

用于全息三维显示的数据获取方法进展

佟卓键, 桂进斌*, 胡磊, 胡先飞

昆明理工大学理学院, 云南 昆明 650500

摘要 全息显示技术可以准确恢复出包含物体全部信息的三维像, 能为使用者带来自然、真实的视觉体验, 被认为是最理想的三维显示技术。准确获取三维真实场景的数据是实现高质量全息三维显示及应用的重要条件。对三维场景数据重建的各类方法进行比较, 介绍主动式、被动式和基于深度学习的重建方法的基本原理, 分析各类方法的特点和优缺点, 综述基于深度学习的三维重建的基本方法, 探讨重建方法与全息显示技术相结合的应用前景。本文内容为全息三维显示的进一步研究提供了参考。

关键词 图像处理; 全息显示; 三维重建; 深度学习

中图分类号 O438.1

文献标志码 A

DOI: 10.3788/LOP232113

Progress on Data Acquisition Methods for Holographic 3D Display

Tong Zhuojian, Gui Jinbin*, Hu Lei, Hu Xianfei

Faculty of Science, Kunming University of Science and Technology, Kunming 650500, Yunnan, China

Abstract Holographic display technology, which is considered to be the most ideal three-dimensional (3D) display technology, can accurately recover 3D image containing all the information of the object and provide the user a natural and real visual experience. Accurate acquisition of 3D real-scene data is important for realizing high-quality holographic 3D display and application. This paper compares various methods of 3D scene data reconstruction, introduces the basic principles of active, passive, and deep-learning-based reconstruction methods, analyzes the characteristics, advantages, and disadvantages of various methods, summarizes the basic methods of 3D reconstruction based on deep learning, and discusses the application prospect of combining them with holographic display technology. This article provides a reference for the further research on holographic 3D displays.

Key words image processing; holographic display; three-dimensional reconstruction; deep learning

1 引 言

人类从现实世界中获取信息的主要方式之一是视觉, 视觉信息在所获全部信息中的占比超过 80%^[1]。传统显示技术所提供的视觉信息通常是缺乏深度信息的二维图像信息, 存在信息量不够和真实感缺乏等缺点, 难以为用户提供良好的视觉体验。三维显示技术能为用户提供更丰富的视觉信息, 从而带来更优质的视觉体验, 现头戴式三维立体显示、体三维显示和合成成像立体显示等三维显示技术已广泛应用于教育、娱乐、医疗、交通等领域^[2-14]。

全息显示技术利用干涉原理, 以干涉条纹的形式记录真实或虚拟场景物光波前的全部信息, 并根据衍

射原理能够完全重建出原物体的波前信息, 可以准确恢复出包含物体全部信息的三维像, 能够表现出和真实物体一样的深度和视差^[15-16]。因此全息三维显示技术被认为是最理想的三维显示技术之一。记录物光波前信息的干涉条纹图样叫作全息图, 按其生成方法, 可以分为光学全息图^[17]和计算全息图^[18-19]两类。光学全息图是根据光学干涉法通过搭建光路将物光波和参考光波叠加干涉后形成的干涉条纹图样, 根据记录介质的特性获得光学全息图或数字全息图; 计算全息图是在计算机中通过模拟真实场景或虚拟的物体的物光波与参考光波的干涉过程得到的数字化的全息图。全息三维显示需要得到三维物体的计算全息图。一直以来, 生成三维物体计算全息图的方法被国内外的相关

收稿日期: 2023-09-14; 修回日期: 2023-10-16; 录用日期: 2023-10-30; 网络首发日期: 2023-11-17

基金项目: 国家自然科学基金(62065010)

通信作者: jinbingui@163.com

学者广泛关注和研究,使用的物体可以是 3DS_MAX 等建模软件建模得到的虚拟模型,或通过扫描、成像等方式得到的三维坐标数据。虚拟模型缺乏真实感,且使用建模软件设计复杂的三维模型需要耗费极大的时间和精力。因此,研究准确快速的三维信息重建算法对全息三维显示具有重要的意义。

近年来三维重建方法^[20-34]可以分为两类:一类是主动式重建,该方法向目标物体发射有规律的信号,通过对比发射信号与反射信号的变化,或探测发出信号与接收到信号的时间差,计算得到目标物体的深度信息;另一类是被动式重建,它的特点是基于图像的重建,对不同角度下相机拍摄所得的图像序列中的匹配点对使用三角法计算获得三维物体上各点的深度,进而获得物体三维模型。主动式重建具有精度高、模型表面完整等优点;被动式重建从图像中重建三维模型,对实验设备要求低,有助于研究者开展研究。本文主要对三维重建方法进行阐述,并分析各种重建方法的基本理论,还对用于三维重建的典型深度学习网络模型进行讨论,探讨其在全息三维显示领域中的适用范围和应用前景。

2 主动式重建

主动式三维重建是一种主动向周围环境或目标物体发射信号探测出物体深度信息,从而重建物体三维模型的技术。它利用激光、声波、电磁波等能量源向周围环境或被测目标发射信号,根据传感器捕获返回的信号来获取物体表面的深度信息,解析计算出物体的三维信息,从而重建出物体的三维模型。

主动式三维重建主要包括激光扫描法^[35-36]、飞行时间(TOF)法^[37]、结构光法^[38]、Kinect技术^[39]等。对于激光扫描法,激光测距仪发射光束到物体表面,探测器

接收返回的信号,然后该方法根据接收信号和发射信号的时间差计算出物体与激光发射器间的距离,从而获得目标物体的三维模型,具有生成模型的精度较高的优点。

TOF法根据发射器的类型可分为脉冲调制法和连续波调制法,根据测量传播时间方式可分为直接测量法和间接测量法。其原理如图1所示。TOF法通过向目标物体发送光脉冲信号,根据发射信号与接收信号的时间差来计算与目标物体的距离。TOF法采用主动红外调制光源,可在无光环境中工作,具有稳定性高、集成度高等优点,但易受环境光的干扰。

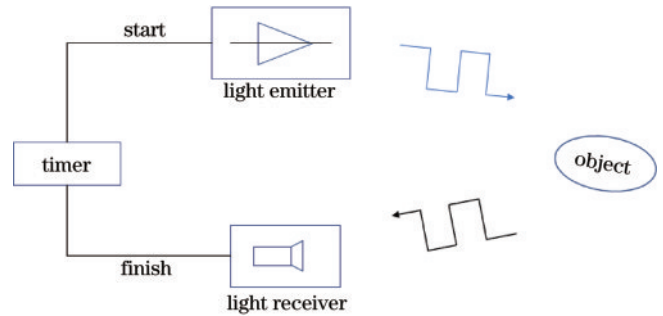


图1 飞行时间法原理图

Fig. 1 Schematic of time-of-flight method

Hoegg等^[40]使用多个TOF相机,基于图像处理单元(GPU),通过预先标定的相机系统进行配准,采用点到平面的迭代最近点进行配准,将合并后的数据集集成到GPU中,实现了快速高效的三维点云重建。微软公司2014年推出了Kinect2.0传感器,采用TOF技术来获取深度信息,从而重建出目标物体的三维点云模型。李诗锐等^[41]提出了一种基于Kinect2.0的实时三维重建系统,该系统利用单台Kinect2.0能够对物体进行实时和精确的重建,结果如图2所示。

