

基于图像条件的二元合成生成算法

张婷

(安徽粮食工程职业学院信息技术系,安徽 合肥 230011)

摘要:生成对抗网络(GAN)是一种无监督学习方法,该算法巧妙地利用博弈的思想来学习生成式模型,但由于GAN通常从单个潜在源采样,因此常常丢失场景中的多个实体交互信息。为了捕获不同对象之间的复杂交互,包括它们的相对缩放,空间布局,遮挡,提出了一种基于图像条件的生成对抗网络,利用“分解—合成”的流程,模型可以根据输入对象的纹理和形状从它们的关节分布生成逼真的合成图像。通过使用Shapenet数据集,在2D和3D图像中分别对55个常见对象类别约51 300个图像模型进行试验,比起传统的SLP和cGAN,算法的图片质量有4%以上的提高。

关键词:生成对抗网络;图像合成;图像纹理;交互信息

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1673-1891(2020)02-0069-04

Binary Synthesis Generation Algorithm Based on Image Conditions

ZHANG Ting

(Department of Information Technology, Anhui Vocational College of Food Engineering,
Hefei, Anhui 230011, China)

Abstract: The Generative Adversarial Network (GAN) can generate images with significant complexity and authenticity, but it is usually constructed to sample from a single potential source, thus ignoring spatial interactions between multiple entities that may exist in the scene. To capture the complex interactions between different objects, including their relative scaling, spatial layout, occlusion or view conversion, in this paper we propose an image condition-based generative adversarial network by using decomposition-synthesis procedures. The model can generate realistic composite images from their joint distributions based on their texture and shape of the input object. Through use of Shapenet data set, we experiment on 51 300 image models out of the respective 55 common objects in the 2D and 3D images. Compared with the traditional SLP and cGAN, the image quality in our algorithm can be improved by 4%.

Keywords: Generative Adversarial Network; image synthesis; image texture; interactive information

0 引言

图像合成技术是计算机视觉的一个重要分支,在影视动画,平面设计领域都有许多应用。国内外很早就展开了关于图像合成技术的研究,吴昊等^[1]综合分析了主流的图像合成技术,并分别在鲁棒性、运算效率方面进行了比较。朱春晓等^[2]采用了将色度信息和亮度信息都分别处理再利用混合融合的方式实现了图像合成。蔡伯林等^[3]基于视觉成像技术和复眼成像原理,利用多摄像机多角度进行光晕校正,实现了图像的合成和标定。王攀等^[4]基于Sort-last并行绘制,充分利用动态规划算法减少了合成相关的开销。Badki等^[5]提出控制构图的2个最重要的参数是相机位置和镜头焦距,对此引入了

计算缩放,并进一步定义了一种多视角相机模型,该模型可以生成物理上无法达到的构图,从而扩展了对诸如不同深度处的对象的相对大小和图片深度感等因素的控制。但是上述基于传统算法的图像合成技术,都面临合成图像的失真或者高昂硬件成本的挑战。如何在降低计算成本的基础上完成提升图片合成的质量,是学术界和工业界是十分关注的问题。

条件生成对抗网络(cGAN, Conditional Generative Adversarial Network)是生成对抗网络的升级,它能根据给定图像,伪造生成逼真的图像,除了图像,cGAN还可以对文本短语或类标签布局输出结果^[6]。这类cGAN的学习目标是利用学习一个映射,该映射从图像源分布转换到给定样本,再输

收稿日期:2020-01-07

作者简介:张婷(1982—),女,安徽淮北人,讲师,学士,研究方向:软件开发、数据库。

出生成结果。在此基础上,cGAN还可以根据需求改变输入图像的样式和纹理^[7]。但是cGAN更多是生成单个图像,而自然界中的图像大多是3D视觉世界中交互的多个对象组合的2D投影,而传统的cGANs并不能捕获这类信息^[8],为了让进一步利用图像信息来生成更加逼真的组合图像,本文探索组合算法在cGAN框架中的作用,并提出一种新的算法,该算法通过训练函数映射,从边缘分布中采样出不同对象的图像到组合样本中,从而捕获对象的联合分布。本文重点研究两两组合的一对物体的合成。由于不同对象之间在相对缩放、空间布局、遮挡或视点变换方面的复杂交互,如何自然和谐地对2个图像进行组合具有较大难度。Senaras等^[7]在GAN框架中使用空间变换器网络,从而通过在几何扭曲参数空间中操作前景对象来分解该问题。但是这种方法仅限于固定的背景,并不考虑现实世界中更复杂的交互。本文考虑将2个输入对象图像组合成一个能捕捉真实交互的联合图像的工作。

