

基于互反射先验的端到端沉浸式投影补偿

胡宗禹¹, 程鹏², 刘建³, 刘洪^{1,4}, 陈虎¹, 张意⁵, 杨红雨^{1,4}

(1. 四川大学计算机学院, 成都 610065; 2. 四川大学空天科学与工程学院, 成都 610065;

3. 中国民用航空局空中交通管理局空域管理中心, 北京 100000;

4. 四川大学视觉合成图形图像技术重点学科实验室, 成都 610065;

5. 四川大学网络空间安全学院, 成都 610065)

摘要: 沉浸式投影系统中的互反射现象会导致投影质量下降, 投影补偿任务通过对输入图像的修改补偿, 减少互反射造成的影响从而得到更好的成像效果. 投影补偿任务是研究输入原图与最终成像之间的关系, 通过对输入图像的修改补偿, 减少互反射造成的影响从而得到更好的成像效果. 本文提出互反射先验, 互反射先验是指发生互反射的区域的临近区域通道像素值十分接近最大值 256, 因此通过图像的区域通道像素值可以获得互反射信息. 本文提出的互反射补偿网络(IRCIN), 通过互反射先验生成互反射掩膜, 消除沉浸式投影环境中的多余互反射. 为了验证 IRCIN 模型, 使用了超过 5000 对投影图像数据集和曲面投影系统以及折面投影系统来完成实验. 本文将 IRCIN 对比多种现有投影补偿方法, 实验结果表明 IRCIN 在 PSNR, RMSE, SSIM, 均值标准差分析等客观评价指标中具有明显优势, 说明本文提出的 IRCIN 能够有效利用互反射先验信息, 对沉浸式投影环境中的互反射现象进行补偿, 在均衡训练时间的基础上增强了投影效果.

关键词: 投影补偿; 神经网络; 互反射; 沉浸式投影

中图分类号: TP389.1 **文献标志码:** A **DOI:** 10.19907/j.0490-6756.2024.012001

End-to-end projection compensation in immersive projection systems based on the inter-reflection prior

HU Zong-Yu¹, CHENG Peng², LIU Jian³, LIU Hong^{1,4},
CHEN Hu¹, ZHANG Yi⁵, YANG Hong-Yu^{1,4}

(1. College of Computer Science, Sichuan University, Chengdu 610065, China;

2. School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China;

3. Airspace Management Center, Air Traffic Management Bureau, Beijing 100000, China;

4. State Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu 610065, China;

5. School of Cyber Science and Engineering, Sichuan University, Chengdu 610065, China)

Abstract: The inter-reflection phenomenon in immersion projection systems seriously affects the effect of projection, the projection compensation task can reduce the influence caused by the inter-reflection by modifying and compensating the input image so as to obtain a better imaging effect. In this paper, an inter-reflection prior is proposed, which means that the color effect of the region where inter-reflection oc-

收稿日期: 2022-10-06

基金项目: 国家自然科学基金民航联合研究基金(U1833128); 成都市重点研发支撑计划项目(2021YF0501788SN)

作者简介: 胡宗禹(1998-), 男, 四川宜宾人, 硕士研究生, 研究方向为图像处理与神经网络.

E-mail: 2020223040012@stu.scu.edu.cn

通讯作者: 陈虎. E-mail: huchen@scu.edu.cn

curs is very close to the maximum 256 due to the light superposition effect. Therefore, inter-reflection information can be obtained through the regional channel pixel of the image. The Inter-Reflection Compensation Network (IRCN) proposed in this paper generates an inter-reflection mask value through the inter-reflection prior, so that the IRCN can eliminates the excess interreflection in immersive projection environments. To verify the IRCN model, the experiment were performed on more than 5000 pairs of projected image datasets and curved projection system, and fold projection system. Compared with the existing projection compensation methods, the experiment results show that IRCN has obvious advantages in objective evaluation indicators such as PSNR, RMSE, SSIM, mean and standard deviation analysis, The proposed IRCN can effectively utilize the inter-reflection prior information to compensate for the inter-reflection phenomenon in the immersive projection environment, and enhance the projection effect based on balancing the training time.

Keywords: Projection compensation; Neural network; Inter-reflection; Immersive projection

1 引言

近年来,沉浸式投影系统发展迅速,特别是各种包围式投影系统飞速兴起。塔台模拟系统、飞行模拟系统等对投影质量有一定要求的大型投影系统在国内外广泛应用,使得投影质量重建工作尤为重要,因此投影补偿成为一个重要的课题^[1-3]。

投影互反射是指入射光在几个反射面间多次反射的现象,在沉浸式投影系统,特别是上文提到的大型包围式投影环境中互反射现象普遍存在并成为影响投影成像质量的主要因素之一。投影补偿是指建模整个投影系统,从输入图片到最终成像的整体改变中得到一定的规律,进而补偿投影仪的输入图片,达到增强投影图像质量的过程。

最近卷积神经网络在投影图像质量增强方面的应用展现了巨大的潜力。Huang 等提出的 CompenNet^[4]第一次使用神经网络解决投影质量重建问题,这种方法假设投影面是不平整且有纹理的,使用一张合适的投影面背光图片作为输入,通过相同的下采样设计,在对输入图下采样的同时补偿投影面的纹理和背景。Li 等人^[5]提出的 SRCN 首次在投影补偿任务中考虑人眼感知, SRCN 在输入端进行伽马矫正,并且在 CompenNet 的基础上增加了超分层和感知损失函数尝试提升图像生成效果。Lei 等人^[6]提出的基于注意力机制的卷积神经网络 PairNet 首次提出使用先验信息对神经网络进行优化设计,类似于通道注意力机制,该方法先验信息的计算区域为 1×1 区域,采用这样的先验信息的注意力机制对互反射现象达到抑制效果。