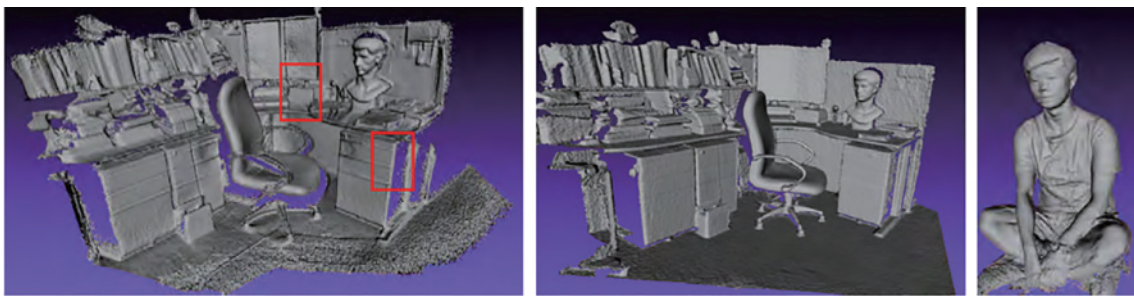


图2 Kinect2.0的场景与人体重建效果图^[41]

Fig. 2 Scene and human reconstruction renderings of Kinect2.0^[41]

结构光法主要通过将特定的结构光图案投影到目标物体上,再利用视觉传感器采集目标物体的二维图像,根据三角测量原理和图像处理技术等处理获得的图像数据,计算求出目标物体的深度信息。根据投影图案的不同,结构光法可以分为:点结构光法、线结构光法、面结构光法和网络结构光法等。结构光法主要通过光源、探测器及反射点构成的三角关系计

算出目标物体的深度信息,从而实现对目标物体的三维重建。2004年,清华大学研制出了基于线结构光的多用途传感器,该传感器可以对运动的物体进行三维重建^[42]。Montusiewicz等^[43]利用基于结构光技术的三维扫描仪对历史服饰进行三维重建,该技术推动了基于文化遗产的三维模型在数字资源空间的推广,结果如图3所示。



图 3 Emir's costume 的重建效果^[43]。(a)基础模型; (b)三角形数量减少 70% 的传播模型; (c)三角形数量减少 85% 的传播模型

Fig. 3 Reconstruction of Emir's costume^[43]. (a) Basic model; (b) propagation model with a 70% reduction in the number of triangles; (c) propagation model with a 85% reduction in the number of triangles

3 多视图几何的三维重建

基于多视图几何的三维重建方法^[44-46]的基本思路是从多个不同视角对目标物体拍摄图像(值得注意的是,拍摄过程中相邻两幅图像之间要保证足够的重叠部分)上的特征点进行特征匹配,根据匹配关系解算出每张图像对应相机的位置姿态,然后进行极线校正,利用三角法测算出匹配点所对应目标点的空间坐标,就重建出了目标物体的稀疏点云模型。该方法的原理如图 4 所示。稀疏点云模型密度低,难以提供物体的细节信息,因此需要对稀疏点云模型使用 PatchMatch^[47]等稠密重建算法,即可获得目标物体的稠密点云模型,再通过 Delaunay 三角化等表面重建(也称网格重建)方法根据点云模型重建出三维网格模型,就得到了目标物体的表面结构信息;还可以通过纹理贴图技术对网格模型进行处理,得到彩色纹理网格模型,以提高模型的真实感,有助于使用者在全息三维显示过程中获取更良好的视觉体验。

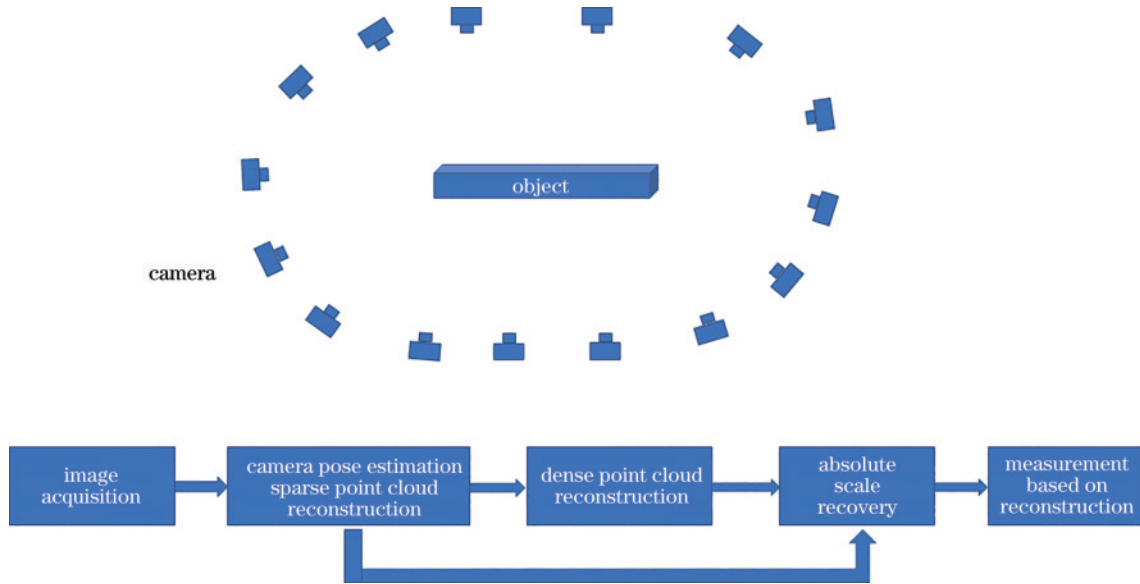


图 4 多视图三维重建示意图

Fig. 4 Diagram of multi-view 3D reconstruction

稠密重建中两种经典的算法是 PatchMatch 和 semi-global matching (SGM) 算法,相关研究工作主要是在这两种算法的基础上进行优化拓展的。2009 年 Barnes 等^[48]最先提出了 PatchMatch 算法,起初该算法是应用在图像间相似图像块匹配上的;2011 年 Bleyer 等^[49]将 PatchMatch 算法应用在立体视觉匹配上。PatchMatch 是一种基于图像编辑原理提出的算法,通过找寻同名特征点偏移量,来高效地重建出物体的深度图。2005 年 Hirschmuller^[50]提出了一种 SGM 算法,该算法使用互信息来评价匹配代价,组合多个一维的约束近似一个全局的二维平滑约束,依据最小代价值,计算得到像素的最终视差,逐个像素进行匹配,最终得

到物体的视差图。SGM 算法的匹配效率较全局立体匹配算法有明显提高,因此受到许多研究者的关注。

表面重建分为显示重建和隐式重建两种。显示重建将各点作为三角面网格的顶点,直接三角化点云模型中的各点来得到物体的网格模型。1934 年 Delaunay 提出了 Delaunay 三角剖分,它满足“最大-最小角”优化准则,即所有最小内角之和最大,同时还满足空圆性、唯一性、区域性、最优性,具有凸多边形的外壳等特性^[51]。1978 年 Green 等^[52]实现了二维空间的 Voronoi 图的计算及 Delaunay 三角化,之后 Bowyer^[53]和 Watson^[54]将二维 Delaunay 三角化推广到高维。它通过构造一个包含所有散点的超级三角形,放入三角

形链表并依次将散点插入链表,再删除所有外接圆包含该插入点的三角形(影响三角形)的公共边,连接该插入点与影响三角形的全部顶点,就得到了一个局部三角形,直到所有散点都插入链表并根据优化准则进行优化后即可得到完整的三角网格模型。Delaunay 三角化是很常见且非常通用的三角网格生成算法之一,它生成的三角网格在满足以上特性的同时还能够调控网格质量和尺寸,因此广泛地应用于三维模型的表面生成中。

