

平行光场：基本框架与流程

王飞跃^{1,2,3}, 孟祥冰¹, 杜思聪^{1,2}, 耿征⁴

(1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190;

2. 青岛智能产业技术研究院平行光学技术创新中心, 山东 青岛 266000;

3. 中国科学院大学人工智能学院, 北京 100049; 4. Genex 科技公司, 美国 马里兰州 肯辛顿 20895)

摘要: 光场是环境中光线的集合, 对光场信息的采集、计算以及显示是极具挑战性的复杂问题。基于 ACP 的平行光场理论为解决此问题提供了一条新的有效途径。平行光场利用采集得到的实际物理世界的光场信息构建了增强的人工世界的光场, 并且以采集得到的实际物理世界的光场信息为指导, 在增强的人工世界的光场中进行光场实验, 得到最优的光场采集方案或者显示虚实方案, 最后通过平行执行的过程建立起物理世界的光场和人工世界的光场的联动优化, 使整个系统成为闭环, 实现智能的光场采集和显示, 建立光场信息处理和应用的虚实结合的全新理论体系。

关键词: 平行; ACP; 光场; 光场显示; 光场采集

中图分类号: TP37

文献标识码: A

doi: 10.11959/j.issn.2096-6652.202112

Parallel light field: the framework and processes

WANG Fei-Yue^{1,2,3}, MENG Xiangbing¹, DU Sicong^{1,2}, GENG Zheng⁴

1. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

2. Parallel Optics Technology Innovation Center, Qingdao Academy of Intelligent Industries, Qingdao 266000, China

3. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

4. Genex Technologies Inc., Kensington, MD 20895, the United States

Abstract: The light field is the collection of lights in the environment, and the acquisition, calculation and display of light field information are extremely challenging and complex issues. The ACP-based theory of parallel light field provides a new and effective way to solve this problem. It used the acquired light field information from the actual physical world to construct an enhanced light field of artificial world. And guided by the light field and other information from the actual physical world, light field experiments were conducted in the enhanced light field in all artificial world to obtain the optimal light field acquisition or display schemes. Finally, the parallel optimization of the physical and artificial light fields was established through the process of parallel execution, so that the entire system become a closed-loop system. In this way, intelligent light field acquisition and display were achieved, and a real-virtual theoretical framework for light field information processing and utilization was established.

Key words: parallel, ACP, light field, light field display, light field acquisition

1 引言

人的感官信息有 70%~80%来源于视觉, 另外有研究表明, 人脑约 50%的信息处理能力被用于处

理视觉信息^[1]。这里涉及的视觉信息来源可以用光场(light field)^[2-3]统一表示。这一理论被广泛应用于计算机视觉^[4]、真三维显示^[5-7]以及计算机图形学^[8]等。类比电磁场表示了三维物理空间的电磁线集

收稿日期: 2020-07-21; 修回日期: 2020-08-25

通信作者: 王飞跃, feiyue.wang@ia.ac.cn

基金项目: 国家重点研发计划基金资助项目(No.2018AAA0101502)

Foundation Item: The National Key Research and Development Program of China(No.2018AAA0101502)

合，光场是三维物理空间中的光线集合的完备表示。三维空间中光场光线全光函数示意如图1所示。

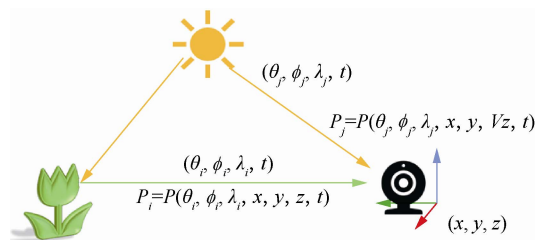


图1 三维空间中光场光线全光函数示意

如图1所示，从光源发出的光线，有些遇到环境中的物体产生了反射、折射等，有些则直接投向采集设备。进入采集设备的光线可以用全光函数（plenoptic function）^[9]表示：

$$P = P(\theta, \phi, \lambda, x, y, z, t) \quad (1)$$

其中，采集设备的坐标为 (x, y, z) ，光线进入采集设备的水平和垂直夹角为 (θ, ϕ) ，光线波长为 λ ，采集时间为 t 。

若要计算某个场景的全光场，就需要获取来自这一场景的所有7维光线的信息（如式（1）所示）：从光场采集角度来说，需要对环境中的所有位置、所有角度进行采集；从光场计算角度来说，计算量取决于环境中光线的数量^[10]。光场显示融合了光场采集和光场计算，因此其计算复杂度与环境的光场中光线的数量呈正相关。

空间中的光场信息计算的复杂度取决于场景中物体以及光源的复杂度。从光场采集的角度来看，如果环境是单色光源、平面（例如白墙）的，那么上述计算便会非常简单；相对地，动态、多光源、混合材质、复杂形状场景下的光场是非常复杂的。而人们生活中的多数场景是复杂的，如果待研究的环境是雨天、雾天等极端天气，并且有复杂的动态光源和物体，那么这种场景下的光场信息将非常庞大。目前很难通过单一的光场采集技术获取上述所有情况下的完备光场信息。从光场显示的角度来看，光场显示的目的是让观众身临其境地感受显示的内容，因此，显示的内容需要根据物理世界的光场进行相应的调整。同样，目前很难有一种方案能做到环境感知下动态全光场显示。综上，物理世界中的光场信息处理是一个非常复杂的任务。

本文提出了平行光场理论，以期解决复杂场景中的光场信息处理任务。平行光场是ACP（即人工社会（artificial societies）、计算实验（computational

experiments）、平行执行（parallel execution）^[11-13]理论在光场信息处理中的应用和拓展。其核心是搭建与实际物理世界的光场平行的虚拟人工世界的光场（A步），通过计算实验^[14]（C步）计算推测适用于实际物理世界光场处理的方法和参数，最终通过平行执行（P步）的方式建立物理世界的光场和人工世界的光场间的闭环反馈，并进行协同更新，实现在线化、长期化的优化处理。

平行光场提出了与物理世界的光场平行的人工世界的光场，通过对物理世界的光场进行模拟和学习，人工世界的光场中不仅可以涵盖常见的物理世界的光场信息，而且还可以演化生成罕见的光场信息，从而进行光场解决方案的测试或实验。除了光场信息外，人工世界的光场中还包含不同约束条件下光场信息处理的解决方案。进而，针对实际物理场景中的复杂问题，可以利用人工世界的光场进行计算实验，从而模拟推演出最优的解决方案。最后在平行执行的过程中，将这一解决方案部署到实际物理系统中，根据实际物理系统的反馈来优化人工世界的光场，并调整计算实验方案，继续推演改进方案。因此，平行光场是人工世界的光场和物理世界的光场的闭环优化系统，是实现复杂场景下光场信息处理的有效途径。

2 相关工作综述

光场信息计算整体流程如图2所示，传统的光场信息处理研究可以分为3个方面：光场采集^[15]、光场计算^[16]以及光场显示^[6]。光场采集是指利用相机阵列或者移动相机对实际物理空间中自然的光场进行采集；光场计算是指利用采集得到的光场，通过立体几何以及计算机视觉等原理恢复出场景的全光场，并逆推出场景中的物体的三维信息，然后根据需求对恢复出的光场信息进行删减、插入等增强工作，最终得到虚拟空间中人工的光场；光场显示是指将增强后的人工的光场通过真三维显示、增强现实（augmented reality, AR）等技术呈现在实际的物理世界中，与物理世界的光场形成混合的光场，并呈现给观众。