受到上述工作的启发,综合考虑多种方法的优劣势,本文提出一种全新的互反射先验,并且基于

互反射先验提出一个投影补偿卷积神经网络。该网络通过互反射先验信息,对受到互反射污染的图片区域精确补偿互反射影响,并且通过一个几何矫正网络的设计,实现端到端的投影图像质量重建。

综上所述,本文有以下贡献:(1) 在多种沉浸式投影环境中分析研究了互反射问题,提出互反射先验用于从投影图像中直接获取互反射信息;(2) 构建了全新的端到端 IRCN 模型,通过融合几何矫正网络实现真正端到端训练网络,降低数据处理工作量;(3) 提出的 IRCN 模型生成的补偿图像重投影质量优于主流基线,能够利用互反射先验信息有效消除互反射现象。

2 相关工作

2.1 传统方法

传统投影补偿方法假设投影仪输入图像和相机捕获图像之间的像素存在一一对应关系,然后建立对应的光传输矩阵^[7,8]。Ren 等人^[9]根据这一假设提出求解光传输矩阵的逆矩阵从而实现投影质量重建的方法。然而投影仪,摄像机各自的双边反射以及光线在投影面上的反射导致像素对应的假设存在一定问题,并且光传输矩阵过于复杂,几乎难以求得精确解。为了解决这些问题,不依赖构建光传输矩阵的方法被提出, Bimber 等人^[10]提出一种建立真实反射模型方法,通过计算互反射,将原图像素对应减去计算所得互反射,直接补偿相应的互反射。Li 等人^[11]提出的方法采用投影仪摄像机系统进行非数值计算的物理矩阵向量乘法模拟,同时该方法采用一种移位克雷洛夫子空间方法来消除物理乘法模拟和非负裁剪过程中产生的累积噪声,从而提升投影质量。但是这些方法很难直接估

计互反射情况, 模拟计算的补偿量并不精确, 在实际环境中容易导致过度补偿等问题。

大规模的矩阵面临计算时间成本过高, 补偿效果不佳等问题, 因此一些依赖额外设备的方法被提出, Takeda 等人^[12]提出了基于投影相机对的空间反射模型, 该模型用一种光致变色化合物绘制屏幕, 当紫外线照射时, 这种化合物会改变颜色。紫外线照射可以通过屏幕后面的紫外线 LED 阵列来控制, 通过借助这些信息达到补偿的效果。这些方法往往不仅需求包括结构光探测器, 紫外线阵列板等昂贵的额外设备, 而且需要精确的几何和光度校准。

2.2 深度学习方法

随着神经网络的发展, 深度学习方法在图像处理方面的应用愈发成熟, 一些基于深度学习的图像处理方法, 如图像风格迁移算法 pix2pix^[13], cyclegan^[14]等以及图像增强算法^[15, 16], 图像超分算法^[17-19]和基于注意力机制的神经网络算法^[20, 21]都展示了神经网络在投影补偿任务中的潜力和可能性。

在投影补偿方面, Huang 等人^[4]第一次提出了 CompenNet 使用深度学习方法解决投影质量增强问题。这种方法采用了类似 U-net^[22]的跳跃连接, 采用卷积神经网络分别采样投影面背光图像和摄像机捕获投影图像, 通过一系列卷积操作, 以投影仪输入图像作为监督, 端到端地完成了投影补偿任务。这种方法设计了背光投影表面图像采样网络, 但是对于同一种投影表面, 使用不同亮度, 色彩的环境光照射会导致结果的巨大差别, 很难找到合适的配置。Li 等人^[5]提出的 SRCN 在 CompenNet 的基础上增加了超分层和感知损失函数尝试提升图像生成效果, 并且在输入端进行伽马矫正使补偿效果更贴合人眼感知, 但是这样的网络结构和 CompenNet 区别并不明显, 在客观评价指标上的提升也有限。Lei 等人^[6]提出的 PairNet 采用先验信息, 并设计两个光度补偿子网分别对前景和背景光针对性补偿, 并融合两个子网的输出与先验信息的乘积作为输出。该方法使用的先验信息仅仅考虑单点像素, 而且通过先验信息获得的掩膜直接与生成图像乘积的做法有待商榷, 并且有着网络较大, 训练时间较长以及数据处理较复杂等问题。这些深度学习方法往往没有针对考虑区域性的互反射问题, 并且普遍面临需要裁剪数据集, 数据整理工作复杂的问题。

3 互反射补偿网络设计

3.1 互反射先验

互反射是沉浸式投影环境中普遍存在的现象, 本文提出的互反射先验受到暗通道先验^[23]和互反射通道先验^[6]的启发, 基于沉浸式投影系统中投影图像由于单点的像素受到复杂光反射影响可能受该点附近区域高亮度像素点影响, 导致最终投影图像比原始图像亮度普遍较高的事实。对于摄像机捕获的与原始图像像素对齐的投影图片 I_{cam} , x 点的互反射先验由以下公式计算而来:

$$\text{Mask}(x) = \frac{\sum_{i=1}^3 f_{y \in Z(x)}(i, \max_{c \in \{r, g, b\}}(I_{\text{cam}}^c(y)))}{3} \quad (1)$$

