- [23] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction[C]//Proceedings of the Fourth Eurographics Symposium on Geometry Processing, June 26-28, 2006, Cagliari, Sardinia, Italy. New York: ACM Press, 2006: 61-70.
- [24] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. *International Journal of Computer Vision*, 2002, 47(1): 7-42.