隐式重建是一种先用曲面来拟合点云数据,再将拟合曲面离散化为网格模型的方法。2006年Kazhdan等^[55]提出了泊松曲面重建算法,通过求解泊松方程来拟合隐式曲面。设点云模型中的所有数据点 p_i 都分布在三维空间中的曲面 M 上,要找到一个近似于曲面 M 的曲面 M' ,曲面 M' 可由隐函数 f 表示,其中 f 满足 $f(p_i)=0, i=1, \dots, n$;选取指示函数 χ ,当数据点在模型外部和表面时,令 $\chi(p_i)=0$,数据点在模型内部时, $\chi(p_i)=1$;将输入的带有法向量的点云数据视为一个

向量场 \mathbf{V} ,该向量场仅在表面上存在非零值,由表面法线所确定,可以将指示函数的构建转为求解指示函数梯度场和向量场 \mathbf{V} 差值的绝对值最小问题;最终通过求解泊松方程来计算指示函数。

为了获得具有真实感的模型,表面重建后得到的网格模型需要使用纹理贴图技术生成彩色贴图网格。纹理贴图技术将二维图像映射到三维几何模型上,生成物体的色彩信息。彩色网格模型主要分为两类:一类是彩色顶点网格模型,这类模型的颜色信息集中在网格顶点,各三角形面的颜色信息是由顶点的颜色插值得到的,模型网格的颜色分辨率取决于顶点分辨率;另一类是彩色贴图网格模型,其三角形面的颜色信息由对应图像的三角形赋值得到,它的颜色分辨率受图像颜色分辨率的影响。在给模型贴图的过程中存在两个问题:一是要为每个三角面选择最适合的视图;二是不同视角下图像的亮暗程度不同,同时可能存在色差,因此需要通过颜色融合处理消除色差。图5显示的是本文对实际物品(日历和墨水盒)组合的场景进行三维重建得到的模型。

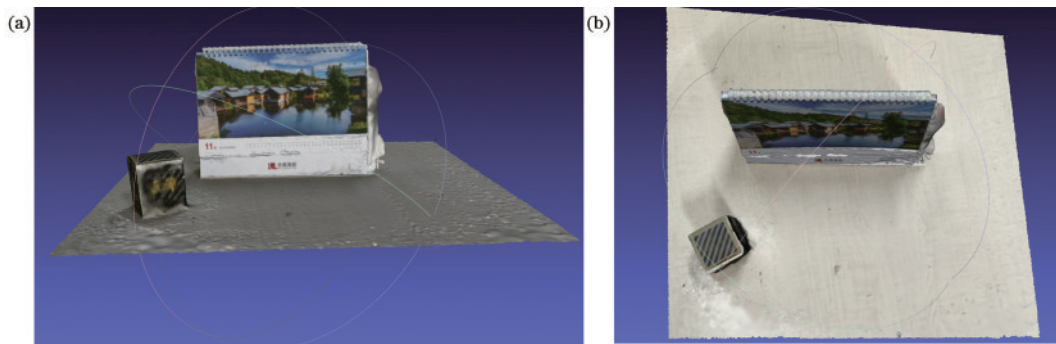


图5 实际物体重建效果。(a)物体主视图;(b)物体俯视图

Fig. 5 Actual object reconstruction effect. (a) Main view of the object; (b) top view of the object

基于多视图几何的三维重建方法仅利用拍摄的图像就能够获得完整的、与建模软件重建质量相近的三维网格模型,在实际应用中操作较为简便,作为一种全息三维显示的数据获取方法,有助于研究实际物品的全息三维显示。

4 基于深度学习的三维重建

深度学习(DL)技术是机器学习中重要的研究领域之一,通过人工神经网络对输入信息进行提取和筛选,可以实现端到端的深度学习体系结构。随着深度学习技术的发展,深度学习算法对数据提取和筛选的精度已经远超传统的机器学习算法,它可以自动且高准确率地筛选数据,提取数据的特征信息。近些年,许多优秀的深度学习模型涌现出来,并广泛应用于图像处理、自动驾驶等多个领域^[56-62]。其中卷积神经网络(CNN)以强大的图像特征提取能力和分类能力广泛应用于计算机视觉领域。由于卷积神经网络取得了良好的进展,研究者们开展了深度学习在立体重建领域的研究。

4.1 MVSNet

2018年香港科技大学的Yao等^[63]提出了一种基于端到端训练方式的MVSNet,开创性地将深度学习与多视图重建相结合。MVSNet结构示意图如图6所示。MVSNet将输入的不同视图分别作为参考视图,对于任意一张参考视图,其邻近的视图称为源视图,源视图的数目可以自行设置,在MVSNet中将这个参数设置为3,即1张参考视图加上2张邻近的源视图一起输入到网络。MVSNet的特征提取网络采用8层二维卷积神经网络,第3层和第6层的卷积步长设置为2,将特征网络分为3个尺度。特征提取网络的参数共享权重,输入的视图经过这个具有共享权重的特征提取网络生成对应的特征图,这源视图的特征图通过单应性变换形成参考视图空间内的特征图。

二维特征提取网络的输入是 N 张图像,输出是 N 个32通道的特征图,输出经过下采样后分辨率是原图的 $1/4$ 。虽然提取到的特征图的分辨率为原图的 $1/4$,但每个剩余像素的邻域信息已经被编码到了32通道

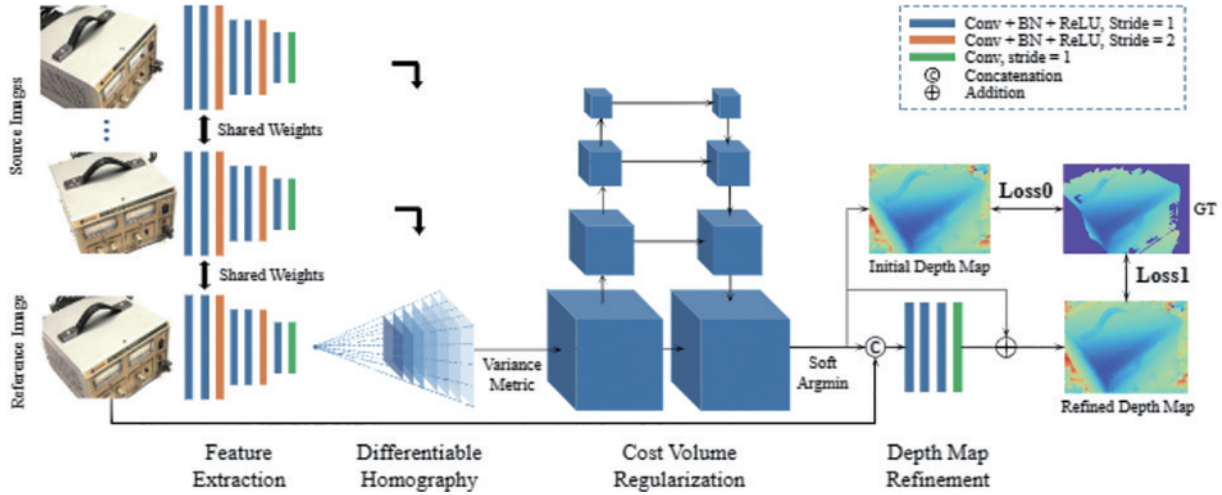


图 6 MVSNet 结构示意图^[63]