其中, c 是 y 点的 rgb 像素值; $Z(x)$ 是围绕 x 的 3×3 区域; $\max_{c \in \{r, g, b\}}(I_{\text{cam}}^c(y))$ 是 y 点 rgb 像素通道中对应最大值的通道, 本文称这个值为互反射像素, $f_{y \in Z(x)}(i, c)$ 是求在 $Z(x)$ 区域内一共 $3 \times 3 = 9$ 个点的互反射像素值从大到小排列后第 i 大的那一个。不同于两种通道先验方法只计算单点像素, 考虑到在投影任务中存在原本自身的通道像素值较高的情况, 因此本文设计了区域性先验。这样的互反射先验计算所得的互反射掩膜能够综合考虑到图像本身高像素区域成像效果以及互反射对成像的影响, 从而进一步增强互反射抑制效果。

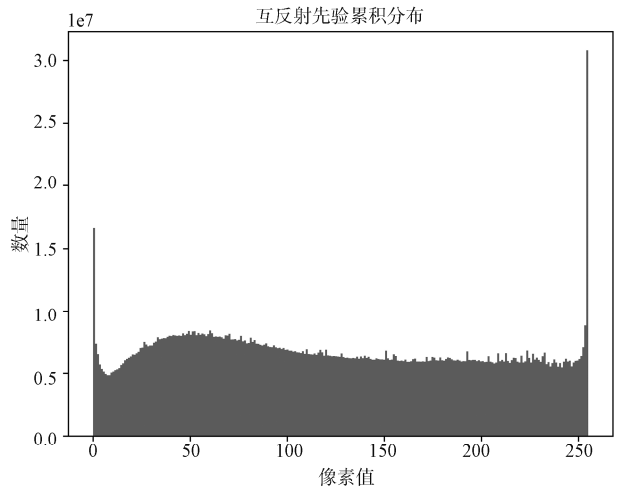


图 1 投影图像互反射先验累积分布
Fig. 1 Cumulative distribution of projected images inter reflection prior

在存在互反射现象的投影图像中, 部分局部的高亮度可能会影响周围大部分区域的投影质量, 互反射先验能够获取这些区域信息。本文分析了超过

4000 张投影图片的互反射先验值的累积分布,从图 1 可以看到有大量区域的互反射先验单点像素趋近 256 最大值,这些区域的成像质量受到极大影响。

3.2 投影系统的数学模型

本文的投影质量补偿系统是由一个投影仪-摄像机对和一个与投影仪距离方向固定的沉浸式投影屏幕构成,如图 2 和图 3 所示。为了在多维验证本方法的可行性,本文采用了曲面屏和 90°折面屏两种投影面构建了两套投影系统。

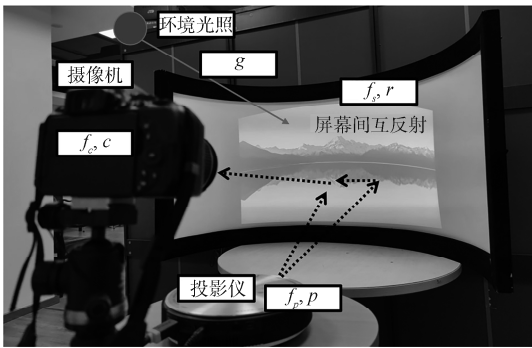


图 2 曲面沉浸式投影系统

Fig. 2 Curved surface immersion projection system

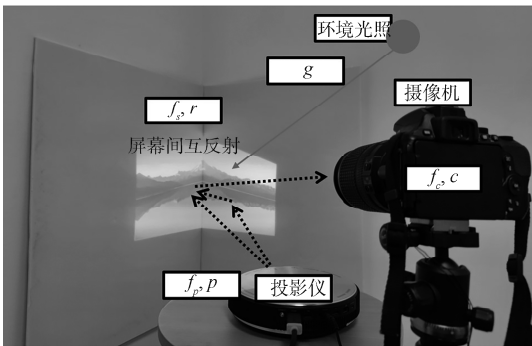


图 3 折面沉浸式投影系统

Fig. 3 Folded surface immersion projection system

本文的投影系统均可以描述为

$$y = G(f_c(f_s(f_p(x, p), g, r), c)) \quad (2)$$

其中, y 是与输入图像 x 像素对齐的摄像机捕获图像; x 是投影仪的输入图像; f_p 是投影仪的几何光度畸变函数; p 是几何光度畸变的参数; f_s 是投影面表面双向反射率分布函数; g 是全局环境光照参数; r 是互反射参数; f_c 是相机的复合捕获函数; c 是对应的参数; G 是从摄像机原始捕获图片到与输入图片像素对齐的几何畸变函数。需要注意在投影补偿流程中输入图片为原始图片,不存在几何畸变,因此在补偿流程里输入图片不需要几何校正。由于沉浸式投影环境一般要求在全局光的情况