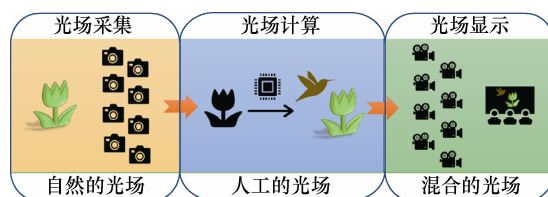


图2 光场信息计算整体流程

2.1 光场采集的现状

光场采集的目的是获取实际物理世界场景中的光场信息。分析式(1)可以得出,与光场采集相关的变量有7个:相机的位置 (x, y, z) 、光线的入射角度 (θ, ϕ) 、光线波长 λ 以及采集时间 t 。因此,单目相机获取的二维图像是一种简单的降维采样后的光场,其仅包含场景中固定位置、单一角度、固定时间的光场。为了尽可能从更多位置、更多角度、更长时间内获取光场信息,目前广泛研究的光场采集方案有3种。图3(a)为美国 Adobe 系统公司研发的基于微透镜阵列的光场采集^[17-19],其在成像传感器与主镜头之间加入由多透镜组成的微透镜阵列,使得来自物体一点上的光线经过主镜头汇聚后,在微透镜阵列处散开,然后被相机记录,这种方案的优点是体积小,但是分辨率较低,并且采集的范围有限。图3(b)为 Bennett W 等人^[20]在斯坦福大学研发的基于相机阵列的光场采集系统,该系统将多个标定后的相机分布在不同位置、不同角度,同时对场景的光场进行采集。该方案的采集分辨率高,但是所需相机数量多,硬件成本非常高,因此曹煊等人^[21]在此基础上提出利用压缩感知和稀疏编码的方案对采集信息进行处理,可以在一定程度上减少相机数量。图3(c)为 Kshitij M 等人^[22]在麻省理工学院设计的基于编码掩模的光场采集相机,其在镜头和成像传感器间加入一片半透明的编码掩模,通过掩模的调制获取环境的信息,进而根据与掩模相关的提前训练好的光场字典恢复场景的光场信息,此采集方案的优点是体积小,缺点是采用半透明掩模会使采集到的光场信噪比较低。

以上技术可以被总结为静态光场采集技术,即通过硬件上改造相机来获取多位置、多角度的光场,除此之外,还可以通过相机的移动来获取多角度、多位置的光场信息,然后再利用相机间的相对位姿信息,将光场信息融合起来,获取环境的全光场信息,因此,此方案的关键在于获取相机移动过程中的位姿信息以及相机合理的移动轨迹。从此角度考虑,同时定位和重建(simultaneous localization and mapping, SLAM)技术^[23-24]可以被看作一种动态光场采集技术,其在相机移动的同时实时计算相机的位姿,如果不考虑场景中光线随时间的变化(局部时间内不变),SLAM 采集方案可以被等效为相机阵列光场采集方案。视觉 SLAM 光场采集技术示意图如图4所示。

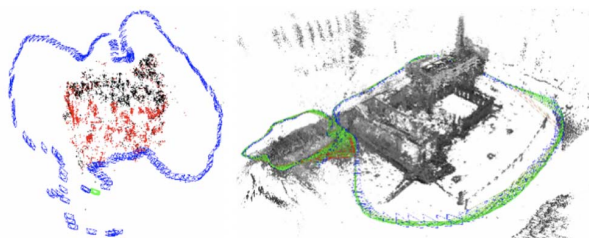


图4 视觉 SLAM 光场采集技术示意图

根据使用传感器的不同,视觉 SLAM 可以被划分为单目视觉 SLAM、双目视觉 SLAM 以及 RGB-D SLAM^[24]。传统的 SLAM 系统由前端和后端两部分组成。前端又称视觉里程计,它根据帧间的匹配关系增量式地恢复相机位姿。根据优化目标的不同,前端使用的方法可被分为直接法和间接法两类^[25]。

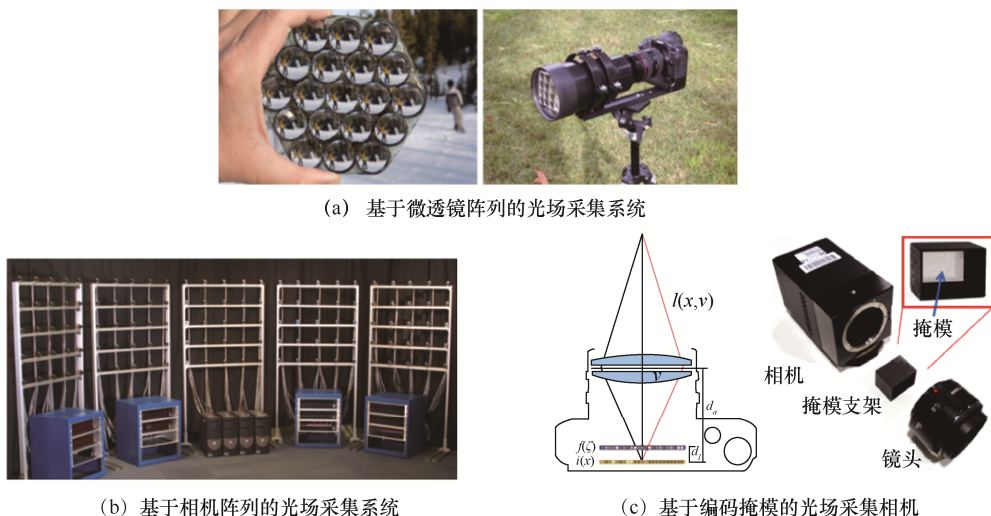


图3 光场采集技术分类^[23]

其中直接法以光度误差为优化目标,对相机位姿和路标点进行优化,而间接法的优化目标为重投影误差。因为系统长时间运行会累积噪声,所以在SLAM后端引入了减小误差的策略,具体来说有两种实现方案:滤波和非线性优化。其中非线性优化按照贝叶斯法则求解相机在每个位置的最大似然估计,准确率较高,现已在SLAM系统中得到了广泛应用。除此之外,为了使结果具有全局一致性,SLAM系统还具有基于外观方法的闭环检测功能。在没有先验信息的场景中,SLAM技术将相机位姿信息和地图数据作为反馈形成闭环,同步迭代,保证了系统的高效稳定运行。此外,随着深度学习技术在计算机视觉领域的发展,业界已逐步开始将深度学习的方法融入视觉SLAM系统,从而为SLAM系统提供更丰富的环境语义信息^[26]。传统的SLAM系统受环境影响较大,在此基础上,孟祥冰等人^[27]提出了平行感知理论框架,为解决视觉SLAM系统在复杂场景下的精度低、泛化能力差等问题指明了方向。

2.2 光场计算的现状

光场计算的目的是对采集获得的自然的光场进行再编辑,并从中恢复出对应的场景信息。光场可以被分为入射光场和反射光场:其中相机或人眼获得的光场是入射光场;光线触及物体表面发生反射形成的光线集合是反射光场,反射光场与物体表面的形状、物体材质等有关。如果将光源看作一种特殊的反射光场,那么入射光场来源于反射光场。因此可以通过采集得到的入射光场反推出反射光场,从而恢复出物体的三维表面结构。获得物体表面的反射光场后,便可以在虚拟世界中模拟实际物理世界的反射光场信息,并对其进行增强等操作。因此,光场计算被广泛应用于增强现实^[28]、动画特效制作^[29]等。美国南加利福尼亚大学ICT Graphic Lab的Paul Debevec领导的研究团队开发的Light Stage^[30]三维采集重建平台是目前光场计算在三维重建中非常前沿的研究之一。

图5(a)为其中一种Light Stage采集系统样机Light Stage 5,通过对设备上光源的调制,相机可以采集到不同光照下的光场信息,进而利用多视几何以及光影模型的原理恢复出物体的三维表面模型^[31-32]。

