

文章编号:1007-2780(2022)05-0654-06

基于 Leap Motion 手势识别的悬浮真 3D 显示 实时交互系统

林星雨, 邢 妍, 张汉乐, 王琼华*

(北京航空航天大学 仪器科学与光电工程学院, 北京 100191)

摘要:提出基于 Leap Motion 手势识别的悬浮真 3D 显示实时交互系统, 可实现对多个悬浮 3D 图像的独立交互。系统由显示和交互两个模块组成, 显示模块利用集成成像 3D 显示屏再现 3D 图像, 再通过二面角反射镜阵列实现悬浮真 3D 显示; 交互模块通过 Leap Motion 识别手部各关节空间位置, 提取手势信息并判断用户操作意图, 利用提出的通道机制, 独立改变 3D 场景内多个物体的空间坐标, 渲染对应的 3D 片源并刷新, 完成实时悬浮真 3D 显示交互, 构成人机交互闭环系统。实验结果表明, 该系统在 3840×2160 分辨率下, 可实现 30 帧/s 的交互帧率。所提系统通过预设手势可实现多物体的独立实时交互, 提高了交互自由度。

关键词: 悬浮 3D 显示; 集成成像; 人机交互

中图分类号: TN27 **文献标识码:** A **doi:** 10.37188/CJLCD.2022-0060

Real-time floating 3D display interaction system based on gesture recognition by leap motion

LIN Xing-yu, XING Yan, ZHANG Han-le, WANG Qiong-hua*

(School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China)

Abstract: A real-time floating 3D display interaction system is proposed, which can realize independent interaction of multiple floating 3D images. The proposed system consists of the display and interaction modules. In the display module, 3D images are reconstructed through integral imaging 3D display screen and then floats into the air through the dihedral corner reflector array. In the interaction module, Leap Motion captures positions of hand joints, extracts the gesture information and judges the motive of the users. After checking interaction channels, positions of multiple 3D objects are changed individually, and the 3D film source is rendered, which forms a human-computer interaction closed loop and realizes the real-time floating 3D display interaction. The experimental results demonstrate that the system can reach 30 frame/s under the resolution of 3840×2160 . The real-time individual interaction of multiple objects can be achieved in the proposed system, and the interaction freedom is increased.

Key words: floating 3D display; integral imaging; human-computer interaction

收稿日期:2022-02-21;修订日期:2022-03-03.

基金项目:国家自然科学基金(No. U21B2034, No. 62105014, No. 62105016)

Supported by National Natural Science Foundation of China(No. U21B2034, No. 62105014, No. 62105016)

*通信联系人, E-mail: qionghua@buaa.edu.cn

1 引言

人类生活在三维(3D)空间中,三维显示相较于目前普及的二维(2D)显示,能够更高效地呈现空间信息,具有更强的真实感和沉浸感。裸眼真3D显示无需佩戴眼镜等辅助设备,无立体观看视疲劳问题,能够提供舒适逼真的3D图像,应用前景十分广阔^[1]。悬浮真3D显示结合裸眼真3D显示与悬浮显示两种技术的优势,利用悬浮器件,将3D图像成像在空中,进一步增强了空间真实感。该技术与手势交互等非接触式交互技术结合^[2],既能进一步发挥3D显示自然直观的优势,又能使人和空间中的3D图像自由交互,可避免不必要的接触,特别有利于减少疾病传播风险,在2019年至今的新冠肺炎大背景下有多种应用场景。

现有的悬浮真3D显示手势交互主要利用全息、光场和集成成像等裸眼真3D显示技术实现3D图像的再现,配合体感检测器件或图像传感器完成手势交互功能。台湾交通大学提出了一种嵌入式集成成像3D显示手势交互系统^[3],利用嵌入在微透镜阵列下的分布式光传感器推算手指位置,实现了手指对3D场景的精确控制。日本东京工业大学实现了一种基于全息3D显示的手势交互系统,采用RGB相机检测手指散射光,对手指位置进行检测,实现了手指与复杂场景的精确配准^[4]。浙江大学提出了一种悬浮真3D显示系统力触觉交互技术,通过指套振动实现力触觉交互,提高了交互的真实感^[5]。尽管目前国内进行了一些悬浮真3D显示手势交互技术研究,但多数仅考虑单个3D物体的交互或3D场景的整体交互,未考虑复杂3D场景中多个3D物体的独立交互,不同3D物体独立交互存在的串扰问题亟待解决。