下部署,本文的投影系统控制全局环境光照参数 $g=0$,同时将各个函数的参数融合进函数中,保留互反射参数,将 f_c, f_s, f_p 融合成一个函数 F :从投影仪输入图片到摄像机捕获图片的光度偏移函数,能得到式(3)。

$$y = G(F(x, r)) \quad (3)$$

考虑互反射先验信息,用 $mask$ 表示与输入图像 x 像素对齐的互反射掩膜,可以将式(3)写为式(4):

$$y = G(F(x) + mask) \quad (4)$$

即相机捕获的与输入图像像素对齐的图像是由输入图像的正常光度变化与互反射额外污染共同作用形成。

投影图像补偿的目的是找到一张补偿图像 x^* ,使得最终成像效果能达到投影仪输入图像 y^* 的水平即:

$$y^* = G(F(x^*) + mask) \quad (5)$$

由于投影仪输入图像已知,因此在补偿流程中补偿图像 x^* 可以表示为:

$$x^* = F^{-1}(G^{-1}(y^*) - mask) \quad (6)$$

3.3 基于深度学习的公式

由于前文提到,互反射信息可以从摄像机捕获图像中计算得到,将从该图像到对应的互反射掩膜的转换函数写为 M ,神经网络训练过程可以从式(6)表示为:

$$\hat{x} = F^{-1}(G^{-1}(y) - \lambda \times M(G^{-1}(y))) \quad (7)$$

其中, \hat{x} 为训练中网络整体输出; y 为摄像机捕获图像,即 λ 控制互反射先验信息对最终图像的影响程度,并产生对应的互反射掩膜,作为先验信息加入网络训练中。将 F^{-1}, G^{-1}, M 建模为卷积神经网络 π_θ :

$$\hat{x} = \pi_\theta(y) \quad (8)$$

其中, $\theta = \{\theta_F, \theta_G, \theta_M\}$ 包含三个网络的训练参数,投影的正过程为从原图像到摄像机捕获图像,而 π_θ 为从摄像机捕获图像 y 到输入投影仪的原图像 y^* 的过程,因此 π_θ 可以直观理解为投影逆过程, \hat{x} 是对应整体网络输入 y 的补偿图像。因此本文的神经网络由原图像,摄像机捕获图像对 (y^*, y) 作为输入进行训练,将训练集分为 N 组原图像,摄像机捕获图像对,则有训练集 $\{(y_i^*, y_i)\}_{i=1}^N$,整体损失函数如下。

$$L_{\text{total}} = L_1 + L_2 + L_{\text{SSIM}} + \lambda_s \times L_2^{\text{mask}} \quad (9)$$

式(9)中, $L_1, L_2, L_{\text{SSIM}}$ 分别为整体网络的 L_1 损失函数, L_2 损失函数和 SSIM 损失函数,具体表示如下。

$$L_1 = \frac{1}{N} \sum_{i=1}^N |\hat{x}_i - y_i| \quad (10)$$

$$L_2 = \frac{1}{N} \sum_{i=1}^N \|\hat{x}_i - y_i\|^2 \quad (11)$$

$$L_{SSIM} = \frac{1}{N} \sum_{i=1}^N (1 - SSIM(\hat{x}_i, y_i)) \quad (12)$$

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \times \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (13)$$

其中, y 是输入投影仪的原图像; \hat{x} 为网络整体输出图像; μ_x 为图像 x 的均值; μ_y 为图像 y 的均值; σ_x 为图像 x 的方差; σ_y 为图像 y 的方差; σ_{xy} 为 x 与 y 的协方差; C_1, C_2 是为了保持稳定避免分母为 0 的两个常数. 式(9)中的 L_2^{mask} 为控制互反射生成网络输出的损失函数; λ_s 为对应的参数; L_2^{mask} 表示

如下.

$$L_2^{mask} = \frac{1}{N} \sum_{i=1}^N (Mask(y_i) - m_i^*)^2 \quad (14)$$

其中, $Mask(y_i)$ 为通过网络输入图像 y 几何校正后计算所得的互反射先验, 具体计算方式见式(1); m^* 为互反射掩膜生成网络的输出.

3.4 卷积神经网络设计

IRCN 由三部分: 一个几何矫正网络, 一个光度矫正网络和一个互反射掩膜生成网络组成. 在训练中, 摄像机捕获图像首先输入几何矫正网络, 生成与投影仪输入图像像素对齐的摄像机图像. 然后这个图像会输入互反射掩膜生成网络, 而互反射掩膜生成网络的输出又与这个图像共同产生光度矫正网络的输入. 具体的网络结构见图 4.