Fig. 6 Diagram of MVSNet structure^[63]

的特征描述符中,因此不会丢失细节信息。相比在原始图像上进行特征匹配的传统方法,MVSNet在重建特征图质量上有着明显的提高。得到特征图后,MVSNet通过将其他源视图的特征通过可微的单应性矩阵校正为参考图像相机前的平行平面,并将其映射在参考图像相机视锥中的 256 个深度平面上,在参考相机视锥上构建代价体。作者采用了基于方差的逐像素求代价体的方式将 N 个特征体聚合为一个代价体。接着 MVSNet 使用多尺度的三维 CNN 对代价体进行正则化,所使用的四维尺度网络类似于 3D 版本的 U-Net 结构。经过正则化后,网络通道数由 32 降为 1,再进行概率归一化,所得到的概率体代表每个像素处每层深度的概率。传统方法直接采用概率最大的深度值作为深度信息,而 MVSNet 将计算得到的深度值的数学期望作为初始深度图。最后,MVSNet 通过将初始

深度图和原始图像串连成一个 4 通道的输入,利用神经网络学习深度残差,并将其添加回初始深度图,从而生成更精细的深度图。对目标物体所有拍摄的视图进行处理得到深度图,再通过数据融合重建出目标物体的三维模型。

4.2 R-MVSNet

MVSNet 在三维重建方面有着远超传统多视图重建方法的性能,而且重建效率也提升了许多倍,但在对代价体进行正则化时采用的三维卷积神经网络会造成非常大的 GPU 消耗。这种 GPU 高消耗问题限制了重建过程中的深度估计范围和精度,且难以训练高分辨率的图像集。

对此,香港科技大学的 Yao 等^[64]提出了基于递归神经网络的 R-MVSNet 深度学习框架,在 MVSNet 的基础上解决了正则化过程中 GPU 消耗大的问题。如图 7

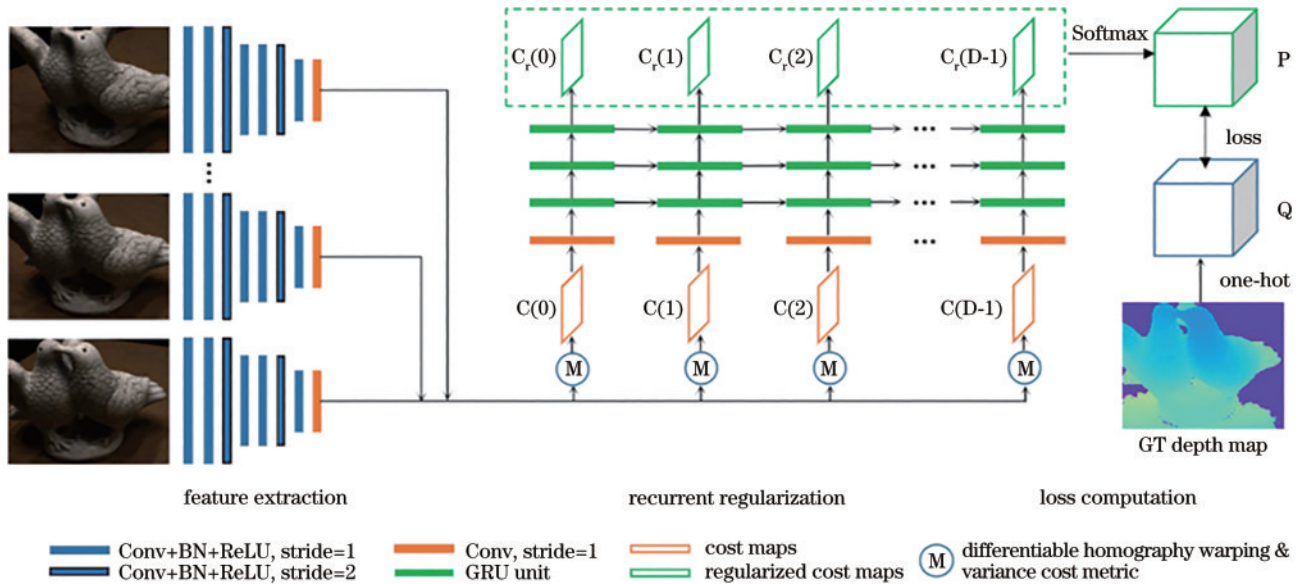


图 7 R-MVSNet 网络结构示意图^[64]

Fig. 7 Diagram of R-MVSNet structure^[64]

所示,不同于一次性正则化整个代价体的 MVSNet, R-MVSNet 使用门控循环单元(GRU)沿着深度方向依次正则化二维代价体,从而极大减少了内存消耗。GRU 是一种循环神经网络结构,和 long-short term memory(LSTM)一样,设有状态传递和“遗忘”机制,便于顺序处理数据。R-MVSNet 将整个代价体沿着深度方向分割成多个代价图,采用 GRU 依次对所有代价图进行过滤,在高效估计深度图的同时降低内存消耗。

4.3 CVP-MVSNet

2020 年 Yang 等^[65]提出了用于多视角图像深度估计的 CVP-MVSNet 深度学习框架,该框架和 Chen 等^[66]提出的 Point-MVSNet 相似,以分辨率由粗到细的方式构建代价金字塔,可以得到更为紧凑、轻量级的网络,如图 8 所示。假设 I_0 为参考图片, $\{I_i\}_{i=1}^N$ 表示 N

张相邻的 source 图片。在以往的工作中,特征提取过程中一般使用高分辨率图像提取多尺度的特征图,输出低分辨率的深度图。由于低分辨率图像包含足够的信息,CVP-MVSNet 设计了一个特征提取金字塔网络,采用低分辨率图像估计低分辨率的深度图,在保证重建质量的同时减少了内存消耗。在参考视图中构建深度估计的代价体时,以往的方法通常以固定的分辨率构建单一的代价体,会导致极大的内存消耗,限制了高分辨图像的重建。因此 CVP-MVSNet 构建代价体金字塔时,首先利用最低分辨率的图像和参考图像所对应的相机的前平行平面的均匀采样为基础,建立一个用于低分辨率深度图估计的代价体,接着通过深度残差优化低分辨率的深度估计得到更高分辨率、更精细的深度图。与 Point-MVSNet 相比,CVP-MVSNet 在保证重建质量和精度的同时,提升了重建的效率。