本文提出一种基于 Leap Motion 的悬浮真3D显示实时手势交互系统,可实现多个悬浮3D物体的独立实时交互,交互帧率为30帧/s。

2 系统结构与工作原理

2.1 系统结构

悬浮真3D显示手势交互系统由集成成像

3D显示屏、二面角反射镜阵列、Leap Motion 和计算机组成,如图1所示。系统原理如图2所示。集成成像3D显示屏与二面角反射镜阵列共同构成系统的显示模块。如图3所示,二面角反射镜阵列是一种悬浮器件^[6-8],将其倾斜45°放置能够使集成成像3D显示屏再现的3D图像“悬浮”于该器件前方^[9];悬浮3D图像相比原3D图像深度反转,可抵消传统集成成像3D显示的深度反转现象,最终形成深度关系正确的悬浮3D图像。

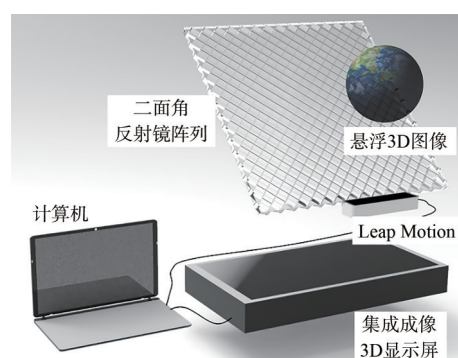


图1 悬浮真3D显示手势交互系统结构

Fig. 1 Structure of the floating 3D display gesture interaction system

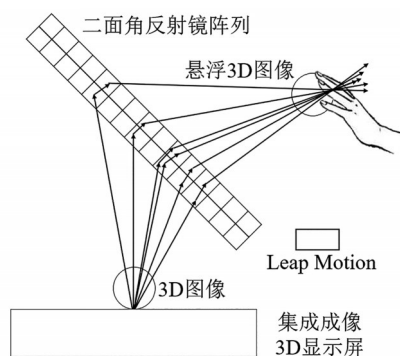


图2 交互系统原理示意图

Fig. 2 Principle of the proposed interaction system

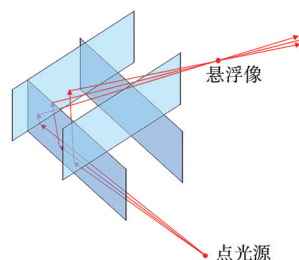


图3 二面角反射镜原理示意图

Fig. 3 Principle of the dihedral corner reflector

Leap Motion 连接计算机,两者共同组成系统的交互模块。Leap Motion 是一款体感控制器,它内置双目相机,基于计算机视觉算法采集手的数量、各个关节节点的空间位置、手指和手掌的朝向等数据,通过应用程序编程接口(API)被计算机读取。为充分利用 Leap Motion 的探测范围,我们将 Leap Motion 水平放置在二面角反射镜阵列底边的正前方。悬浮真 3D 显示手势交互系统的手势交互流程如图 4 所示。

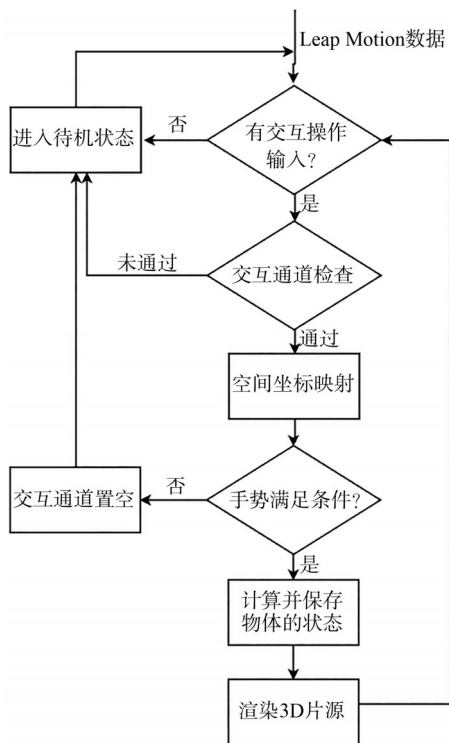


图 4 手势交互流程图

Fig. 4 Flow diagram of the gesture interaction

在一次交互循环流程中,计算机通过 API 从 Leap Motion 读取手势数据,然后利用交互通道机制检查物体是否被占用。若未被占用,将读取到的关节空间坐标映射到开放图形库(OpenGL)世界坐标系下,与预设手势进行比对,确定交互类型。之后计算并更新每个参与交互的物体的空间位置与姿态,并渲染 3D 片源,刷新显示在 LCD 屏上,使得悬浮 3D 图像中的物体发生位移、旋转等变化,实现人机交互。