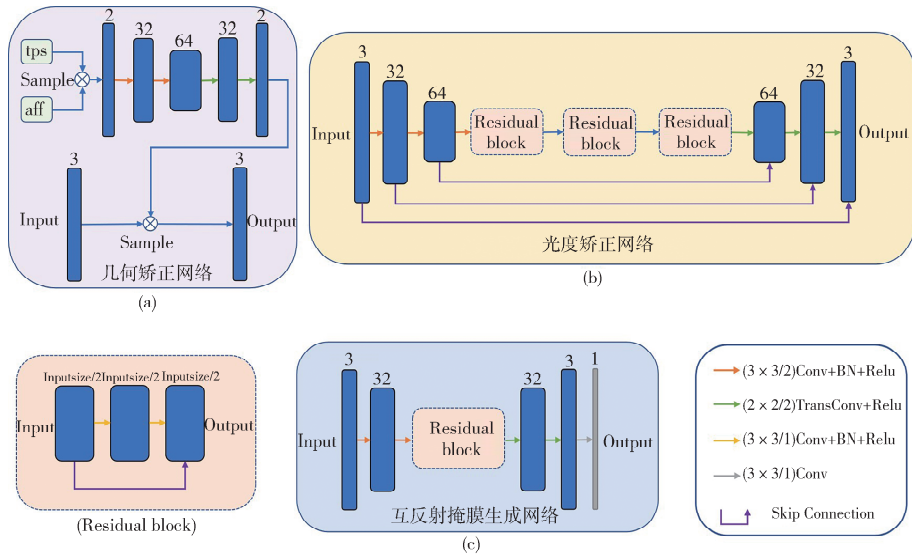


图 4 网络结构
Fig. 4 The network structure

由于沉浸式投影环境中, 输入图像最终成像形状会产生畸变, 摄像机捕获图像同时包含输入图像以外的屏幕背景信息, 所以需要使用几何矫正网络将摄像机捕获的图像转变为与投影仪输入图像像素对齐的图像. IRCN 的几何矫正网络首先包含薄板样条参数^[24]和仿射变换参数这两个可学习模块, 受其他几何矫正网络的启发^[5, 25, 26], 如图 4a 所示, IRCN 设计仿射变换参数 aff , 薄板样条参数 tps 两个可学习参数分别通过网络生成器生成对应的采样网格, 然后通过网格采样将两个网格结合后输入一个网格细化网络, 以调整整体矫正效果. 最终输出的采样网格通过双线性插值将摄像机捕

获图像几何矫正为与输入图像对应的图像.

如图 4c 所示, 互反射掩膜生成网络是一个轻量级网络, 由于投影系统中的真实互反射很难直接获取, 受到一些注意力机制方法^[27, 28]的启发, 本文设计互反射掩膜生成网络, 如图 5 所示. 采用摄像机捕获图像信息通过前文提到的互反射先验计算得到互反射先验图作为监督, 使用一个带参数 $\lambda_s = 0.2$ 调节其影响程度的 L_2 损失函数控制互反射掩膜生成网络的输出.

如图 4b, 按照前文的公式推导, IRCN 的光度矫正网络将前两个网络输出做差结合, 通过互反射先验信息, 将补偿了互反射影响的摄像机捕获图像

输入,在使用卷积和反卷积的下采样,上采样之间采用跳跃连接结构,具体结构参考图 4,采用三个残差块的卷积神经网络能够学习到对应的光度变化信息,在训练中,式(7)和图 5 里控制互反射先验信息的参数 λ 被设计为一个可学习参数,将其范围控制在 $[0.1, \text{init_value}]$, 初始值 init_value 被设计为 0.3.

综上所述,IRCN 训练中以摄像机捕获图像为

输入,投影仪输出图像为监督. 在投影补偿流程中,将投影仪输入图像作为训练好的 IRCN 的输入,如图 5 所示. 由于补偿流程中输入图像并不需要几何校正,因此忽略训练好的几何校正网络,直接输入网络其余部分,IRCN 生成相应的补偿图像,将该补偿图像输入投影仪以增强投影效果,整体的互反射补偿流程如图 5 所示.

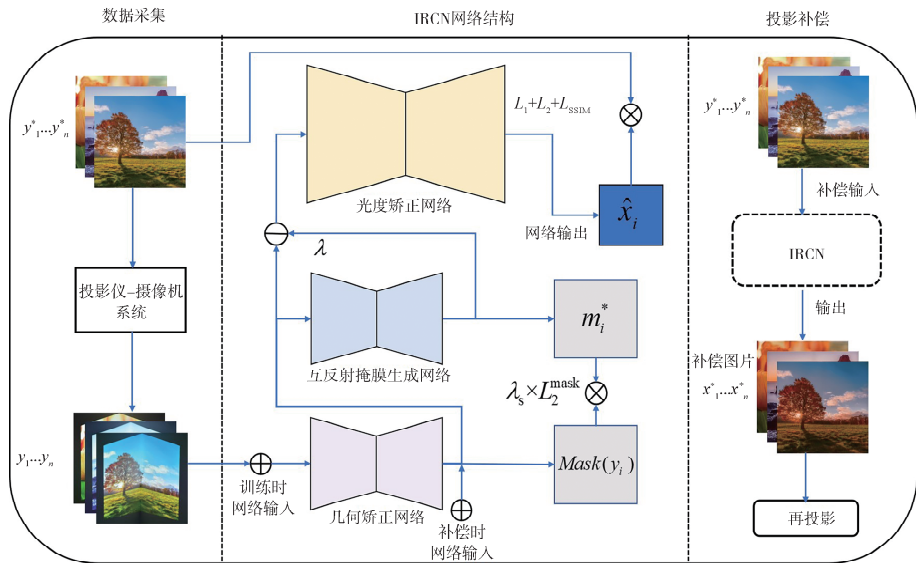


图 5 投影补偿流程

Fig. 5 Projection compensation process

4 实验过程和分析

4.1 评价指标和对比方法

本文对照一种传统方法 (Bimber 的方法^[10]) 和三种深度学习方法 (CompenNet^[4], SRCN^[5], PairNet^[6]), 采用常用于投影重建任务中的像素级别图像质量评价指标, PSNR, SSIM, RMSE 以及像素均值标准差分析. 所有方法都在曲面投影系统和折面投影系统中测试, 投影补偿测试流程如图 5 所示. 测试集经过再投影, 摄像机拍摄后采用已训练好的几何校正网络对采集图片进行校正, 与原图对齐后比较各个客观评价指标.

4.2 数据集

由于投影补偿任务缺乏公开数据集, CompenNet 提出的数据集在特殊纹理投影面上采集, 并不适用于本文实验要求, 因此本文构建了自己的数据集. 本文实验均控制环境光照, 在黑暗环境下进行数据采集, 为了确保方法的适用性, 分别在 90° 折面屏和曲面屏投影面上, 使用超过 4000 张 512×512 的多种类 RGB 图像进行训练, 并且额外采集

超过 500 张图像进行测试和结果分析.