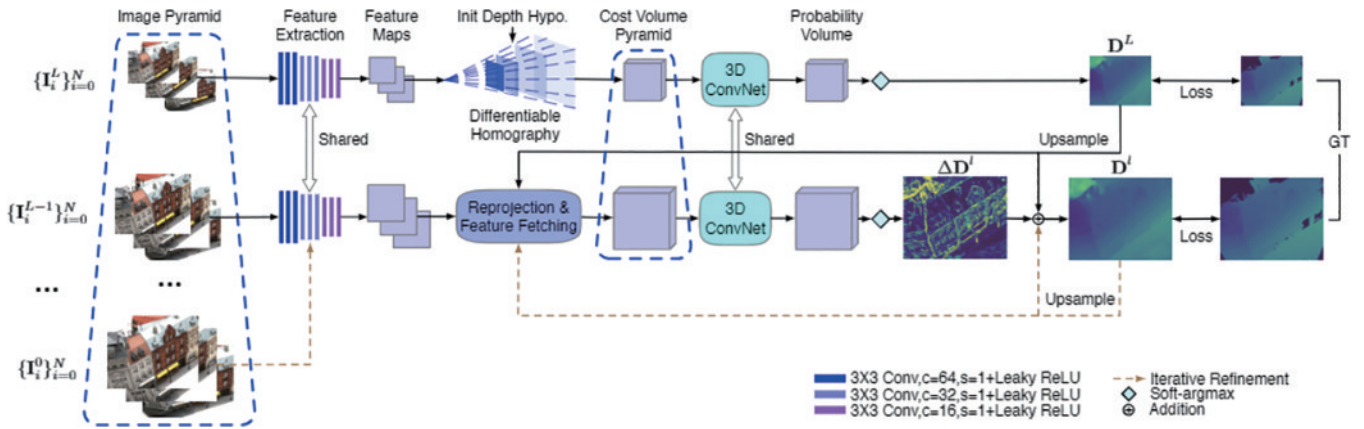


图 8 CVP-MVSNet 结构示意图^[65]

Fig. 8 Diagram of CVP-MVSNet structure^[65]

基于深度学习的三维重建方法通过端到端的学习方式,将深度学习和多视图三维重建相结合,根据不同视角图像作为输入,能够高效地生成目标物体高精度、高完整度的三维模型。根据重建得到的模型,可以生成计算全息图。由于深度学习模型性能强大,建议将计算全息图的生成与深度学习三维重建相结合,端到端地生成三维物体计算全息图,即输入目标物体的多视角图像,输出目标物体计算全息图。

4.4 基于深度学习的三维重建模型效果分析

基于深度学习的三维重建方法对计算力和内存有很高的成本要求,大大限制了重建质量和效率。许多研究人员开始研究在有限资源的条件下提升重建质量和效率的方法^[67-86]。2020 年 Gu 等^[67]提出了 CasMVSNet,在由粗到精的深度推断中构建三级代价体,并缩小深度假设范围,实现高分辨率深度图估计,从而提高重建精度。Cheng 等^[68]提出了一种基于自适应分辨率代价体的深度估计方法,通过上一层的像素预测的不确定性来得到每层的深度假设空间。除了模型的重建质量,重建效率也是现实场景中重建的重要因素。Yu

等^[70]提出的 Fast-MVSNet 可以实现快速和准确的多视图深度估计。Xu 等^[71]提出的 PVSNet 在计算聚合的代价体之前估计不同源图像的像素可见性信息,基于可见性信息,将两个视图的代价体聚合为一个加权的代价体,减少了重建过程中相邻图像包含强烈的视图变化时的影响。Weilharter 等^[73]提出了一种具有金字塔编码器-解码器结构的卷积神经网络(HighRes-MVSNet),它通过更小的分辨率编码图像特征,大大降低了内存需求。DTU 数据集是一个包括受控实验室环境中的 128 个场景的大型数据集,使用结构光扫描仪对物体进行多角度拍摄,获得每个视角下相机内和外参数。每个场景都有 49 或 64 个相机位置,获得分辨率为 1200×1600 像素的 7 种不同照明条件下的 RGB 图像。它涵盖了多种室内的对象和材料,因此 DTU 数据集适合用于训练和测试深度学习三维重建方法。各种重建方法在 DTU 数据集上的效果对比总结如表 1 所示。

准确性(accuracy)和完整性(completeness)是用于评估重建结果和真实物体之间一致性的度量指标。其

表1 不同基于深度学习的三维重建方法在DTU数据集上的测试效果

Table 1 Test effects of different 3D reconstruction methods based on deep learning on DTU dataset unit: mm

Method	Accuracy	Completeness	Overall score
MVSNet ^[63]	0.396	0.527	0.462
R-MVSNet ^[64]	0.383	0.452	0.417
Point-MVSNet ^[66]	0.342	0.411	0.376
CVP-MVSNet ^[65]	0.296	0.406	0.351
CasMVSNet ^[67]	0.325	0.385	0.355
UCSNet ^[68]	0.338	0.349	0.344
BP-MVSNet ^[69]	0.333	0.320	0.327
Fast-MVSNet ^[70]	0.336	0.403	0.370
PVSNet ^[71]	0.337	0.315	0.326
CER-MVS ^[72]	0.359	0.305	0.332
HighRes-MVSNet ^[73]	0.354	0.393	0.373
EPP-MVSNet ^[74]	0.413	0.296	0.355
AACVP-MVSNet ^[75]	0.357	0.326	0.341
PatchmatchNet ^[76]	0.427	0.277	0.352
DRI-MVSNet ^[77]	0.432	0.327	0.379
IterMVS ^[78]	0.373	0.354	0.363
ASPPMVSNet ^[79]	0.334	0.360	0.347
UniMVSNet ^[80]	0.352	0.278	0.315
NP-CVP-MVSNet ^[81]	0.356	0.275	0.315
ADIM-MVSNet ^[82]	0.344	0.298	0.321
MVSFormer ^[83]	0.327	0.251	0.289
JDACS ^[84]	0.571	0.515	0.543
RC-MVSNet ^[85]	0.396	0.295	0.345
KD-MVS ^[86]	0.359	0.295	0.327

中:准确性表示重建结果和真实点云的接近程度,即每个重建点到真实点云中最近点绝对距离的均值或中值;完整性度量了重建模型对真实物体的覆盖程度,即从真实点云到重建点云的绝对距离的均值或中值,较高的完整性表示重建模型能够较好地恢复真实场景中的点云信息。将准确性和完整性的平均值作为总分(overall score)来表示重建模型的整体质量,总分值越低表示重建效果越好。

5 不同类型数据生成计算全息图的适用性分析

讨论前面各种重建技术所得数据应用于生成计算全息图的适用性。主动式和被动式重建都可以得到包含物体表面空间坐标的点云模型,并可以进一步通过表面重建算法得到网格模型。目前生成计算全息图的方法主要有点源法、面元法和层析法。点源法的核心思想是将三维物体采样离散为许多个点,把物光波视为这些点发出光波的叠加;面元法的思想是将三维物

体分割为许多不同形状的面元,把物光波视为这些面元发出光波的叠加;层析法的基本思想是将三维物体分解为一系列平行于全息图的平面,每个平面作为一个独立的计算单元,然后利用平行平面之间的衍射计算来获得三维物体的全息图。

“虚实结合”是未来全息显示技术应用的趋势,无论是虚拟现实、增强现实还是元宇宙等方面的发展都离不开三维显示技术,这不仅需要对工业设计领域从业者建模的虚拟物体进行三维显示,对实际环境和物体的三维显示也存在很大需求。对于主动式三维重建、传统的多视图三维重建以及深度学习三维重建方法得到的物体模型,都可以采用点源法、面元法、层析法以及深度学习等方法生成计算全息图^[87-98]。然而先从主动式三维重建和传统的多视图三维重建方法得到物体的三维数据,再利用三维数据生成全息图的成本很高。深度学习直接生成全息图具有效率高、计算成本低等优势,深度学习三维重建得到的三维数据不仅可以运用得更加灵活,而且能通过渲染、纹理添加和光照调节等操作,得到不同光照条件下具有不同纹理的物体模型。进一步,这些经过操作的物体模型能够被用于深度学习生成全息图,实现多样化的全息显示效果。因此,将深度学习三维重建方法与深度学习生成计算全息图相结合,即从二维彩色图像计算出深度信息,再根据二维彩色图形和深度信息生成计算全息图的方案具有很大的应用前景和研究价值。