集成成像 3D 显示屏的 3D 片源即图像阵列。为确保实时性和手势交互模块设计的便捷性,采用 OpenGL,通过设置多视口模拟相机阵列的拍

摄,生成图像阵列。渲染过程应用了双缓冲和顶点缓冲对象(VBO)技术,以保证 3D 片源的实时生成。

2.2 显示空间与交互空间坐标映射

为精确模拟现实世界的“触碰”等手势交互动作,需要建立显示空间与交互空间的关联。显示空间指悬浮 3D 图像所处的空间,交互空间指 Leap Motion 可探测到的手关节所处的空间。然而手关节坐标基于 Leap Motion 自身的坐标系,无法直接得到该坐标与悬浮 3D 图像之间的坐标关系。因此需要将 Leap Motion 坐标映射到显示空间,才能进行触碰与否的判断。根据集成成像原理可知,3D 拍摄与显示过程互逆,显示空间坐标系与渲染过程的 OpenGL 世界坐标系统一,悬浮 3D 图像坐标与待渲染的 3D 模型坐标完全一致。因此,只要确定 Leap Motion 坐标系和 OpenGL 世界坐标系的映射规则,即可实现显示空间与交互空间的坐标映射。

假设 OpenGL 世界坐标系 $O-XYZ$ 的 OX 、 OY 、 OZ 轴在 Leap Motion 坐标系 $O'-X'Y'Z'$ 下的单位方向矢量分别为 $(X'_x, Y'_x, Z'_x)^T$, $(X'_y, Y'_y, Z'_y)^T$, $(X'_z, Y'_z, Z'_z)^T$, 坐标系 $O-XYZ$ 与 $O'-X'Y'Z'$ 的标度比为 K_s , 则可计算出从坐标系 $O'-X'Y'Z'$ 到 $O-XYZ$ 的自由向量变换矩阵 $C^{[10]}$:

$$C = K_s \begin{bmatrix} X'_x & Y'_x & Z'_x & 0 \\ X'_y & Y'_y & Z'_y & 0 \\ X'_z & Y'_z & Z'_z & 0 \\ 0 & 0 & 0 & 1/K_s \end{bmatrix}. \quad (1)$$

在 Leap Motion 中经过标定得到原点 O' 的坐标,记作 $(X'_{origin}, Y'_{origin}, Z'_{origin})^T$, 则点 O' 映射到 OpenGL 世界坐标系的变换矩阵为:

$$T = \begin{bmatrix} 1 & 0 & 0 & -X'_{origin} \\ 0 & 1 & 0 & -Y'_{origin} \\ 0 & 0 & 1 & -Z'_{origin} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

由公式(1)和(2),可将 Leap Motion 坐标映射到 OpenGL 世界坐标系中:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = C \cdot T \cdot \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}. \quad (3)$$

2.3 手势交互的实现

本系统预设了 3 种手势交互方式:抓取平移、

划动旋转和按键平移/旋转。3种手势的交互过程均包括3个阶段:选定物体、发生运动、释放物体。以抓取平移动作为例,人在用手移动一个体积较小的物体时,主要是单手的拇指、食指、中指弯曲并发力,夹持物体,移动一定距离后,手指略微伸直而释放。那么手指弯曲到一定程度夹持住物体为选定对象,移动为发生运动,手指松开则为释放对象。

本文以抓取平移为例介绍交互实现过程,其他两种手势原理与此类似,不再赘述。

首先选定对象,采用地球和月球两个球体模型,两个球体球心均位于各自模型局部坐标系的原点处,其在OpenGL世界坐标系的坐标表示为 $(X_{obj}, Y_{obj}, Z_{obj})^T$ 。当手的映射坐标处在以 $(X_{obj}, Y_{obj}, Z_{obj})^T$ 为中心、半径为 d_f 的球域内,并做出抓握动作时,视为选定交互对象。选定交互对象的位置判断条件为:

$$\left| C \cdot T \cdot \begin{bmatrix} X'_f \\ Y'_f \\ Z'_f \\ 1 \end{bmatrix} - \begin{bmatrix} X_{obj} \\ Y_{obj} \\ Z_{obj} \\ 1 \end{bmatrix} \right| \leq d_f, \quad (4)$$