本文的投影仪和摄像机分别是 JMGO G7 投影仪和 NIKON DX VR 相机. 本文的沉浸式投影系统中, 为了达到较好的投影效果和相机捕获效果, 摄像机和投影仪的距离控制在 60 cm, 投影面的中心与摄像机投影仪的垂直距离控制在 90 cm. 相机具体设置为: 自动白平衡, $1/45$ 快门速度, ISO 为 400, 对焦 $f=5.6$. 投影仪采用自动对焦, 投影模式为默认. 为了端到端训练, 摄像机图像在直接输入网络之前自动剪裁为 512×512 作为输入.

4.3 实现细节

对于传统方法, 本文直接采用原文方法进行测试. 而对于深度学习方法, 首先由于 CompenNet, SRCN, PairNet 三种方法都是光度矫正算法, 需要对采集图像几何校正才能进行训练. 因此本文单独采用本文方法中的几何校正网络对训练集和测试集几何校正, 形成新的数据集进行训练. 由于 CompenNet 是针对投影面有纹理和颜色的投影环境的方法, 因此需要一张背光投影表面图像作为额外输入, 实际上不同的表面图像会对最终的补偿效

果产生巨大的影响,因此本文测试多种灰度,色彩的表面图像,选择效果最佳的进行对比.所有方法的数据采集和投影补偿效果测试流程可见图 5.

本文所有深度学习方法都在 Pytorch 框架下,训练 20 个 epoch, batch-size 为 8, 采用 D Kingma 等提出的 Adam 函数进行优化,使学习率从 $1e-3$ 下降到 $1e-4$. 所有深度学习方法都按照以上的设置在两张 RTX2080ti 上完成.

4.4 现有方法对比

本文在客观评价指标的基准上,将本文提出的 IRCN 与四种目前主流的投影补偿方法进行了比较.对比的方法分别是 Bimber 提出的方法^[10], CompenNet^[4], SRCN^[5], Lei 提出的方法^[6]. 所有方法均通过投影补偿流程,采集测试集 500 张测试图像并矫正为 512×512 尺寸进行对比.图 6 展示这些方法在不同色调的投影图像上的视觉效果.

Bimber 的方法将投影面分为多个子区域直接计算区域间互反射.但是正如所有不依赖光传输矩阵的传统方法的弊病,该方法需要包括摄像机镜头畸变,投影面双边反射,投影补偿系统的相对坐标信息等额外细节.但是在实际场景中,这些信息很难准确获取,并且可能改变,同时计算所得互反射精确度也不能保证因此该方法在实际的沉浸式投影系统中的效果并不佳.如表 1 所示,在客观评价指标方面,本文的方法明显优于该方法.

本文通过对比三种深度学习方法,说明 IRCN 在投影质量增强和互反射补偿方面的优势.所有的深度学习方法都在统一数据集上训练.正如前文所说,CompenNet 和 SRCN 需要背景光信息,而采用不同背景光信息会导致实验结果的巨大差别,因此这两种方法不能兼顾多种不同色调的补偿效果,虽然 SRCN 通过感知损失函数和伽马矫正最终成像上有一定改善,但是整体的区别并不明显,如图 6b,由于背景信息的加入需要光照后采集,两种方法的低亮度区域被抑制,导致整体颜色偏暗,特别是低饱和度区域色调上表现较差,与原图的灰度相似性大幅下降,在客观评价指标上也劣于 IRCN. Lei 提出的方法由于采用 1×1 区域的最大像素值作为先验信息,没有考虑到局部区域互反射和原本高亮度区域,对局部互反射抑制并没有达到很好的效果.双子网的输出直接融合的网络设计导致针对前景与背景的补偿难以平衡,见图 6. 在各种色调下的生成图像亮度较原图更高,并且双子网设计让训练时间大幅度增加.而 IRCN 在输入端学习到互

反射信息,只使用一个光度补偿子网的设计有效补偿了相应的互反射.从图 7 可以看出,在像素总体分布上也更接近原图分布,并且兼顾了训练时间,在整体评价指标上领先于其他方法.

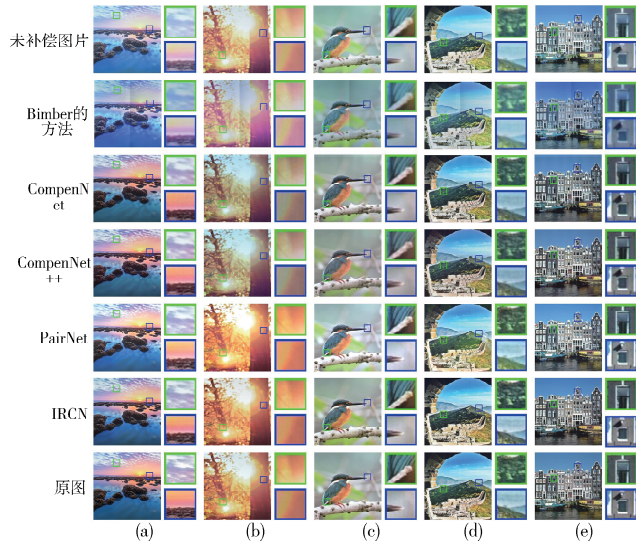


图 6 实验结果对比

Fig. 6 Comparison with state of the art

表 1 现有方法评价指标对比

Tab. 1 Comparison with state of the art

方法	PSNR	RMSE	SSIM	训练时间/min
未补偿	15.458	0.368	0.778	
Bimber 方法	16.044	0.327	0.774	
CompenNet	19.758	0.253	0.775	2h52
SRCN	20.194	0.242	0.821	3h22
PairNet	21.571	0.234	0.825	4h26
IRCN	21.807	0.229	0.848	3h18