6 结束语

阐述了目前主要的三维重建方法,分别对主动式重建、基于多视图几何的重建方法和基于深度学习的重建方法的基本理论和效果进行分析。基于多视图几何的重建方式通过立体匹配、稠密重建、表面重建和纹理贴图来实现真实感三维模型的生成,重建质量较好,但在弱纹理区域存在匹配歧义问题,会导致模型完整度下降。基于深度学习的重建方式通过引入语义信息,可以得到完整度高、精度高的三维模型,同时提高了模型重建效率。近年来,在三维重建领域,研究者构建出许多性能强大的深度学习模型,如果将深度学习相关算法和计算全息图生成技术结合,可以为全息三维显示技术提供新的研究思路。

全息三维显示作为一种理想的三维显示技术,是三维显示领域中的研究热点。随着计算机技术、图像处理技术等相关研究领域的发展,全息三维显示技术不断朝着分辨率更高、大场景、高效便捷的方向进步,直到无论虚拟物品还是实际物体都可以得到高效、高质量的显示。全息三维显示技术将在医疗卫生和游戏娱乐等多个领域为用户带来优质的视觉体验。

参 考 文 献

- [1] Marieb E N, Hoehn K N. Human anatomy & physiology [M]. 11th ed. London: Pearson, 2018.
- [2] 曾吉勇. 折反射全景立体成像[D]. 成都: 四川大学, 2003.
Zeng J Y. Catadioptric panoramic stereo imaging[D]. Chengdu: Sichuan University, 2003.
- [3] 徐韡, 刘向东, 刘旭, 等. 基于 3DS MAX 信息源的体三维显示技术[J]. 浙江大学学报(工学版), 2005, 39(11): 1723-1726.
Xu W, Liu X D, Liu X, et al. Volumetric 3D display technique based on image information derived from 3DS MAX[J]. Journal of Zhejiang University (Engineering Science), 2005, 39(11): 1723-1726.
- [4] 邢建芳, 龚华军, 沈春林, 等. 基于数字显微镜和旋转扫描技术的体三维显示器[J]. 光电子·激光, 2008, 19(8): 1011-1015.
Xing J F, Gong H J, Shen C L, et al. A volumetric display system based on DMD and swept volume technique[J]. Journal of Optoelectronics·Laser, 2008, 19(8): 1011-1015.
- [5] 谢小燕, 刘旭, 林远芳. 体三维显示系统中基于几何近似法的灰度偏差研究[J]. 光学学报, 2009, 29(6): 1508-1513.
Xie X Y, Liu X, Lin Y F. Study on gray scale deviations in three-dimensional volumetric display system based on geometric approximation[J]. Acta Optica Sinica, 2009, 29(6): 1508-1513.
- [6] 于瀛洁, 蔡明义, 张之江. 柱透镜光栅自由立体显示中几何参数间关系分析[J]. 光电子技术, 2010, 30(1): 10-15.
Yu Y J, Cai M Y, Zhang Z J. Analysis for geometric parameter relation of lenticular sheet based autostereoscopic[J]. Optoelectronic Technology, 2010, 30(1): 10-15.
- [7] 黄涛, 朱秋东, 王涌天, 等. 立体显示中的垂轴放大率[J]. 光子学报, 2010, 39(1): 116-118.
Huang T, Zhu Q D, Wang Y T, et al. Paraxial magnification of stereoscopic display[J]. Acta Photonica Sinica, 2010, 39(1): 116-118.
- [8] Kleinberger P, Kleinberger I D. Systems for three-dimensional viewing and projection: US20020030888[P]. 2002-03-14.
- [9] Ezra D, Woodgate G J, Omar B A, et al. New autostereoscopic display system[J]. Proceedings of SPIE, 1995, 2409: 10-14.
- [10] Pastoor S, Liu J, Renault S. An experimental multimedia system allowing 3-D visualization and eye-controlled interaction without user-worn devices[J]. IEEE Transactions on Multimedia, 1999, 1(1): 41-52.
- [11] Sang X Z, Fan F C, Jiang C C, et al. Demonstration of a large-size real-time full-color three-dimensional display [J]. Optics Letters, 2009, 34(24): 3803-3805.
- [12] 田丰, 廖薇, 刘锦高. 基于 DLP 的自由立体显示系统构建[J]. 电视技术, 2010, 34(3): 31-34.
Tian F, Liao W, Liu J G. Realization of auto-stereoscopic display system based on DLP[J]. Video Engineering, 2010, 34(3): 31-34.
- [13] Bove V M. Display holography's digital second act[J]. Proceedings of the IEEE, 2012, 100(4): 918-928.
- [14] Dufaux F, Pesquet-Popescu B, Cagnazzo M. Emerging technologies for 3D video: creation, coding, transmission, and rendering[M]. Singapore: John Wiley & Sons, Ltd, 2013.
- [15] Javidi B, Ferraro P, Hong S H, et al. Three-dimensional image fusion by use of multiwavelength digital holography[J]. Optics Letters, 2005, 30(2): 144-146.
- [16] Onural L, Gotchev A, Ozaktas H M, et al. A survey of signal processing problems and tools in holographic three-dimensional television[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(11): 1631-1646.
- [17] 王丁, 李劲松. 白光编码法制作真彩色全息图[J]. 重庆大学学报(自然科学版), 2003, 26(11): 101-104.
Wang D, Li J S. 2-D real color hologram making with white light recording method[J]. Journal of Chongqing University (Natural Science Edition), 2003, 26(11): 101-104.
- [18] 潘武, 田贻丽. 光学全息的数字实现[J]. 光学精密工程, 2005, 13(S1): 15-20.
Pan W, Tian Y L. Digital implementation scheme of optical holograms[J]. Optics and Precision Engineering, 2005, 13(S1): 15-20.
- [19] 曹玉茹. 全息显示的计算原理、方法及系统实现[D]. 合肥: 安徽大学, 2006.
Cao Y R. Calculation principle, method and system implementation of holographic display[D]. Hefei: Anhui University, 2006.
- [20] 龙霄潇, 程新景, 朱昊, 等. 三维视觉前沿进展[J]. 中国图象图形学报, 2021, 26(6): 1389-1428.
Long X X, Cheng X J, Zhu H, et al. Recent progress in 3D vision[J]. Journal of Image and Graphics, 2021, 26(6): 1389-1428.
- [21] 丁少闻, 张小虎, 于起峰, 等. 非接触式三维重建测量方法综述[J]. 激光与光电子学进展, 2017, 54(7): 070003.
Ding S W, Zhang X H, Yu Q F, et al. Overview of non-contact 3D reconstruction measurement methods[J]. Laser & Optoelectronics Progress, 2017, 54(7): 070003.
- [22] 张文明, 刘彬, 李海滨. 基于双目视觉的三维重建中特征点提取及匹配算法的研究[J]. 光学技术, 2008, 34(2): 181-185.
Zhang W M, Liu B, Li H B. Characteristic point extracts and the match algorithm based on the binocular vision in three dimensional reconstruction[J]. Optical Technique, 2008, 34(2): 181-185.
- [23] Garcia E, Dugelay J L. Low cost 3D face acquisition and modeling[C]//Proceedings International Conference on Information Technology: Coding and Computing, April 2-4, 2001, Las Vegas, NV, USA. New York: IEEE Press, 2002: 657-661.
- [24] Yao Y, Li S W, Zhu S Y, et al. Relative camera

