# 基于GAN的阴影生成

武汉大学

2019/12/03

## 1 项目开发目的和意义

增强现实技术将计算机生成的虚拟信息如图片、文字、三维物体或声音等无缝地与真实环境进行叠加融合。近年来，随着科技的飞速发展，硬件的计算能力大幅度提升，这使得增强现实技术在娱乐、教育与医疗等领域具有巨大的需求量。增强现实技术尽可能让合成图像或视频具有较强的真实感，而真实感的体现是增强现实技术中非常具有挑战性的难题。增强现实合成图像的真实感主要体现在合成图像的几何一致性与光照一致性两个方面。其中，几何一致性要求插入的虚拟物体具有与背景图像相一致的几何透视关系，与现实环境相容的刚体位置与姿态等。光照一致性要求被插入物体能呈现出与背景真实环境相一致的明暗及其四周的阴影效果，否则被插入物体将表现出严重的悬空感。而被插入物体的阴影是合成图像光照一致性的最直接体现，与真实环境相一致的阴影效果能大幅度提升合成图像的真实感，如图1-1。



图1-1 虚拟物体的阴影提升真实感

通常，虚拟物体阴影生成需要采用逆渲染的方式，即需要估计整个场景的三维几何结构信息、反射率属性以及光照信息，但这些信息的估计难度很大，而且不精确的估计会造成最终生成的阴影具有奇怪的外观。本项目利用生成对抗网络（Generative Adversarial Network，GAN）的生成与对抗训练，以一种端到端（end-to-end）的方式直接为输入的合成图像生成虚拟物体的阴影，而不需要除了RGB图像以外的信息与显式的逆渲染过程，更加简单直观方便易用。本项目旨在华为Atlas200DK上实现虚拟物体的阴影生成，能够对输入的无虚拟物体阴影合成图像进行处理得到有虚拟物体阴影的合成图像，以此提升合成图像的真实感。虚拟物体阴影的推断与生成通过GAN完成，然而如果要实现这样的目标，相应的深度神经网络通常需要较大的参数量，可是较大的参数量往往会导致计算时间成本较高，计算延时太长将会违背增强现实应用所要求的实时性特性进而影响用户体验。因此，如何研究一个速度快、参数少、实用性强的阴影生成方案十分具有挑战性。

## 2 总体设计

本项目包含两个重要部分：数据集的构建与深度神经网络的训练。由于现在缺少相应的阴影生成数据集，我们需要自己收集构建相应的阴影生成数据集。由于深度神经网络的特点，其性能高度依赖于训练数据集，因此我们约定收集的数据需要具备一定的特征并符合相应的分布。

## 3 算法设计

整个项目涉及的算法包括数据集的创建方法与深度神经网络的设计与训练两部分。其中，数据的收集与数据集的制作是该项目的重点。

### 3.1 数据收集约束



图3-1 原始数据采集示意图

由于缺少相应的阴影生成数据集，本项目需要收集数据并处理以创建相应的数据集。拟使用Logitech C920 USB摄像头，以640x480的分辨率采集原始背景图像数据（已标定相机矩阵参数：cx=319.5，cy=239.5，f=680.0，单位：像素），如图3-1所示。本项目的核心思想是通过真实环境中的线索（即真实阴影与其投射物）来推断虚拟物体的阴影。因此，收集的原始背景图像数据必须能够提供足够的线索用于推断虚拟物体的阴影。除此之外，原始背景图像数据还需要满足严格的约束，具体分为如下三点：

1. 较为明显的真实阴影与其投射物必须同时存在，因为只有阴影的背景图像无法从中推断出虚拟物体的阴影。
2. 不得包含大面积的黑色或暗色非阴影区域，因为这类区域会干扰真实环境中有助于推断虚拟物体阴影的线索。
3. 暂时不考虑形状太复杂的阴影以及虚拟物体阴影与真实阴影互相作用的情况。

典型的正面案例与反面案例如图3-2所示。



图3-2从左至右：正面案例、缺少投射物的反面案例、大面积黑色非阴影区域的反面案例、复杂形状阴影的反面案例

### 3.2 标定与图像合成

我们所需要的数据集中，每条记录包括图像三元组：不含虚拟物体阴影的合成图像、虚拟物体的掩码（mask）、相应的包含虚拟物体阴影的合成图像，如图3-3。如前所述，产生这样的图像需要一套完整的逆渲染过程：相机内参数标定、相机位姿估计、场景几何反射率估计、光照估计。有了这些参数才能成功渲染出合理的虚拟物体阴影。