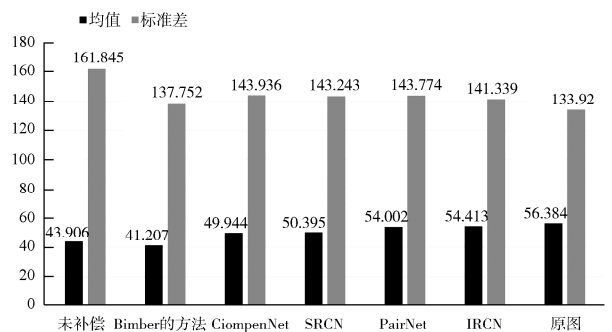


图 7 均值标准差分析

Fig. 7 Mean-standard deviation analysis

4.5 消融实验

本文设计了两个消融实验证明 IRCN 的合理

性. 第一个是直接采取计算所得互反射先验作为互反射掩膜代替掩膜生成网络, 在后文称为直接先验网络. 另一个是将互反射先验的监督计算从 3×3 的区域修改为 1×1 的区域, 在后文称为 1×1 网络.

由于互反射先验能够得到高亮度区域的互反射信息, 但是在沉浸式投影环境中仍然有一些受互反射影响较小, 像素过渡比较平滑的区域. 从图 8 可以看到, 直接先验网络采用计算得到的互反射先验作为掩膜可能会导致这些受互反射影响较小的区域的补偿量过大, 整体颜色偏暗, 从表 2 和图 8 也可以看出, 直接先验网络的评价指标表现也较差. 在图 8 中可以看到 1×1 网络没有考虑局部高亮度对附近区域的影响, 导致局部区域的像素补偿效果较差, 并且丢失原图部分细节, 存在互反射区域最终成像饱和度降低的现象.

从表 2, 图 8 和图 9 可以看出, 本文的互反射先验能够有效获取投影互反射信息, 通过 3×3 区域互反射先验对互反射掩膜生成网络的监督, 能够兼顾局部和整体的互反射抑制效果. 同时本文方法在最终成像质量上以及细节保留上有更好的表现.

表 2 消融实验客观评价指标对比

Tab. 2 Comparison of objective evaluation indexes of ablation experiment

方法	PSNR	RMSE	SSIM
完整网络	21.807	0.229	0.848
1×1 网络	20.943	0.248	0.844
直接先验网络	19.498	0.259	0.807



图 8 消融实验对比

Fig. 8 Ablation experiment comparison

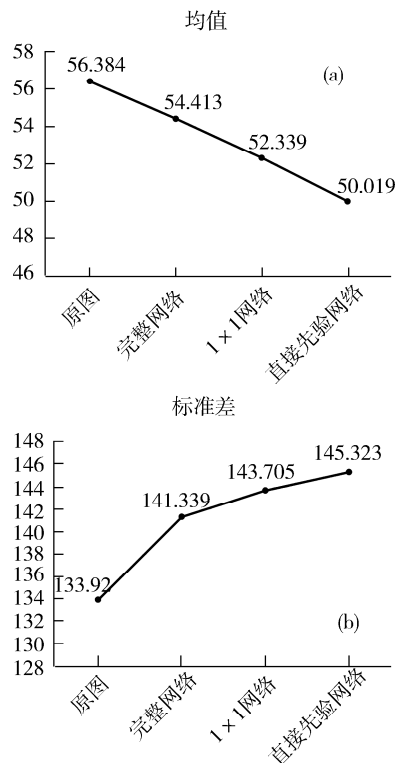


图 9 消融实验均值标准差分析

Fig. 9 Mean-standard deviation analysis of Ablation experiments

5 结论

本文提出了一个全新的基于互反射先验的沉浸式投影补偿网络 IRCN, 以解决沉浸式投影系统中的互反射问题. 采用提出的 3×3 区域互反射先验, 获取相应的沉浸式投影导致的互反射的信息, 同时本文提出的 IRCN 能够端到端训练网络, 融合几何矫正网络的整体结构使得数据处理难度降低, 并且使网络结果重投影的几何结构复原效果更良好. 本文在训练超过 4000 对图像的基础上, 分析对比 500 张测试图像. 与目前存在的投影补偿方法对比, 本文提出的 IRCN 能够有效抑制互反射的产生, 使得投影质量大幅提升, 色彩亮度更接近原图, 综合训练时间的客观评价指标表现也优于其他方法.

但是本文的工作仍然存在一定局限性, 首先, 本文的方法依然存在卷积神经网络图像生成局部模糊问题, 这个问题在大面积高分辨率投影中可能会比较明显. 其次, 包含几何矫正网络的整体网络在没有预训练的情况下收敛速度较慢. 上述问题都是后续研究的重点.