其中 $(X'_f, Y'_f, Z'_f)^T$ 表示将拇指 $(X'_{thumb}, Y'_{thumb}, Z'_{thumb})^T$ 、食指 $(X'_{index}, Y'_{index}, Z'_{index})^T$ 和中指 $(X'_{middle}, Y'_{middle}, Z'_{middle})^T$ 作为整体考虑的一个加权平均坐标,其表达式为:

$$\begin{bmatrix} X'_f \\ Y'_f \\ Z'_f \\ 1 \end{bmatrix} = A_{thumb} \begin{bmatrix} X'_{thumb} \\ Y'_{thumb} \\ Z'_{thumb} \\ 1 \end{bmatrix} + A_{index} \begin{bmatrix} X'_{index} \\ Y'_{index} \\ Z'_{index} \\ 1 \end{bmatrix} + A_{middle} \begin{bmatrix} X'_{middle} \\ Y'_{middle} \\ Z'_{middle} \\ 1 \end{bmatrix}, \quad (5)$$

其中 $A_{thumb} + A_{index} + A_{middle} = 1$ 。除此之外还要满足3个手部形态判断条件:

$$\left| \begin{bmatrix} X'_{thumb} \\ Y'_{thumb} \\ Z'_{thumb} \\ 1 \end{bmatrix} - \begin{bmatrix} X'_f \\ Y'_f \\ Z'_f \\ 1 \end{bmatrix} \right| \leq d_{thumb}, \quad (6)$$

$$\left| \begin{bmatrix} X'_{index} \\ Y'_{index} \\ Z'_{index} \\ 1 \end{bmatrix} - \begin{bmatrix} X'_f \\ Y'_f \\ Z'_f \\ 1 \end{bmatrix} \right| \leq d_{index}, \quad (7)$$

$$\left| \begin{bmatrix} X'_{middle} \\ Y'_{middle} \\ Z'_{middle} \\ 1 \end{bmatrix} - \begin{bmatrix} X'_f \\ Y'_f \\ Z'_f \\ 1 \end{bmatrix} \right| \leq d_{middle}, \quad (8)$$

当拇指、食指和中指远节的中心点与它们加权平均坐标的距离小于各自的给定阈值 d_{thumb} 、 d_{index} 、 d_{middle} 时,视为抓取到该物体。

在运动发生过程中,3D图像跟随手部动作发生平移运动。两帧之间的手部位移被映射到OpenGL坐标系上,位移值表示为 $(\Delta X_f, \Delta Y_f, \Delta Z_f)^T$,对应模型跟随手部同步运动,位移值同样为 $(\Delta X_f, \Delta Y_f, \Delta Z_f)^T$ 。

最后阶段为释放对象,只需使选定物体的4个条件不同时满足。需要注意的是,手在“抓取”和“释放”的过程中,拇指与其余4指的运动幅度不同,拇指运动距离明显小于其余4指。因此,在计算平均坐标 $(X'_f, Y'_f, Z'_f)^T$ 时给予权重系数 A_{thumb} 更大数值。

2.4 多物体独立交互

多物体独立交互指同一时刻3D场景中的每个物体都能独立响应一只手的手势交互,互不干扰。当交互过程中手与手、物体与物体的空间位置非常接近甚至部分重合时,由于Leap Motion动态的手势数据存在不可避免的漂移和抖动,即使采用物理碰撞检测,也会出现手势选定物体的错乱,无法达到理想的交互效果。本文提出通道模型解决多物体交互串扰问题。如图5所示,每个待交互物体参与手势交互的状态都被储存,称为交互通道。交互通道在物体未参与交互时取值为-1,表示“未占用”;参与交互时取值为与物体交互的手的标识ID,表示物体被占用并正在参