图5(b)为其重建结果,其中左图为人脸表面法向量图,中间的图为恢复的人脸三维模型,右图为加入皮肤细节的三维模型。可以看出,此种重建方案的效果非常好,但是其对采集设备以及采集时

人静止状态的要求非常高。Light Stage 5已经在《本杰明·巴顿》《蜘蛛侠》等电影特效中得到应用。除此之外,Light Stage还有其他的计算方案,如Light Stage 6、Light Stage X等。

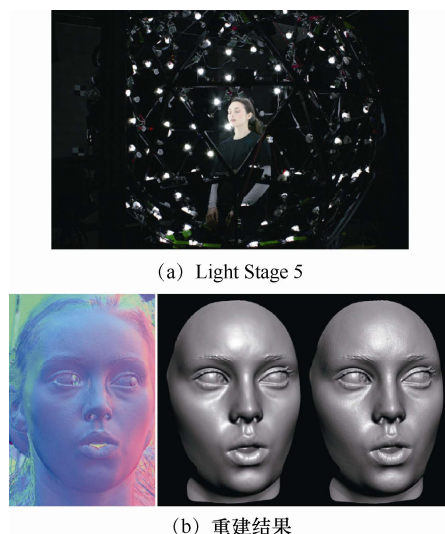


图5 Light Stage 5 及其重建结果

2.3 光场显示的现状

光场显示将增强后的光场通过显示设备呈现给观众,如图6所示,根据观众获取的信息的来源不同,光场显示可以被分为虚拟现实^[33]、混合现实^[34]以及增强现实^[35]。其中虚拟现实的内容由人工世界生成的光场构成,混合现实的内容由人工世界生成的光场和从实际物理世界采集到的光场融合而成。这两种方案在显示过程中均无须感知实际的物理场景。而增强现实的内容则由人工世界生成的光场投影到实际物理世界而成,其在显示过程中需要获取周围实际的物理场景,以决定人工世界的光场在实际物理世界的光场中的投射位姿。



图6 光场显示分类

以上3种显示方案均需要显示器设备将内容显示出来。普通的二维显示器只能提供仿射、遮挡、光照阴影、纹理、先验知识5个方面的心理上的三维感知^[6]。此外,真三维光场显示^[5-6]还能提供双目视差、移动视差、聚焦模糊3个方面的生理上的三维感知,可以从本质上提升观众体验的真实感,光场显示示意图如图7所示,其包含体三维显示、多视投影阵列、集成成像、数字全息、多层液晶张量显示5种显示方式。

如图8所示,我国在真三维显示领域成果显著,中国科学院自动化研究所在真三维显示领域先后开发了体三维显示器、基于柱面光栅的集成成像显示、多视投影显示阵列^[36]、多层液晶张量显示^[37]等多种显示方式,并在基于柱面光栅的集成成像显示中提出了多种优化方案^[38-39],提升了

视觉效果。该团队还探索了光场显示在外科手术中的应用,通过增强现实显示的方案来提升医生在手术过程中的视觉精准度,并为医生提供实时的信息指引^[40-41]。

除此之外,还有头戴式显示方案,如图9所示,HTC VIVE头戴式显示设备针对的是虚拟现实以及混合现实的显示,而HoloLens和Magic Leap头戴式显示设备针对的是增强现实的显示。

3 平行光场的基本框架

平行光场理论建立在ACP理论架构模型上,处理复杂环境中光场的采集、计算以及显示任务^[42],基于ACP理论的平行光场理论框架如图10所示。ACP理论是王飞跃^[11]于2004年提出的,该理论的核心是解决复杂系统的建模和调控问题。