参考文献:

- [1] Fujimoto Y, Sawabe T, Kanbara M, *et al.* Structured light of flickering patterns having different frequencies for a projector-event-camera system [C]//Proceedings of the Conference on Virtual Reality and 3D User Interfaces. Christchurch: IEEE, 2022; 582.
- [2] Park J, Jung D, Moon B. Projector compensation framework using differentiable rendering [J]. IEEE Access, 2022, 10: 44461.
- [3] Narita G, Watanabe Y, Ishikawa M. Dynamic projection mapping onto deforming non-rigid surface using deformable dot cluster marker [J]. IEEE T Vis Comput Gr, 2017, 23: 1235.
- [4] Huang B, Ling H. End-to-end projector photometric compensation [C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019; 6810.
- [5] Li X F, Hu Z H, Gu H Y, *et al.* Multiprojector interreflection compensation using a deep convolution network [J]. Sci Programming-Neth, 2022, 2022: 1.
- [6] Lei Q H, Yang T, Cheng P. An attention mechanism based inter-reflection compensation network for immersive projection system [J]. J Imag Grap, 2022, 27: 1238. [雷清桦, 杨婷, 程鹏. 注意力机制的曲面沉浸式投影系统补偿. 中国图象图形学报, 2022, 27: 1238.]
- [7] Grundhofer A, Iwai D. Robust, error-tolerant photometric projector compensation [J]. IEEE T Image Process, 2015, 24: 5086.
- [8] Sajadi B, Lazarov M, Majumder A. ADICT: accurate direct and inverse color transformation [C]//Proceedings of the 11th European Conference on Computer Vision. Heraklion: Springer, 2010 (6314): 72.
- [9] Ren N, Ramamoorthi R, Hanrahan P. All-frequency shadows using non-linear wavelet lighting approximation [J]. ACM T Graphic, 2003, 22: 376.
- [10] Bimber O, Grundhofer A, Zeidler T, *et al.* Compensating indirect scattering for immersive and semi-immersive projection displays [C]//Proceedings of the IEEE Virtual Reality Conference (VR 2006). Alexandria: IEEE, 2006; 151.
- [11] Li Y, Yuan Q, Lu D. Perceptual radiometric compensation for inter-reflection in immersive projection environment [C]//Acm Symposium on Virtual Reality Software & Technology. Singapore: ACM, 2013; 201.
- [12] Takeda S, Iwai D, Sato K. Inter-reflection compensation of immersive projection display by spatiotemporal screen reflectance modulation [J]. IEEE T Vis Comput Gr, 2016, 22: 1424.
- [13] Isola P, Zhu J Y, Zhou T, *et al.* Image-to-image translation with conditional adversarial networks [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017; 5967.
- [14] 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. Venice: IEEE, 2017; 2242.
- [15] He L, Long W, Li Y Y, *et al.* A low-light image enhancement model based on gray level mapping and fusion technology [J]. J Sichuan Univ (Nat Sci Ed), 2022, 59: 022001. [何磊, 龙伟, 李炎炎, 等. 一种基于灰度映射与融合技术的微光图像增强模型[J]. 四川大学学报(自然科学版), 2022, 59: 022001.]
- [16] Liu S X, Long W, Li Y Y, *et al.* Non-linear low-light image enhancement based on fusion color model spade [J]. J Sichuan Univ (Nat Sci Ed), 2021, 58: 012003. [刘寿鑫, 龙伟, 李炎炎, 等. 融合彩色模型空间的非线性低照度图像增强[J]. 四川大学学报(自然科学版), 2021, 58: 012003.]
- [17] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks [C]//Proceedings of the 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016; 1646.
- [18] Wang X, Yu K, Dong C, *et al.* Recovering realistic texture in image super-resolution by deep spatial feature transform [C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018; 606.
- [19] Zhang D, Shao J, Hu G, *et al.* Sharp and real image super-resolution using generative adversarial network [C]//Proceedings of the International Conference on Neural Information Processing. Guangzhou: Springer, 2017, 3: 217.
- [20] Zhang Y, Kan Z W, Shao Z W, *et al.* Remote sensing image denoising based on attention mechanism

- and perceptual loss [J]. *J Sichuan Univ (Nat Sci Ed)*, 2021, 58: 042001. [张意, 阚子文, 邵志敏, 等. 基于注意力机制和感知损失的遥感图像去噪 [J]. *四川大学学报(自然科学版)*, 2021, 58: 042001.]
- [21] Jin R L, Qing L B, Wen H Q. Emotion recognition of the natural scenes based on attention mechanism and multi-scale network [J]. *J Sichuan Univ (Nat Sci Ed)*, 2022, 59: 012003. [晋儒龙, 卿粼波, 文虹茜. 基于注意力机制多尺度网络的自然场景情绪识别 [J]. *四川大学学报(自然科学版)*, 2022, 59: 012003.]
- [22] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation [C]//*Proceedings of the International Conference on Medical Image Computing and Computer-assisted Intervention*. Munich: Springer, 2015: 234.
- [23] He K, Sun J, Tang X. Single image haze removal using dark channel prior [J]. *IEEE T Pattern Anal*, 2011, 33: 2341.
- [24] Donato G, Belongie S. Approximate thin plate spline mappings [C]//*Proceedings of the 7th European Conference on Computer Vision*. Berlin, Heidelberg: Springer, 2002: 21.
- [25] Feng H, Zhou W, Deng J, *et al.* Geometric representation learning for document image rectification [C]//*Proceedings of the 17th European Conference on Computer Vision*. Cham, Switzerland: Springer, 2022.
- [26] Rocco I, Arandjelovic R, Sivic J. Convolutional neural network architecture for geometric matching [J]. *IEEE T Pattern Anal*, 2019, 41: 2553.
- [27] Cheng Q, Li H, Wu Q, *et al.* Learn to pay attention via switchable attention for image recognition [C]//*Proceedings of the 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR 2020)*. Shenzhen: IEEE, 2020: 291.
- [28] Fei W, Jiang M, Chen Q, *et al.* Residual attention network for image classification [C]//*Proceedings of the 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 3156.