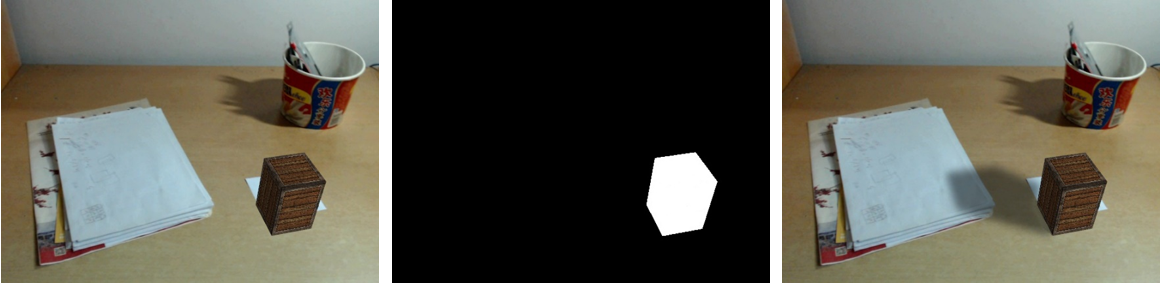
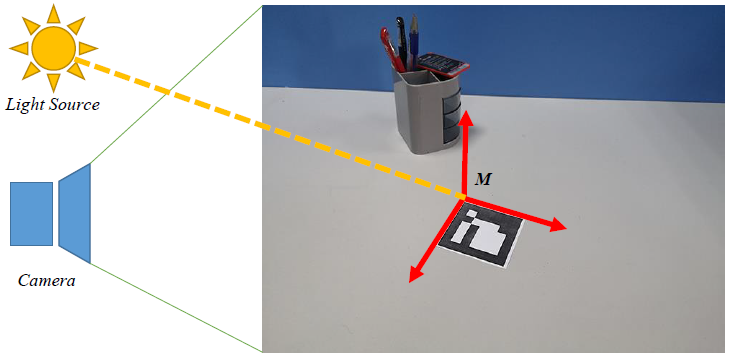


图3-3数据记录：无虚拟阴影图像、虚拟物体mask、有虚拟阴影图像

数据标定工作拟采用黑白方形标记物的识别辅助完成。以黑白方形标记物的几何中心为原点，以其边为坐标轴建立笛卡尔坐标系。通过识别黑白方形标记物，可以从图像中提取到标记物的4个角点，从而计算相机的位姿，完成从相机坐标系（简称C坐标系）到黑白方形标记物坐标系（简称M坐标系）的互相转换。

光源信息的标定也利用黑白方形标记完成。假定室内光源为最主要的单个点光源，室外光源（太阳光）为无限远的点光源，即方向光源，可在M坐标系下测量光源的位置与方向信息。针对室内环境，拟通过逐个关闭或遮挡找出产生最明显真实阴影的光源，并测量其几何中心的位置。针对室外环境，拟采用预先测定好的直尺，通过直尺的边缘角点与其影子相应角点之间的匹配计算太阳光的方向向量。虚拟物体的渲染在M坐标系内完成，且令虚拟物体的阴影投射在M坐标系所定义的平面上，并使用Phong光照模型绘制虚拟物体，用ShadowMapping方法绘制虚拟物体的阴影。如图3-4所示。

图3-4数据标定与图像合成

有了标定好的几何与光源信息，将不同的3D物体用于合成图像即可产生如图3-3所示的数据三元组。通过把物体颜色设置为纯白，背景设置为纯黑的方式可以渲染得到精确的物体mask。

### 3.3 黑白方形标记物识别

黑白方形标记物的识别是构建数据集的重要步骤，黑白标记的识别分为图像二值化、轮廓提取、角点检测、匹配识别这四大步骤。如图3-5所示。

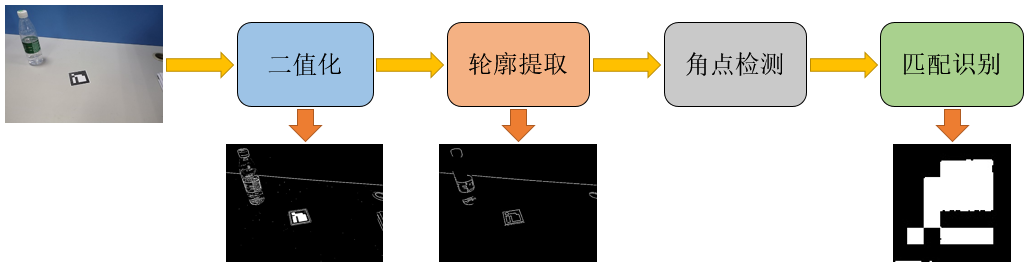


图3-5黑白方形标记物识别

图像二值化利用自适应阈值方法，将图片分为前景和背景两部分。本项目中，图像二值化的目的是为了将黑白方形标记的外观特征从RGB图像中提取出来，将背景统一变为纯色。轮廓提取将二值化图像中的所有的轮廓检测出来并放入列表，进而可以筛选出组成四边形的直线，实现角点的检测与提取。通过四个角点，可将提取得到的四边形用透视变换（Perspective Transformation）恢复为正方形，经过0/1编码后与黑白方形标记物的编码进行匹配。若两者编码的Hamming距离满足一定的阈值条件，则说明从图像中成功识别出了一个黑白方形标记物。黑白方形标记物的0/1编码为一个7x7的矩阵，如图3-6所示。