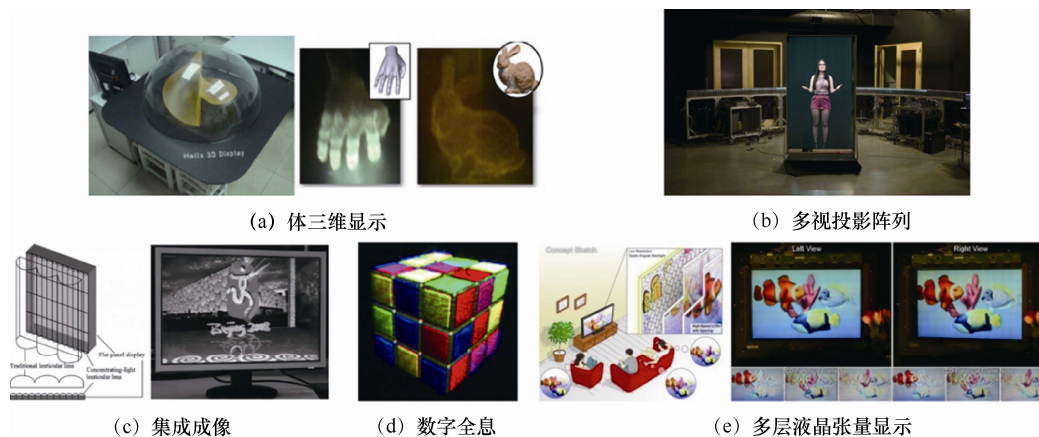


图7 光场显示示意

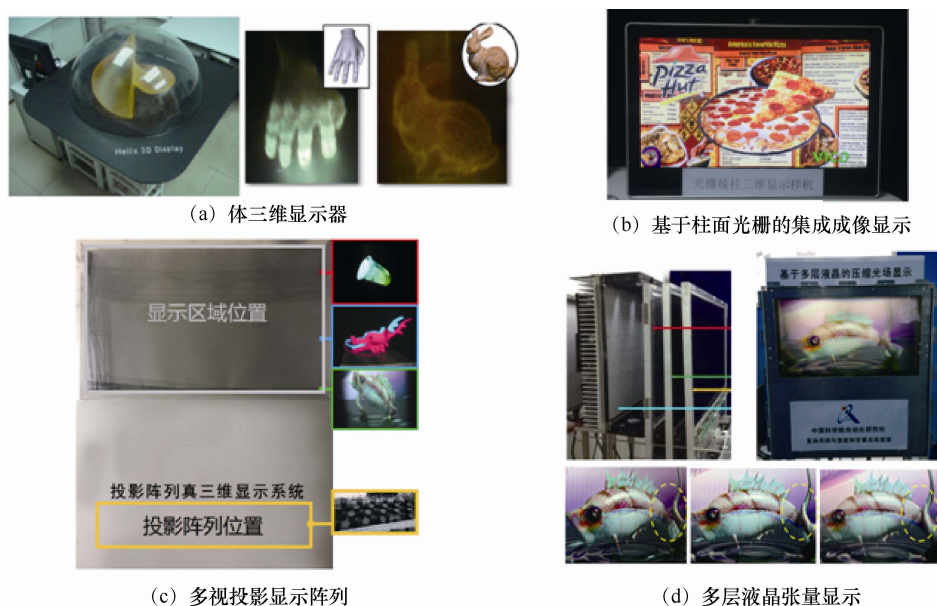


图8 国内光场显示相关成果

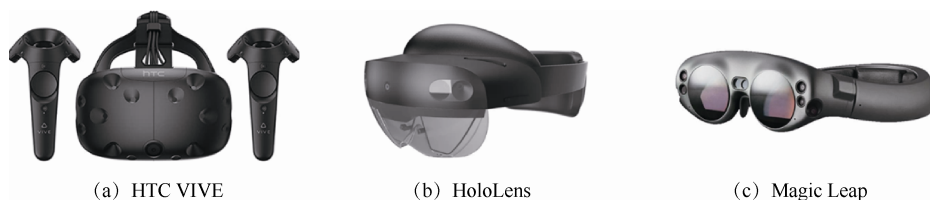


图 9 头戴式显示方案

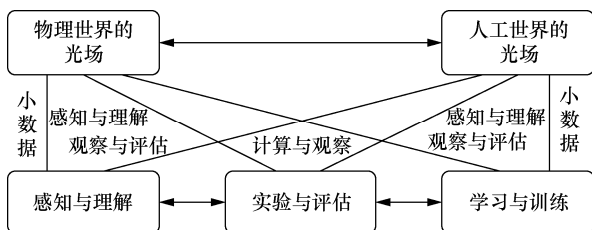


图 10 基于 ACP 理论的平行光场理论框架

ACP 理论通过建立与实际场景平行的虚拟人工场景，在人工场景中通过计算实验求解出适合实际物理场景的最优解决方案。最终通过平行执行过程，利用人工场景与实际场景间的虚实互动对整个系统进行闭环式的管理与控制。此理论方案已经被应用于农业^[43]、交通^[44]、医疗^[45]等领域，有效地提升了原始方案的效率。本文提出的平行光场理论将 ACP 理论推广到光场信息处理过程中，其目的是提升光场信息处理方案对实际物理环境的适应能力，通过计算实验获取个性化的解决方案，并利用平行执行的过程，建立光场的信息处理（光场的采集、计算以及显示）系统中人工世界的光场环境和实际物理环境间的双向沟通，形成闭环反馈。

平行光场的系统框架如图 11 所示，不同于传统的光场信息的开环式处理方式，该系统框架以人工世界的光场为中心，在人工世界的光场中进行光场实验，将光场生成和光场融合设计为虚实互动的闭环系统，从而提升系统执行的效率。

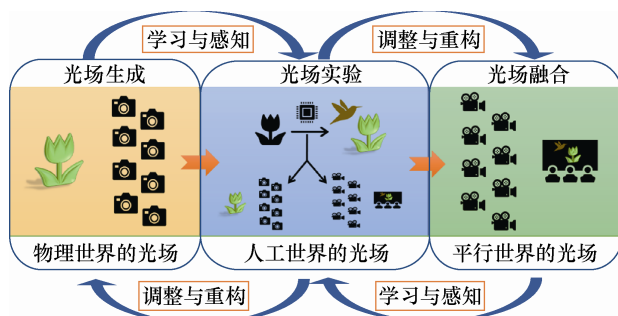


图 11 平行光场的系统框架

在平行光场的系统框架中，光场可以被分为：物理世界的光场，即物理世界中真实存在的光场；人工世界的光场，即在虚拟的人工世界中设计的光场，其一部分来源于对物理世界的逼真模拟，另一部分来源于对物理世界光场的增强和扩展；平行世界的光场，即将物理世界的光场和人工世界的光场融合在一起，并在真实的物理世界中被展示出来的光场。

不同于传统的光场信息处理流程，平行光场理论包含 3 个环节：光场生成、光场实验以及光场融合。光场生成是物理世界的光场和人工世界的光场互相沟通的桥梁。传统的光场采集仅仅是物理世界的光场到人工世界的光场的单向流动，而在平行光场理论中，光场生成不仅将物理世界的光场采集到人工世界，其采集过程还受到人工世界的光场的指引，光场采集过程更高效。同时，光场生成还会根据任务需要主动向环境中投射信息，以提升采集效率。在此过程中，人工世界的光场对来自物理世界的光场进行学习和感知，利用计算实验过程获取针对此场景的最优采集策略，然后利用此策略指引实际采集生成过程。光场融合则将人工世界的光场和物理世界的光场融合起来。传统的光场显示过程仅仅将人工世界的光场单向显示在物理世界中，很少有将真实物理世界的光场作为约束条件纳入人工世界的光场显示中的。平行光场系统中的光场融合则将物理世界的光场信息作为引导，调整人工世界中的待融合的光场，并指引人工世界的光场的显示过程，使人工世界的光场与所在物理世界的光场以及观众充分融合，提升观众的观感体验。

综上所述，平行光场包含 3 个重要组成部分：人工世界的光场、光场实验，以及将物理世界的光场、人工世界的光场和平行世界的光场串联起来的光场生成和光场融合过程。对比 ACP 理论，笔者将光场生成和光场融合统称为平行执行。

3.1 平行光场中的 A：人工世界的光场

人工世界的光场对应于 ACP 理论中的 A，其目的是高度逼真地模拟实际物理世界的光场，并在此

基础上对其进行增强和扩展。

在复杂的物理世界中研究光场采集和显示,往往受限于现实的物理硬件资源,而得不到最优的结果。而在人工世界中,因其本身的灵活性,人们可以进行近乎无约束的实验探索,这也是平行光场中计算实验的基本依赖条件。因此,人工世界的光场可以被用于引导实际物理世界中的光场采集和显示过程。如图 11 所示,人工世界的光场对物理世界的光场的感知和学习是实现上述过程的关键因素,也是人工光场建立的基本方法。人工世界的光场通过感知和学习物理世界的光场,建立起逼真的人工光场数据库,如图 12 所示,其中包含光场数据以及各种光场条件下的光场生成方案和光场融合方案。

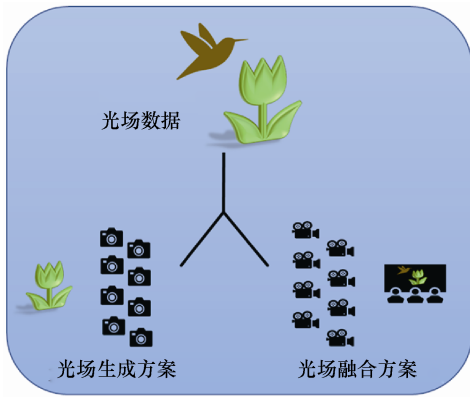


图 12 人工光场数据库

综上所述,人工世界的光场来源于实际物理世界的光场,并通过光场实验指引光场生成和光场融合。

3.2 平行光场中的 C: 光场实验

光场实验对应于 ACP 理论中的 C,其包含更新人工世界的光场以及更新平行执行中的策略两个主要过程。

如图 2 所示,传统的光场计算处理采集到的自

然光场信息是光场实验的一部分,被用来构建或更新人工世界的光场。如式(2)所示,光场计算过程获取处理后的自然光场的信息 p_p 和当前人工光场的信息 p_{a0} ,并以此为条件信息对人工世界的光场进行更新。

$$\hat{p}_a = \arg \max_{p_a} P(p_a | p_{a0}, p_p) \quad (2)$$

另外,如式(3)所示,光场实验借助已经构建起的人工世界的光场 p_a ,根据实际物理世界的光场信息 p_p ,通过大量的模拟计算获取最优的光场信息处理方案 λ_e ,从而引导实际物理世界的光场信息处理。因此,一个完善的人工世界的光场是进行光场实验的基础条件。

$$\hat{\lambda}_e = \arg \max_{\lambda_e} P(\lambda_e | p_a, p_p) \quad (3)$$

光场实验的系统流程如图 13 所示。其中,光场生成过程中的计算实验通过初始采集得到的物理世界的光场信息,以人工世界的光场数据库为参考信息,感知得到该场景对应的场景类型,进而根据该场景类型结合人工设定的采集目标得到目标优化任务。求解该任务时,以人工世界的光场中已有的最近距离的采集方案为初始值,求解得出针对该场景的最优光场生成方案,其包含相机的部署以及移动轨迹等采集过程最核心的架构和参数。

光场融合过程中的计算实验的目的是得到适合物理环境的最优显示方案,因此,该过程首先需要对当前的物理环境进行感知来获取当前环境的场景类型信息,然后以人工世界的光场中已有的最近距离的显示方案为初始值,结合采集得到的物理世界的光场对待融合的人工光场内容进行增强优化,得到最优光场融合内容,同时求解最优光场融