- refinement for accurate dense reconstruction[C]//2017 International Conference on 3D Vision (3DV), October 10-12, 2017, Qingdao, China. New York: IEEE Press, 2018: 185-194.
- [25] Cremers D, Kolev K. Multiview stereo and silhouette consistency via convex functionals over convex domains [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(6): 1161-1174.
- [26] Al-Temeemy A A, Al-Saqal S A. Laser-based structured light technique for 3D reconstruction using extreme laser stripes extraction method with global information extraction[J]. *Optics & Laser Technology*, 2021, 138: 106897.
- [27] Goesele M, Snavely N, Curless B, et al. Multi-view stereo for community photo collections[C]//2007 IEEE 11th International Conference on Computer Vision, October 14-21, 2007, Rio de Janeiro, Brazil. New York: IEEE Press, 2007.
- [28] Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(8): 1362-1376.
- [29] Campbell N D F, Vogiatzis G, Hernández C, et al. Using multiple hypotheses to improve depth-maps for multi-view stereo[M]//Forsyth D, Torr P, Zisserman Z. *Computer vision-ECCV 2008. Lecture notes in computer science*. Heidelberg: Springer, 2008, 5302: 766-779.
- [30] Sinha S N, Mordohai P, Pollefeys M. Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh [C]//2007 IEEE 11th International Conference on Computer Vision, October 14-21, 2007, Rio de Janeiro, Brazil. New York: IEEE Press, 2007.
- [31] Vogiatzis G, Hernandez Esteban C, Torr P H S, et al. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(12): 2241-2246.
- [32] Schönberger J L, Zheng E L, Frahm J M, et al. Pixelwise view selection for unstructured multi-view stereo[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9907: 501-518.
- [33] Galliani S, Lasinger K, Schindler K. Massively parallel multiview stereopsis by surface normal diffusion[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 873-881.
- [34] Tola E, Strecha C, Fua P. Efficient large-scale multi-view stereo for ultra high-resolution image sets[J]. *Machine Vision and Applications*, 2012, 23(5): 903-920.
- [35] Kraus K, Pfeifer N. Determination of terrain models in wooded areas with airborne laser scanner data[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 1998, 53(4): 193-203.
- [36] Göbel W, Kampa B M, Helmchen F. Imaging cellular network dynamics in three dimensions using fast 3D laser scanning[J]. *Nature Methods*, 2007, 4(1): 73-79.
- [37] Park J, Kim H, Tai Y W, et al. High quality depth map upsampling for 3D-TOF cameras[C]//2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE Press, 2012: 1623-1630.
- [38] Rocchini C, Cignoni P, Montani C, et al. A low cost 3D scanner based on structured light[J]. *Computer Graphics Forum*, 2001, 20(3): 299-308.
- [39] Khoshelham K, Elberink S O. Accuracy and resolution of Kinect depth data for indoor mapping applications[J]. *Sensors*, 2012, 12(2): 1437-1454.
- [40] Hoegg T, Lefloch D, Kolb A. Time-of-Flight camera based 3D point cloud reconstruction of a car[J]. *Computers in Industry*, 2013, 64(9): 1099-1114.
- [41] 李诗锐, 李琪, 李海洋, 等. 基于 Kinect v2 的实时精确三维重建系统[J]. *软件学报*, 2016, 27(10): 2519-2529.
- Li S R, Li Q, Li H Y, et al. Real-time accurate 3D reconstruction based on Kinect v2[J]. *Journal of Software*, 2016, 27(10): 2519-2529.
- [42] 罗先波, 钟约先, 李仁举. 三维扫描系统中的数据配准技术[J]. *清华大学学报(自然科学版)*, 2004, 44(8): 1104-1106.
- Luo X B, Zhong Y X, Li R J. Data registration in 3-D scanning systems[J]. *Journal of Tsinghua University (Science and Technology)*, 2004, 44(8): 1104-1106.
- [43] Montusiewicz J, Miłosz M, Kęsik J, et al. Structured-light 3D scanning of exhibited historical clothing: a first-ever methodical trial and its results[J]. *Heritage Science*, 2021, 9(1): 74.
- [44] 潘其锋, 任悦萍. 基于多视图几何的三维测量技术研究[J]. *建材与装饰*, 2015(47): 192-193.
- Pan Q F, Ren Y P. Research on 3D measurement technology based on multi-view geometry[J]. *Construction Materials & Decoration*, 2015(47): 192-193.
- [45] 汪秉论. 基于单目视觉多视图几何的三维点云重建算法研究[D]. 武汉: 华中科技大学, 2019.
- Wang B Q. Research on 3D point cloud reconstruction algorithm based on monocular vision and multi-view geometry[D]. Wuhan: Huazhong University of Science and Technology, 2019.
- [46] 李在林. 基于多视图几何的三维重建技术研究[D]. 深圳: 中国科学院大学(中国科学院深圳先进技术研究院), 2020.
- Li Z L. Research on 3D reconstruction technology based on multi-view geometry[D]. Shenzhen: Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 2020.
- [47] Morse B, Howard J, Cohen S, et al. PatchMatch-based content completion of stereo image pairs[C]//2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, October 13-15, 2012, Zurich, Switzerland. New York: IEEE Press, 2012: 555-562.
- [48] Barnes C, Shechtman E, Finkelstein A, et al. PatchMatch: a randomized correspondence algorithm for structural image editing[J]. *ACM Transactions on Graphics*, 2009, 28(3): 24.