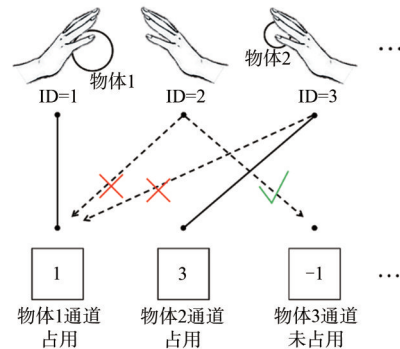


图5 多物体交互的通道模型

Fig. 5 Channel model for multiple objects interaction

与交互。标识 ID 在每只手进入 Leap Motion 交互范围时被自动分配。

当一只手进入某一物体的交互范围时,需要遍历所有物体的交互通道,增加额外的判断条件决定物体是否受到该手的控制:(1)物体的交互通道如果被除当前手以外的手的 ID 占用,物体不会对当前手的手势产生响应;(2)物体的交互通道被当前手的 ID 占用,物体将保持接受当前手的控制;(3)物体的交互通道没有被任何手的 ID 占用,则当前手势图与该物体建立新的控制与被控关系,即有可能将当前手的 ID 放入对应的通道内;(4)当前手的 ID 未在其他物体的交互通道里,建立一一对应的控制与被控关系;(5)多只手同时进入物体的交互范围,按照 ID 从小到大分配从高到低的优先级,这是由于先被 Leap Motion 追踪的手 ID 更小,后被追踪的更大,满足“先来后到”的规则。

上述一系列判断流程在交互的选定物体阶段完成。只有建立物体与手的交互关系后,才会进入下一阶段,计算物体对交互手势产生的响应。

3 实验结果

实物系统如图 6 所示。右侧为悬浮 3D 显示模块,从上至下依次为二面角反射镜阵列和集成成像 3D 显示屏。集成成像 3D 显示屏内使用的 LCD 屏与左侧的便携计算机连接。二面角反射镜阵列倾斜 45° 放置。为便于展示光路,Leap Motion 被置于桌面。实际交互时,Leap Motion 被水平放置在二面角反射镜阵列底边的正前方。悬浮 3D 图像位于二面角反射镜阵列前方,其中

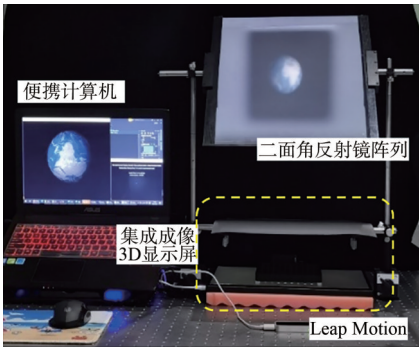


图 6 系统实物图

Fig. 6 Experiment setup of the proposed system

心与二面角反射镜阵列中心的距离为 259 mm。实验系统的其他参数如表 1 所示。

表 1 系统关键参数

Tab. 1 Parameters of the proposed system

便携计算机		二面角 反射镜阵列	Leap Motion	
CPU	GPU	有效尺寸/ mm	双目相机 视角/(°)	最大探测 距离/cm
Intel CORE i5 7 Gen	NVIDIA GeForce RTX1050	300×300	140×120	60~80

多物体独立交互效果如图 7 所示。悬浮 3D 图像包括“地球”、“月球”、简易的按键和状态栏。“月球”的独立交互实验效果如图 7(a)和 7(b)所示,交互动作作为划动旋转。从“月球”表面的纹理变化可以看出,“月球”跟随用户食指逆时针绕自轴旋转,此时“地球”表面纹理未发生变化。“地球”的独立交互实验效果如图 7(c)和 7(d)所示,交互动作作为抓取平移,“地球”跟随做出抓握动作的手自左下向右上发生平移,此时“月球”位置未发生变化。以上实验结果表明,所提系统可实现对 3D 场景内多物体的独立交互。在 3D 片源分辨率为 3 840×2 160 像素下,系统交互帧率为 30 帧/s。