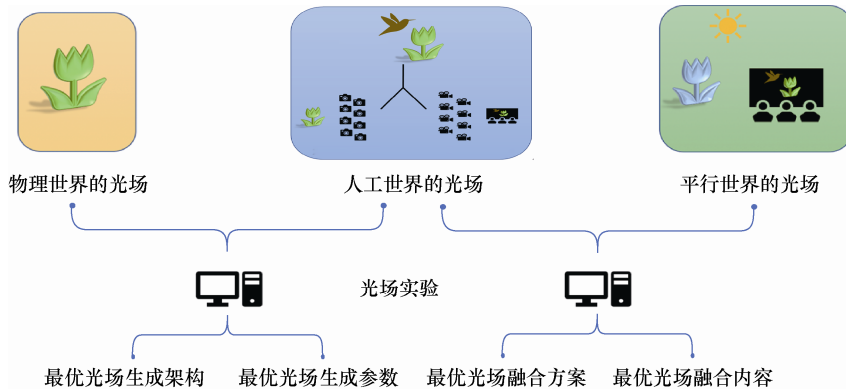


图 13 光场实验的系统流程

合方案（包含光场融合的系统方案及对应参数）。

3.3 平行光场中的 P：平行执行

平行执行对应 ACP 理论的 P，其目的是将光场实验中求解得到的最优方案应用于实际物理世界的光场信息处理中，同时将实际物理世界的光场处理信息反馈给光场实验来更新人工世界的光场，并调整光场实验方案，因此，平行执行是平行光场的最终执行环节，其建立起人工世界的光场和实际物理世界的光场间的虚实互动，最终形成持久优化的光场信息处理过程。需要特别指出的是，在策略执行过程中，需要光场信息处理的设备具有可调节、可重构的能力，以响应策略执行命令：

$$p_i = f(\hat{\lambda}_e) \quad (4)$$

其中， $f(\cdot)$ 为策略执行函数， p_i 为执行后获取的信息，其中光场生成得到物理世界的光场，光场融合获得融合后的平行世界的光场信息。

光场生成的平行执行过程如图 14 所示，其包含策略执行以及数据生成感知两个主要过程。在实际物理世界的光场生成过程中会产生来自物理环境以及生成方案本身的信息数据，光场实验便会利用这些信息数据在人工世界的光场中对光场生成方案进行实验优化，得到适合该环境的最优方案，经过参数调整或者系统重构后被应用于实际的物理世界的光场生成过程中。同时，人工世界的光场利用采集得到的物理世界的有效信息，通过学习和

感知过程进行自我更新（包含光场数据以及生成和融合方案的更新），至此光场生成过程便成为持续优化更新的闭环过程。

例如，基于视觉 SLAM 的动态光场采集通过视觉 SLAM 获取当前场景的相机的位姿以及采集到的光场，并将其反馈给人工世界，光场实验根据此光场信息在人工光场数据库中寻找与之最匹配的信息，以此为优化初值，然后结合物理世界和人工世界的光场搜索得到最优的采集路径以及最优的视觉 SLAM 方案，通过物理世界中的执行端进行采集，进而形成由人工世界引导的动态采集过程。此方案充分考虑了环境对采集算法的影响，并且充分利用了人工世界的先验知识，既保证了采集的稳定性，又提升了采集精度。

光场融合平行执行过程如图 15 所示，其包含策略执行和数据采集感知两个主要过程。与光场生成方案不同的是，光场融合平行执行过程的策略执行过程是调整或者重构显示设备，并且依赖光场生成的过程获取平行世界的光场信息。通过先后获取未融合人工世界的光场信息的物理世界的光场信息和融合了人工世界的光场信息后的平行世界的光场信息，可以对平行世界的光场信息进行分离，以获取人工世界的光场信息在物理世界中的分布。除此之外，采集的数据还需要包含观众的肢体、表情、注视方向以及位姿等信息。然后，光场实验便会根据这些信息对显示设备以及显示内容进行实时

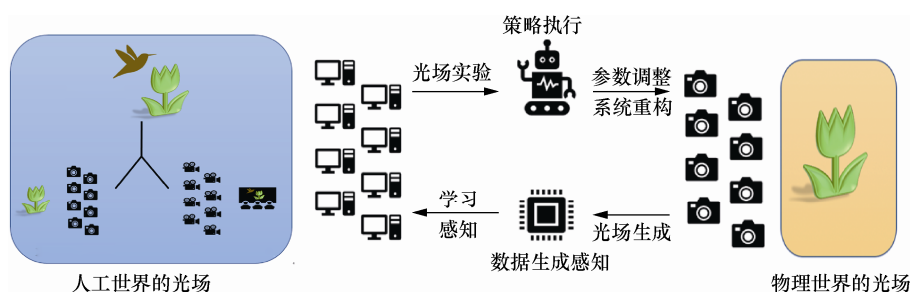


图 14 光场生成的平行执行过程

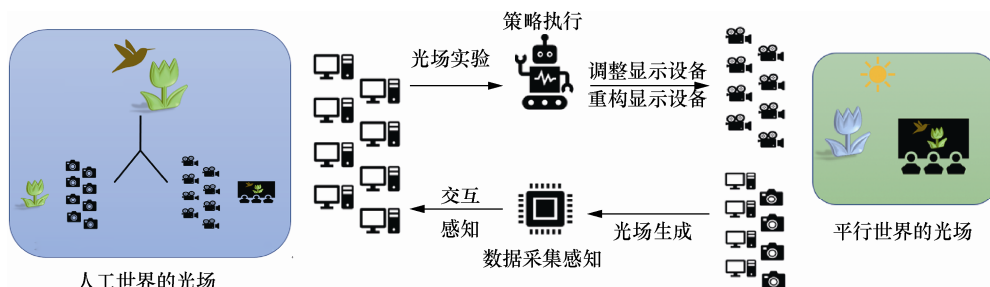


图 15 光场融合平行执行过程

的策略优化调整,以提升当前环境下观众的观感体验。

例如,在舞台灯光的透射过程中,首先需要采集当前舞台的光场信息,将其反馈给人工世界的光场,之后光场实验根据来自物理世界的光场信息,结合已有的人工世界的光场信息,同时以舞台上的表演内容信息为引导,计算适合当前场景的透射光场信息以及根据当前舞台设备求解最优的透射方案,然后通过透射设备将设计好的光场投射到实际物理舞台。最后光场实验会根据光场采集的平行世界的光场信息实时调整透射方案,使舞台上的光影效果达到最佳。

4 平行光场的关键技术

平行光场的关键技术如图 16 所示,共分为 3 个部分:描述智能,用于构建人工世界的光场;预测智能,用于进行光场实验;引导智能,用于平行执行中的环境感知以及执行优化。

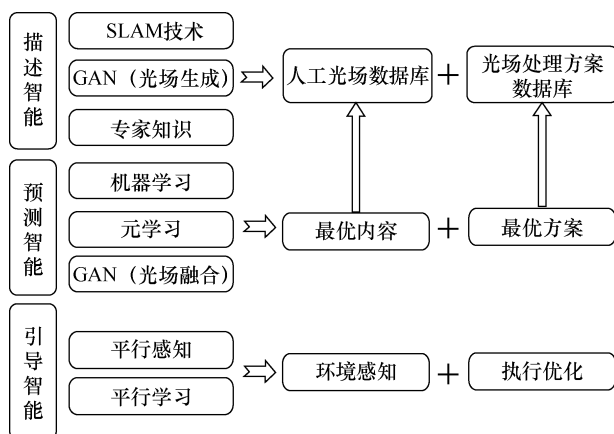


图 16 平行光场中的关键技术

4.1 平行光场的描述智能

平行光场的描述智能是人工世界的光场的关键技术,人工世界的光场的数据集的信息来源包含两个方面:通过光场生成过程获取的实际物理世界的光场信息以及通过学习实际物理世界的光场的构成规律而增强生成的光场信息。除了人工光场数据集之外,人工世界的光场中还包含不同场景下光场生成以及光场融合的策略方案。这些策略方案产生的来源既包括人工设计的方案,还包括光场实验中产生的最优策略方案历史。因此,构建完善的人工世界的光场需要人工光场生成、光场生成和光场融合等关键技术的支持。