- [49] Bleyer M, Rhemann C, Rother C. PatchMatch stereo-stereo matching with slanted support windows[C]// Proceedings of the British Machine Vision Conference 2011, August 29-September 2, 2011, Dundee. London: British Machine Vision Association, 2011.
- [50] Hirschmuller H. Accurate and efficient stereo processing by semi-global matching and mutual information[C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June 20-25, 2005, San Diego, CA, USA. New York: IEEE Press, 2005: 807-814.
- [51] 高莉. 改进的 Delaunay 三角剖分算法研究[D]. 兰州: 兰州交通大学, 2015.
Gao L. Research on improved Delaunay triangulation algorithm[D]. Lanzhou: Lanzhou Jiatong University, 2015.
- [52] Green P J, Sibson R. Computing dirichlet tessellations in the plane[J]. The Computer Journal, 1978, 21(2): 168-173.
- [53] Bowyer A. Computing dirichlet tessellations[J]. The Computer Journal, 1981, 24(2): 162-166.
- [54] Watson D F. Computing the n-dimensional delaunay tessellation with application to voronoi polytopes[J]. The Computer Journal, 1981, 24(2): 167-172.
- [55] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction[C]//Eurographics Symposium on Geometry Processing 2016, June 20-24, 2016, Berlin, Germany. Utrecht: Utrecht University, 2016.
- [56] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4): 193-202.
- [57] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [58] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.
- [59] Huang G B, Lee H, Learned-Miller E. Learning hierarchical representations for face verification with convolutional deep belief networks[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 2518-2525.
- [60] Ranzato M, Hinton G E. Modeling pixel means and covariances using factorized third-order boltzmann machines[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE Press, 2010: 2551-2558.
- [61] Ranzato M, Mnih V, Susskind J M, et al. Modeling natural images using gated MRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(9): 2206-2222.
- [62] Courville A, Bergstra J, Bengio Y. Unsupervised models of images by spike-and-slab RBMs[C]// Proceedings of the 28th International Conference on Machine Learning, June 28-July 2, 2011, Bellevue, Washington, DC, USA. Madison: Omnipress, 2011.
- [63] Yao Y, Luo Z X, Li S W, et al. MVSNet: depth inference for unstructured multi-view stereo[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer Vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11212: 785-801.
- [64] Yao Y, Luo Z X, Li S W, et al. Recurrent MVSNet for high-resolution multi-view stereo depth inference[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 5520-5529.
- [65] Yang J Y, Mao W, Alvarez J M, et al. Cost volume pyramid based depth inference for multi-view stereo[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 4876-4885.
- [66] Chen R, Han S F, Xu J, et al. Point-based multi-view stereo network[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 1538-1547.
- [67] Gu X D, Fan Z W, Zhu S Y, et al. Cascade cost volume for high-resolution multi-view stereo and stereo matching [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 2492-2501.
- [68] Cheng S, Xu Z X, Zhu S L, et al. Deep stereo using adaptive thin volume representation with uncertainty awareness[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 2521-2531.
- [69] Sormann C, Knöbelreiter P, Kuhn A, et al. BP-MVSNet: belief-propagation-layers for multi-view-stereo [C]//2020 International Conference on 3D Vision (3DV), November 25-28, 2020, Fukuoka, Japan. New York: IEEE Press, 2021: 394-403.
- [70] Yu Z H, Gao S H. Fast-MVSNet: sparse-to-dense multi-view stereo with learned propagation and gauss-newton refinement[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1946-1955.
- [71] Xu Q S, Tao W B. PVSNet: pixelwise visibility-aware multi-view stereo network[EB/OL]. (2020-07-15)[2023-03-05]. <https://arxiv.org/abs/2007.07714>.
- [72] Ma Z Y, Teed Z, Deng J. Multiview stereo with cascaded epipolar RAFT[M]//Avidan S, Brostow G, Cissé M, et al. Computer vision-ECCV 2022. Lecture notes in computer science. Cham: Springer, 2022, 13691: 734-750.
- [73] Weilharter R, Fraundorfer F. HighRes-MVSNet: a fast multi-view stereo network for dense 3D reconstruction

- from high-resolution images[J]. *IEEE Access*, 2021, 9: 11306-11315.
- [74] Ma X J, Gong Y, Wang Q R, et al. EPP-MVSNet: epipolar-assembling based depth prediction for multi-view stereo[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 5712-5720.
- [75] Yu A Z, Guo W Y, Liu B, et al. Attention aware cost volume pyramid based multi-view stereo network for 3D reconstruction[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 175: 448-460.
- [76] Wang F, Galliani S, Vogel C, et al. PatchmatchNet: learned multi-view patchmatch stereo[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 14189-14198.
- [77] Li Y, Li W Y, Zhao Z J, et al. DRI-MVSNet: a depth residual inference network for multi-view stereo images [J]. *PLoS One*, 2022, 17(3): e0264721.
- [78] Wang F, Galliani S, Vogel C, et al. IterMVS: iterative probability estimation for efficient multi-view stereo[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 8596-8605.
- [79] Saeed S, Lee S, Cho Y, et al. ASPPMVSNet: a high-receptive-field multiview stereo network for dense three-dimensional reconstruction[J]. *ETRI Journal*, 2022, 44 (6): 1034-1046.
- [80] Peng R, Wang R J, Wang Z Y, et al. Rethinking depth estimation for multi-view stereo: a unified representation [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 8635-8644.
- [81] Yang J Y, Alvarez J M, Liu M M. Non-parametric depth distribution modelling based depth inference for multi-view stereo[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 8616-8624.
- [82] Cheng W, Bai Z Y, Li J J, et al. ADIM-MVSNet: adaptive depth interval multi-view stereo network for 3D reconstruction[C]//Proceedings of the 2022 5th International Conference on Image and Graphics Processing, January 7-9, 2022, Beijing, China. New York: ACM, 2022: 281-287.
- [83] Cao C J, Ren X L, Fu Y W. MVSFormer: multi-view stereo with pre-trained vision transformers and temperature-based depth[EB/OL]. (2022-08-04) [2023-06-03]. <https://arxiv.org/abs/2208.02541>.
- [84] Xu H B, Zhou Z P, Qiao Y, et al. Self-supervised multi-view stereo via effective co-segmentation and data-augmentation[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(4): 3030-3038.
- [85] Chang D, Božič A, Zhang T, et al. RC-MVSNet: unsupervised multi-view stereo with neural rendering [M]//Avidan S, Brostow G, Cissé M, et al. *Computer vision-ECCV 2022. Lecture notes in computer science*. Cham: Springer, 2022, 13691: 665-680.
- [86] Ding Y K, Zhu Q T, Liu X Y, et al. KD-MVS: knowledge distillation based self-supervised learning for multi-view stereo[M]//Avidan S, Brostow G, Cissé M, et al. *Computer vision-ECCV 2022. Lecture notes in computer science*. Cham: Springer, 2022, 13691: 630-646.
- [87] Gao C, Liu J, Li X, et al. Accurate compressed look up table method for CGH in 3D holographic display[J]. *Optics Express*, 2015, 23(26): 33194-33204.
- [88] Pan Y J, Wang Y T, Liu J, et al. Improved full analytical polygon-based method using Fourier analysis of the three-dimensional affine transformation[J]. *Applied Optics*, 2014, 53(7): 1354-1362.
- [89] Wang F, Shimobaba T, Zhang Y P, et al. Acceleration of polygon-based computer-generated holograms using look-up tables and reduction of the table size via principal component analysis[J]. *Optics Express*, 2021, 29(22): 35442-35455.
- [90] Jia J, Si J, Chu D P. Fast two-step layer-based method for computer generated hologram using sub-sparse 2D fast Fourier transform[J]. *Optics Express*, 2018, 26(13): 17487-17497.
- [91] Horisaki R, Takagi R, Tanida J. Deep-learning-generated holography[J]. *Applied Optics*, 2018, 57(14): 3859-3863.
- [92] Lee J, Jeong J, Cho J, et al. Deep neural network for multi-depth hologram generation and its training strategy [J]. *Optics Express*, 2020, 28(18): 27137-27154.
- [93] Eybposh M H, Caira N W, Atisa M, et al. DeepCGH: 3D computer-generated holography using deep learning [J]. *Optics Express*, 2020, 28(18): 26636-26650.
- [94] Shi L, Li B C, Kim C, et al. Towards real-time photorealistic 3D holography with deep neural networks [J]. *Nature*, 2021, 591(7849): 234-239.
- [95] Wu J C, Liu K X, Sui X M, et al. High-speed computer-generated holography using an autoencoder-based deep neural network[J]. *Optics Letters*, 2021, 46(12): 2908-2911.
- [96] Chang C L, Wang D, Zhu D C, et al. Deep-learning-based computer-generated hologram from a stereo image pair[J]. *Optics Letters*, 2022, 47(6): 1482-1485.
- [97] Chang C L, Zhu D C, Li J M, et al. Three-dimensional computer holography enabled from a single 2D image[J]. *Optics Letters*, 2022, 47(9): 2202-2205.
- [98] Shui X H, Zheng H D, Xia X X, et al. Diffraction model-informed neural network for unsupervised layer-based computer-generated holography[J]. *Optics Express*, 2022, 30(25): 44814-44826.