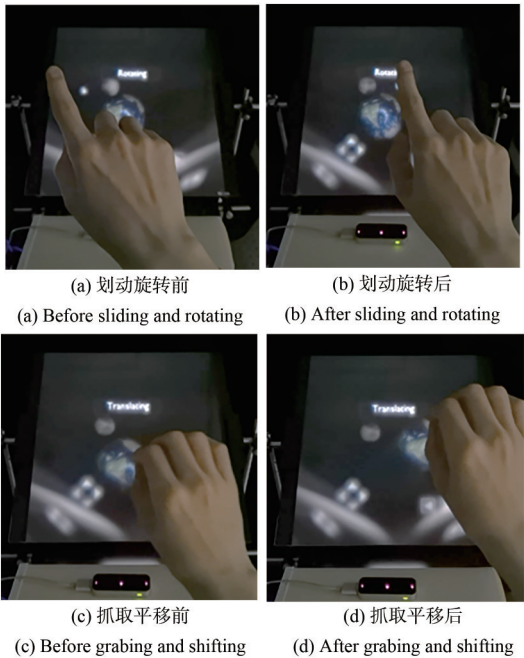


图 7 手势交互效果图

Fig. 7 Effect of the gesture interaction

4 结 论

本文提出了一种基于 Leap Motion 的悬浮真 3D 显示实时手势交互系统。该系统能够实现不同手势的实时交互,3D 片源分辨率为 $3\,840\times$

2 160 时,交互帧率为 30 帧/s。系统还可满足多手同时对多物体的独立交互,交互动作相互不受干扰,解决了多物体独立交互存在的串扰问题,使得交互更加灵活。所提系统为悬浮 3D 显示人机交互的应用提供了可能性。

参 考 文 献:

- [1] 王琼华. 3D 显示技术与器件[M]. 北京:科学出版社,2011.
WANG Q H. 3D *Display Technology and Device* [M]. Beijing: Science Press, 2011. (in Chinese)
- [2] 薛钊鸿,孙梓瀚,成泽锐,等. 基于人眼检测优化的立体人机交互技术的研究[J]. 液晶与显示,2018,33(11): 958-964.
XUE Z H, SUN Z H, CHENG Z R, *et al.* 3D human-machine interaction based on human eye detection [J]. *Chinese Journal of Liquid Crystals and Displays*, 2018, 33(11): 958-964. (in Chinese)
- [3] CHEN C H, TING C H, HUANG Y P. 3D interactive system using integral photography by embedded optical sensors for portable devices [J]. *Journal of Display Technology*, 2016, 12(11): 1329-1334.
- [4] CHAVARRÍA I A S S, NAKAMURA T, YAMAGUCHI M. Interactive optical 3D-touch user interface using a holographic light-field display and color information [J]. *Optics Express*, 2020, 28(24): 36740-36755.
- [5] 李炜,李海峰,徐良,等. 基于体感检测器件的悬浮真三维显示力触觉交互[J]. 光学学报,2020,40(9):0915004.
LI W, LI H F, XU L, *et al.* Suspended three-dimensional display haptic interaction based on Leap Motion [J]. *Acta Optica Sinica*, 2020, 40(9): 0915004. (in Chinese)
- [6] MAEKAWA S, NITTA K, MATOBA O. Transmissive optical imaging device with micromirror array [C]//*Proceedings of SPIE 6392, Three-Dimensional TV, Video, and Display V*. Boston: SPIE, 2006: 63920E.
- [7] MAEKAWA S, NITTA K, MATOBA O. Advances in passive imaging elements with micromirror array [C]//*Proceedings of SPIE 6803, Stereoscopic Displays and Applications XIX*. San Jose: SPIE, 2008: 68030B.
- [8] MIYAZAKI D, MAEDA Y, MAEKAWA S. Floating three-dimensional image display using micro-mirror array imaging element [C]//*Proceedings of SPIE 9495, Three-Dimensional Imaging, Visualization, and Display 2015*. Baltimore: SPIE, 2015: 949508.
- [9] 张汉乐. 集成成像 3D 显示性能提升方法[D]. 成都:四川大学,2020.
ZHANG H L. Methods for improving performance of integral imaging 3D display technology [D]. Chengdu: Sichuan University, 2020. (in Chinese)
- [10] DUNN F, PARBERRY I. *3D Math Primer for Graphics and Game Development* [M]. 2nd ed. Boca Raton: CRC Press, 2011.

作者简介:



林星雨(1998—),男,广东佛山人,硕士研究生,2021 年于北京航空航天大学获得学士学位,主要从事 3D 显示方面的研究。E-mail: linxingyu1998@buaa.edu.cn



王琼华(1969—),女,四川巴中人,博士,教授,2001 年于电子科技大学获得博士学位,主要从事显示与成像方面的研究。E-mail: qionghua@buaa.edu.cn