如图 16 所示,描述智能中包含 SLAM 技术、

生成对抗网络 (generative adversarial network, GAN) 以及专家知识。

传统的光场采集方案包括微透镜阵列、阵列相机、掩模光场相机等。场景中的光场是稠密的,若要采集到场景中所有位置的光场,需要在场景中所有位置都布置相机,因此实际光场采集需要在采集的稠密性和相机的数量间进行权衡。而移动采集方案 (例如视觉 SLAM^[23]、运动恢复结构 (structure from motion, SFM)^[46] 等) 则通过相机的移动在一定程度上解决了静态场景下的光场采集的稠密性问题。SFM 和 SLAM 的区别在于 SLAM 需要实时在线地计算相机位姿和场景结构,而 SFM 则可以在前期完成数据的采集,在后期根据数据完成场景的重建和相机位姿的计算。因此,阵列相机方案结合移动采集方案是获取稠密光场的有效途径。在这个方案中,阵列相机布置方案以及相机的移动方案是稠密采集的关键,因此,需要人工光场的指引,从而构成光场生成的平行执行过程。

当前已有很多成熟的游戏引擎和仿真工具能够完成光场生成的任务,例如 Unity、Half-Life 2、Delta3D、OpenGL、Panda3D、Google 3D Warehouse、3DS MAX、OVVV、VDrift 等。然而,人工世界的光场中生成的光场需要满足实际物理光场的规律,这些工具均无法直接获取实际物理光场的规律。GAN^[47] 由 Goodfellow 提出,其通过生成器和判别器的对抗博弈来获得能够生成足够逼真的数据的生成器,并且已经在图像生成、风格迁移、图像转换、场景合成领域被广泛应用。在平行光场中,笔者采用 GAN,以实际的物理光场为风格 (真) 信息,以人工生成的光场为内容信息,利用不断获取的数据对网络进行优化更新,以获取符合实际物理光场规律的光场信息。

光场生成和光场融合方案策略均是根据实际物理世界中的光场分布规律进行设计的,因此,可以利用符合实际物理规律的专家知识生成光场信息以及光场生成和融合方案,并将其加入人工光场数据库中。除此之外,基于数据的强化学习方案也是生成策略的有效途径。专家知识除了可以被用于构建策略之外,还可以通过知识图谱将颜色光学^[48-51] 等已有知识加入人工光场数据库,使光场实验可以充分考虑环境因素以及人的心理因素。研究表明,色彩对人的情绪有较大的影响^[48-51]。色彩作用于人的感

官，能够刺激人的神经，进而对人类的情绪和心理产生影响。比如橙色使人感到温暖，蓝色使人冷静，红色使人容易激动，绿色能够帮助人们缓解紧张的情绪、提高工作效率等。因此，利用此知识便可以通过光场实验设计出针对人的当前状态以及当前环境的最优的光场融合方案。

4.2 平行光场的预测智能

平行光场的预测智能是进行光场实验的关键技术，光场实验的核心思想是利用计算机进行各种实验，以获取最优的策略和最优参数。如图 13 所示，针对光场生成，光场实验包含最优采集架构以及最优参数；针对光场融合，光场实验包含最优的融合方案和最优的融合内容。

如图 16 所示，预测智能中包含机器学习、元学习等关键技术。

光场实验首先需要感知实际物理世界的光场，并根据获得的物理世界的光场搜索对应的人工世界的光场。因此，在得到实际物理世界的光场之后，光场实验便利用机器学习技术在已经构建好的人工世界的光场中进行搜索，度量物理世界的光场和人工世界的光场间的距离，从而获取距离最近的人工世界的光场场景。进而，通过得到的人工世界的光场场景，获取其对应的人工世界的光场中的策略方案，以此为初值，以当前的物理场景的光场信息为约束目标，对此策略方案进行优化更新，得到适合当前场景的最优策略。元学习^[52]便可以利用现有的光场信息、已有的人工世界的光场中的经验优化学习得到适合当前场景的最优策略和方案。

为了获取与当前环境相融合的显示内容，还可以采用风格迁移的方案，以环境光场为风格内容，将要显示的内容迁移到当前的环境风格下。其中的关键技术包含 GAN^[47]、深度卷积生成对抗网络（deep convolution generative adversarial network, DCGAN）^[53]、CycleGAN^[54]等。同时，根据感知到的环境中观众的位姿，依靠光场基础理论获取最优的显示器布置方案。最终的目的是提升观众的观感体验，因此需要获取观众在观看过程中的信息，其关键技术包含人体姿态识别^[55]、手势识别^[56]、微表情识别^[57]等与深度学习相关的技术。进而，根据观众的反应信息有针对性地调整光场显示，增强观众和人工世界的交互体验。

4.3 平行光场的引导智能

平行光场的引导智能是进行平行执行的关键

技术，平行执行是平行光场真正的执行阶段，其关键技术除了人工世界的光场构建以及光场实验所需的关键技术外，还包括实际物理世界的光场中执行方案的关键技术。

如图 16 所示，引导智能包含平行感知和平行学习，在平行执行过程中，需要获取并感知实际物理世界的光场信息，将其作为反馈信息反馈给人工世界的光场以及光场实验，并且需要在实际物理空间中执行在光场实验中得到的策略。

平行感知理论^[27]给出了在复杂场景中获取场景信息以及当前采集相机位姿的解决方案，得到当前的场景信息后，便可以利用 Image Caption^[58]以及 Video Summary^[59]等技术获取场景的类别和内容信息，并将其反馈给人工世界的光场。

若要将光场实验求解得出的策略应用于实际物理世界中，需要实际物理世界中的设备具有可调节、可重构的能力。传统静态的采集或显示系统显然无法满足这一要求，但是可以将采集或显示设备安装在可调节的机械臂末端，利用机械臂的可调节性来执行策略。另外，基于人工世界的光场获取的最优策略，光场实验若要在实际物理世界中执行该策略，需要充分考虑当前物理世界的光场信息，因此需要将此策略迁移到现实的物理世界中，而平行学习理论^[60]将物理世界中的学习和人工世界的学习融合起来，使策略在实际物理场景中可以很好地执行。

5 平行光场的应用

5.1 ParallelEye 数据集

图 17 所示是中国科学院自动化研究所王飞跃团队提出的 ParallelEye 数据集^[61]中的样本图像，该数据集的构建利用 OpenStreetMap、CityEngine、Unity3D 等计算机图形学工具，首先根据真实地图构建城市交通场景（包含车辆、行人、交通设施等），从而构建起人工场景模型；进而利用此人工场景模型进行计算实验，提升复杂场景下算法的鲁棒性。此人工场景模型是基于真实的物理光场信息构建起来的，并在此基础上建立了增强的车辆、行人等信息。然而，此数据集仅采集了现实场景中的部分光场信息（二维图像），没有将入射光场和反射光场（物体材质信息等）的全部信息考虑在内，数据集中的反射光场和物理世界的光场存在差异，因此，ParallelEye 数据集是不完备的平行光场中人工世界的光场。

图 17 ParallelEye 数据集中的样本图像^[61]