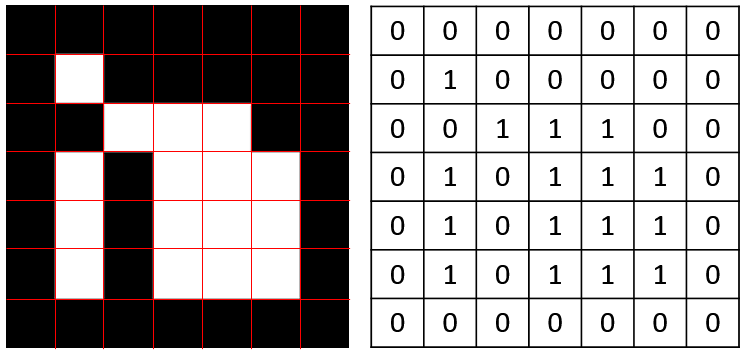


图3-6黑白方形标记物编码

黑白方形标记物识别的所有步骤均通过计算机视觉库OpenCV所提供的开源算法实现，满足本项目的需求。

### 3.4 虚拟物体的阴影生成

虚拟物体的阴影生成通过生成对抗网络完成。GAN在构建好的训练集上进行对抗训练，学习虚拟物体阴影与真实环境之间的映射关系。GAN网络包含一个生成器与一个判别器。给定输入图像，生成器负责生成虚拟物体的阴影，判别器负责判别包含虚拟物体阴影的合成图像是否为“真”。在每一轮训练过程中，生成器与判别器互相竞争、彼此提升能力，最终达到一种均衡。如图3-7所示，虚拟物体阴影生成采用冗余式结构（residual）表示虚拟物体的阴影，生成器的直接输出与输入原图相加得到最终的包含虚拟阴影图像。

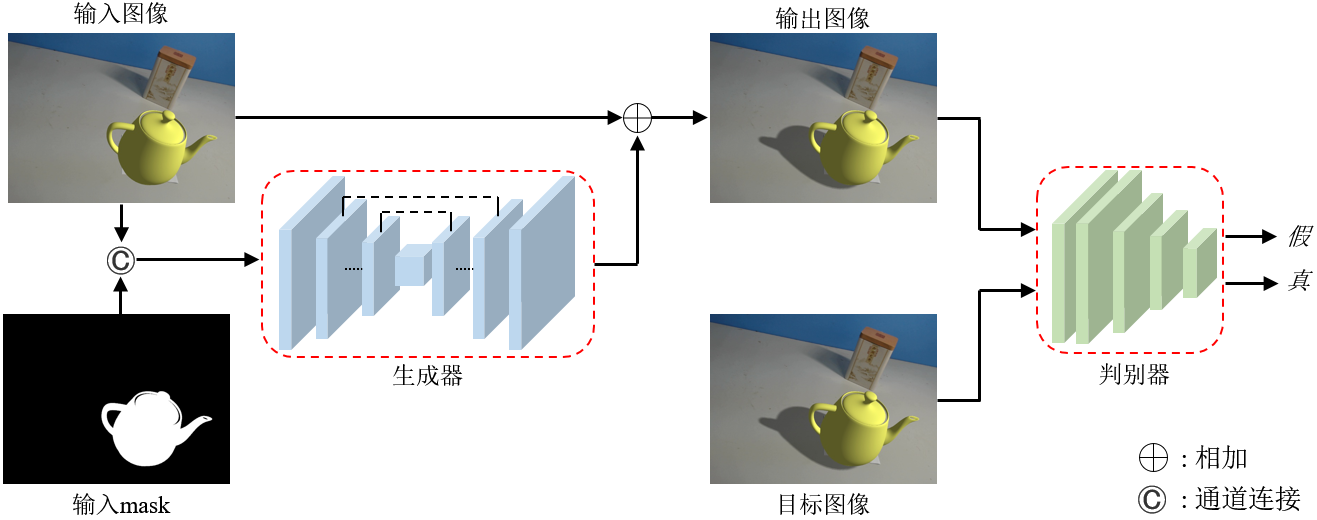


图3-7生成对抗网络总体结构

**3.4.1 生成器结构**

GAN网络的生成器采用典型的U-Net结构，有5个相互对称的降采样与上采样模块组成。降采样模块通过卷积操作提取特征并将特征图的尺寸减半，上采样模块操作将特征图的尺寸逐渐加倍，最终恢复成初始大小。降采样模块输出的特征图被直接连接到对称的上采样模块的输入中，通过这种短连接（Skip Connection）方式可防止降采样过程中的特征消失。生成器输入无虚拟阴影图像与虚拟物体mask，最终输出有虚拟阴影图像。

**3.4.2 判别器结构**

GAN网络的判别器采用Patch-GAN的形式。Patch-GAN是一种判别器，这种判别器由Phillip Isola等人2017年于文献“Image-to-Image Translation with Conditional Adversarial Nets”中提出。这种判别器将一个特征图块直接输出，而非一般的全连接操作输出一个值，这种方式能增大判别器的感受野（Receptive Field）从而使其兼顾图像的局部与全局信息。最终，由全局平均池化操作将特征图块转化为图像为“真”的概率。