5.2 平行光场在增强现实中的应用

图 18 所示是由中国科学院自动化研究所王飞跃团队开发的增强现实系统^[27]在肾脏手术中的应用。该系统通过多目相机实时地采集腹腔中的光场信息,通过 SLAM 技术对该场景信息进行重建,进而获得实时的腹腔场景的三维结构信息。利用该三维结构信息和人工场景中的肾脏模型进行匹配定位,建立起人工场景和实际场景在空间上的关联,进而通过对人工场景进行增强(渲染手术所需的肿瘤等信息),获取融合后的光场信息,最终通过 AR 或真三维显示技术投射到真实的物理世界中,形成平行世界的光场,然后根据物理世界的光场信息的变化,实时调整光场的采集和增强策略。因此该应用几乎涵盖了平行光场从光场采集到光场显示的每个步骤,是平行光场在医疗手术中的典型应用。然而,该应用并没有建立一套完备的人工世界的光场,而是单纯地将采集后的局部光场信息融合到术前采集的信息中。因此,该应用仅可以进行简单的光场实验,是平行光场理论在医学领域的初步探索。

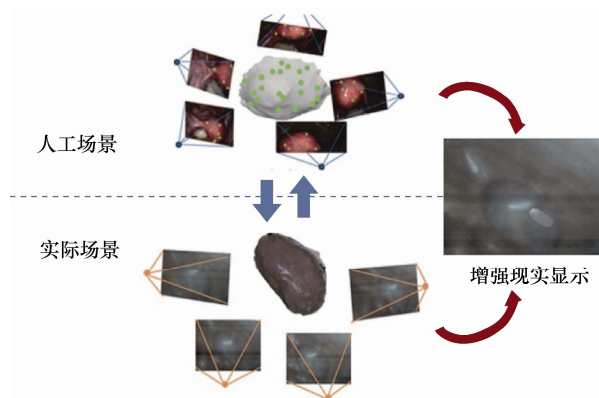


图 18 增强现实系统在肾脏手术中的应用

6 结束语

本文将 ACP 理论推广应用到光场信息处理中(包含采集、计算以及显示),提出了平行光场的基本理论框架,并指出了其中的关键技术。平行光场的基本框架是构建与物理世界的光场平行的人工世界的光场,通过光场计算实验求解得出最优的光场生成以及光场融合方案,最终通过平行执行过程建立人工世界的光场和物理世界的光场间的虚实互动,形成闭环的光场生成和光场融合系统。

将 ACP 理论与光场信息处理结合是一种全新的研究思路,其试图解决光场采集的稠密性和设备的简易性不能兼得的问题,以及提高在雾天、雨天等极端环境下的采集能力。在平行光场的理论中,光场融合将实际物理环境以及观众考虑在内,使人工世界的光场真正融入实际物理世界的光场,形成平行世界的光场,并建立与观众间的交流互动,大大提升观众的观感。笔者相信平行光场将成为光场信息处理领域一个重要的研究方向,平行光场理论和深度学习的结合将会推动越来越多的智能光场的应用。

致谢:

本文由腾讯优图实验室的曹焱博士订正,并参考中国科学院自动化研究所王蓉、裴仁静、张赵行、张梅等研究人员的研究工作,由多方协助共同完成,在此对其工作表示感谢。

参考文献:

- [1] ELAINE M, KATJA H, PEARSON E, et al. Human anatomy & physiology: pearson new international edition/interactive physiology 10-syste[J]. Enfermedades Infecciosas y Microbiología Clínica, 2013, 29(1): 72-73.
- [2] GERSHUN A. The light field[J]. Studies in Applied Mathematics,

- 1939, 18(1-4): 51-151.
- [3] MOON P, SPENCER D E. The photic field[M]. Cambridge: MIT Press, 1981.
 - [4] HONAUER K, JOHANNSEN O, KONDERMANN D, et al. A dataset and evaluation methodology for depth estimation on 4d light fields[C]//Proceedings of the Asian Conference on Computer Vision. Berlin: Springer, 2016.
 - [5] GENG J. Volumetric 3D display for radiation therapy planning[J]. Journal of Display Technology, 2008, 4(4): 437-450.
 - [6] GENG J. Three-dimensional display technologies[J]. Advances in Optics and Photonics, 2013, 5(4): 456-535.
 - [7] LEE B. Three-dimensional displays, past and present[J]. Physics Today, 2013, 66(4): 36-41.
 - [8] LEVOY M. Light field rendering[C]//Proceedings of the ACM SIGGRAPH Conference. New York: ACM Press, 1996.
 - [9] ADELSON E H. The plenoptic function and the elements of early vision[C]//Proceedings of the Computational Models of Visual Processing. [S.l.:s.n.], 1991: 3-20
 - [10] CHAI J X, TONG X, CHAN S C, et al. Plenoptic sampling[C]//Proceedings of the ACM SIGGRAPH Conference. New York: ACM Press, 2000: 307-318.
 - [11] 王飞跃. 平行系统方法与复杂系统的管理和控制[J]. 控制与决策, 2004, 19(5): 485-489, 514.
WANG F Y. Parallel system methods for management and control of complex systems[J]. Control and Decision, 2004, 19(5): 485-489, 514.
 - [12] WANG F Y. Parallel control and management for intelligent transportation systems: concepts, architectures, and applications[J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 11(3): 630-638.
 - [13] 王飞跃. 平行控制: 数据驱动的计算控制方法[J]. 自动化学报, 2013(4): 5-14.
WANG F Y. Parallel control: a method for data-driven and computational control[J]. Acta Automatica Sinica, 2013(4): 5-14.
 - [14] 王飞跃. 计算实验方法与复杂系统行为分析和决策评估[J]. 系统仿真学报, 2004, 16(5): 893-897.
WANG F Y. Computational experiments for behavior analysis and decision evaluation of complex systems[J]. Journal of System Simulation, 2004, 16(5): 893-897.
 - [15] LAM EDMUND Y. Computational photography with plenoptic camera and light field capture: tutorial[J]. Journal of the Optical Society of America a Optics Image Science & Vision, 2015, 32(11): 2021-2032.
 - [16] SHI L X, HASSANIEH H, DAVIS A, et al. Light field reconstruction using sparsity in the continuous fourier domain[J]. ACM Transactions on Graphics, 2014, 34(1): 1-13.
 - [17] GEORGEIV T, ZHENG K C, CURLESS B, et al. Spatio-angular resolution tradeoffs in integral photography[C]//Proceedings of the 17th Eurographics Symposium on Rendering Techniques, Nicosia, Cyprus. [S.l.]: The Eurographics Association, 2006.
 - [18] GEORGEIV T, INTWALA C. Light field camera design for integral view photography[R]. 2008.
 - [19] LUMSDAINE A, GEORGIEV T, SYSTEMS A. Full resolution lightfield rendering[R]. 2008.
 - [20] WILBURN B, JOSHI N, VAISH V, et al. High performance imaging using large camera arrays[J]. ACM Transactions on Graphics, 2005, 24(3): 765-776.
 - [21] CAO X, GENG Z, LI T T. Dictionary-based light field acquisition using sparse camera array[J]. Optics Express, 2014, 22(20): 24081-24095.
 - [22] MARWAH K, WETZSTEIN G, BANDO Y, et al. Compressive light field photography using over complete dictionaries and optimized projections[J]. ACM Transactions on Graphics, 2013, 32(4): 1-12.
 - [23] 刘浩敏, 章国锋, 鲍虎军. 基于单目视觉的同时定位与地图构建方法综述[J]. 计算机辅助设计与图形学学报. 2016, 28(6): 855-868.
LIU H M, ZHANG G F, BAO H J. A survey of monocular simultaneous localization and mapping[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(6): 855-868.
 - [24] CADENA C, CARLONE L, CARRILLO H, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age[J]. IEEE Transactions on Robotics, 2016, 32(6): 1309-1332.
 - [25] BRESSON G, ALSAYED Z, YU L, et al. Simultaneous localization and mapping: a survey of current trends in autonomous driving[J]. IEEE Transactions on Intelligent Vehicles, 2017, 2(3): 194-220.
 - [26] YOUNES G, ASMAR D, SHAMMAS E, et al. Keyframe-based monocular SLAM: design, survey, and future directions[J]. Robotics and Autonomous Systems, 2017: S0921889017300647
 - [27] 孟祥冰, 王蓉, 张梅, 等. 平行感知: ACP 理论在视觉 SLAM 技术中的应用[J]. 指挥与控制学报, 2017, 3(4): 350-358.
MENG X B, WANG R, ZHANG M, et al. Parallel perception: an ACP-based approach to visual SLAM[J]. Journal of Command and Control, 2017, 3(4): 350-358.
 - [28] LEE S, JANG C, MOON S, et al. Additive light field displays: realization of augmented reality with holographic optical elements[J]. ACM Transactions on Graphics, 2016, 35(4): 1-13.
 - [29] DEBEVEC P, WENGER A, TCHOU C, et al. A lighting reproduction approach to live-action compositing[C]//Proceedings of the 29th International Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2002: 547-556.
 - [30] TIM H, JONATHAN C, CHRIS T, et al. Light stage 2.0[C]//SIGGRAPH Technical Sketches. [S.l.: s.n.], 2001.
 - [31] ALEXANDER O, ROGERS M, LAMBETH W, et al. Creating a photoreal digital actor: the digital Emily project[C]//Proceedings of the 2010 Conference for Visual Media Production. Piscataway: IEEE Press, 2010: 176-187.
 - [32] ALEXANDER O, ROGERS M, LAMBETH W, et al. The digital Emily project: achieving a photorealistic digital actor[J]. IEEE Computer Graphics & Applications, 2010, 30(4): 20-31.
 - [33] STEUER J. Defining virtual reality: dimensions determining telepresence[J]. Journal of Communication, 1992, 42(4): 73-93.
 - [34] MILGRAM P, KISHINO F. A taxonomy of mixed reality visual displays[J]. IEICE Transactions on Information and Systems, 1994, 77(12): 1321-1329.
 - [35] AZUMA R, BAILLOT Y, BEHRINGER R, et al. Recent advances in augmented reality[J]. IEEE Computer Graphics and Applications, 2001, 21(6): 34-47.
 - [36] ZHANG A Z, GENG Z, LIA T, et al. Integration of real-time 3D image acquisition and multiview 3D display[J]. Proceedings of SPIE the International Society for Optical Engineering, 2014: 8979.
 - [37] CAO X, GENG Z, LI T T. Dictionary-based light field acquisition using sparse camera array[J]. Optics Express, 2014, 22(20): 24081-24095.
 - [38] PEI R J, GENG Z, ZHANG Z X. Subpixel multiplexing method for

- 3D lenticular display[J]. Journal of Display Technology, 2016, 12(10): 1.
- [39] PEI R J, GENG Z, MA K, et al. Three dimensional lenticular display synthetic image rendering based on light field acquisition[J]. Journal of the Society for Information Display, 2017, 25(1-3): 117-125.
- [40] WANG R, GENG Z, ZHANG Z X, et al. Visualization techniques for augmented reality in endoscopic surgery[C]//Proceedings of the International Conference on Medical Imaging and Augmented Reality. Berlin: Springer, 2016.
- [41] WANG R, ZHANG M, MENG X B, et al. 3D tracking for augmented reality using combined region and dense cues in endoscopic surgery[J]. IEEE Journal of Biomedical and Health Informatics, 2017: 1540-1551.
- [42] 王飞跃. 平行光场与平行光学, 从光学计算实验到光学引导智能[R]. 2018.
- WANG F Y. Parallel light field and parallel optics, from optical computing experiment to optical guided intelligence[R]. 2018.
- [43] 康孟珍, 王秀娟, 华净, 等. 平行农业: 迈向智慧农业的智能技术[J]. 智能科学与技术学报, 2019, 1(2): 107-117.
- KANG M Z, WANG X J, HUA J, et al. Parallel agriculture: intelligent technology toward smart agriculture[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(2): 107-117.
- [44] 吕宜生, 陈圆圆, 金峻臣, 等. 平行交通: 虚实互动的智能交通管理与控制[J]. 智能科学与技术学报, 2019, 1(1): 21-33.
- LYU Y S, CHEN Y Y, JIN J C, et al. Parallel transportation: virtual-real interaction for intelligent traffic management and control[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(1): 21-33.
- [45] 王飞跃, 张梅, 孟祥冰, 等. 平行手术: 基于 ACP 的智能手术计算方法[J]. 模式识别与人工智能, 2017, 30(11): 961-970.
- WANG F Y, ZHANG M, MENG X B, et al. Parallel surgery: an acp-based approach for intelligent operations, 2017, 30(11): 961-970.
- [46] AGARWAL S, SNAVELY N, SIMON I, et al. Building rome in a day[J]. Communications of the ACM, 2011, 54(10): 105-112.
- [47] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems, 2014, 3: 2672-2680.
- [48] 束越新. 颜色光学基础理论[M]. 济南: 山东科学技术出版社, 1981.
- SHU Y X. Basic theory of color optics[M]. Jinan: Shandong Science and Technology Press, 1981.
- [49] WYSZECKI G, STILES W S. Color science[M]. New York: Wiley, 1982.
- [50] FAIRCHILD M D. Color appearance models[M]. New York: Wiley, 2013.
- [51] BERNIS R S. Billmeyer and Saltzman's principles of color technology[M]. New York: Wiley, 2019.
- [52] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[J]. arXiv preprint, 2017, arXiv: 1703.03400.
- [53] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. Computer Science, 2015.
- [54] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2242-2251.
- [55] MITRA S, ACHARYA T. Gesture recognition: a survey[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C, 2007, 37(3): 311-324.
- [56] RAUTARAY S S, AGRAWAL A. Vision based hand gesture recognition for human computer interaction: a survey[J]. Artificial Intelligence Review, 2015, 43(1): 1-54.
- [57] LIU Y J, ZHANG J K, YAN W J, et al. A main directional mean optical flow feature for spontaneous micro-expression recognition[J]. IEEE Transactions on Affective Computing, 2015, 7(4): 299-310.
- [58] XU K, BA J, KIROUS R, et al. Show, attend and tell: neural image caption generation with visual attention[J]. Computer Science, 2015: 2048-2057.
- [59] POTAPOV D, DOUZE M, HARCHAOUI Z, et al. Category-specific video summarization[C]//Proceedings of the 13th European Conference on Computer Vision. Berlin: Springer, 2014: 540-555.
- [60] 李力, 林懿伦, 曹东璞, 等. 平行学习——机器学习的一个新型理论框架[J]. 自动化学报, 2017, 43(1): 1-8.
- LI L, LIN Y L, CAO D P, et al. Parallel learning: a new framework for machine learning[J]. Acta Automatica Sinica, 2017, 43(1): 1-8.
- [61] LI X, WANG K F, TIAN Y L, et al. The paralleleye dataset: a large collection of virtual images for traffic vision research[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20(6): 2072-2084.

[作者简介]



王飞跃 (1961—), 男, 博士, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室主任, 中国科学院大学中国经济与社会安全研究中心主任, 青岛智能产业技术研究院院长, 主要研究方向为平行系统的方法与应用、社会计算、平行智能以及知识自动化。



孟祥冰 (1990—), 男, 博士, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员, 主要研究方向为视觉 SLAM、真三维显示。



杜思聪 (1996—), 男, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室硕士生, 主要研究方向为视觉 SLAM。



耿征 (1957—), 男, 博士, Genex 科技公司创始人, 主要研究方向为真三维显示、三维成像、生物医学工程、智能视频监控系統、智能机器人。