1 算法架构

1.1 基于条件的GAN

给定随机噪声向量 Z ,GAN使用生成器 G 生成特定分布的图像 C ,生成器 G 相对于判别器 D 进行博弈训练。当生成器试图产生逼真图像时,判别器通过学习区分真实对象来对抗生成器从而识别出生成图像的真假,最终生成器和判别器经过博弈达到纳什均衡。在条件GAN模型(cGAN)中,辅助信息 x 以图像或标签的形式与噪声向量一起被输入到模型中。因此,cGAN的目标是将对抗性损失函数表示为式(1)。

$$L_{cGAN}(G, D) = E_{x,c}[\log D(x,c)] + E_{x,z}[1 - \log D(x,c)] \quad (1)$$

其中, D 和 G 分别是判别器和生成器。如果增加损失惩罚所生成图像与其基础事实的偏差,则上述GAN目标的收敛以及因此产生的图像质量将得到改善。因此,生成器的目标函数将总结为式(2)。

$$G^* = \arg \min(G) \max(D) L_{cGAN}(G, D) + \lambda E_{x,c,z}[\|c - G(x,z)\|] \quad (2)$$

cGAN已经广泛应用于多个图像转换场景,例如白天到夜晚,马到斑马,以及草图到描绘。组合问题比仅将图像从一个场景转换到另一个场景更具挑战性,因为模型还需要处理各个对象的相对缩放,空间布局,遮挡和视点变换以生成逼真的合成图像。这里,本文提出了用于在给定2个单独的对象图像的情况下生成合成图像的合成GAN。设 x 是包含第1个对象的图像, y 是第2个对象的图像, c 是来自它们的组合图像。在训练期间,本文从2个

对象的边缘分布给出数据集 $X=\{x_1, \dots, x_n\}$ 和 $Y=\{y_1, \dots, y_n\}$,并且 $C=\{c_1, \dots, c_n\}$ 来自包含2个对象的联合分布。本文利用组合GAN模型以2个输入图像 (x,y) 为条件以生成图像,其目标分布为 $p_{data}(c)$ 。算法的目标是确保生成的合成图像 c 包含具有相同颜色,纹理和结构的 x,y 中的对象,同时对于集合 C 也看起来是真实的。由于条件GAN不足以在空间上转换对象,本文通过基于相对空间变换网络(STN)将对象图像 (x,y) 转换为 (x^r,y^r) 来明确地建模比例和移位变换^[9-10]。

1.2 算法原理

本文研究将2个输入对象图像组合成1个能捕捉真实交互的组合图像。例如,给定椅子和桌子的图像,理想的算法能够生成排列的椅子-桌子图像。本文采用分解算法作为监督函数来训练组合算法,从而通过组合分解(CoDe)的网络实现自我一致性约束。此外,本文在测试时使用这种CoDe网络进行特定于示例元细化(ESMR)方法,通过每个给定的测试示例微调组合网络的权重,以生成更清晰、更准确地合成图像。分解网络可以提供自监督,并通过定性和定量实验,在训练场景中评估本文提出的Compositional-GAN方法。

1.3 通过分解算法监督合成算法

本文算法的核心思想是通过一致性损失函数监督2个图像 (x^r,y^r) 的组成元素,从而确保生成的合成图像 c 可以进一步分解回相应的个体物体图像。使用条件GAN(合成判别器和合成生成器)完成图像合成,将2个RGB图像作为输入串联连接以生成相应的合成输出 c ,2个输入图像合理地组合并生成图像,然后将该生成的图像输入到另一个条件GAN(分解判别器和分解生成器),以使用损失函数将其分解回其组成对象。本文实现的分解合成(CoDe)网络的示意图如图1所示。

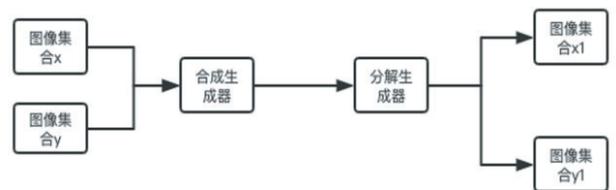


图1 分解合成网络

除了分解网络之外,生成的合成图像将提供给预测网络来预测合成图像中的每个像素属于每个输入对象或背景的概率,同时本文设定2个分解生成器在其编码器网络中共享权重。

1.4 ESMR合成图像精细化

组合GAN模型不仅可以组合2个对象,同时保

留合成图像中各个对象的颜色纹理和位置等其他信息。本文的训练算法充分利用组合GAN的优势,通过将合成图像分解回单个对象来进行监督,使用相同的监督函数来细化生成的合成图像以用于测试结果。在此基础上本文继续使用生成的图像分解成2个测试对象来优化网络参数,以去除伪像并生成更清晰的结果。该算法也被称为基于示例的元细化(ESMR)算法,如图2所示。

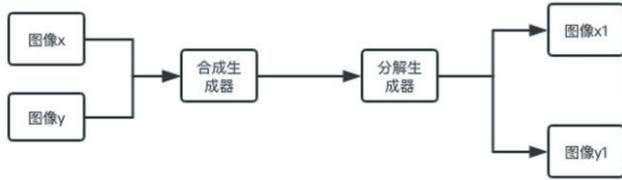


图2 ESMR细化算法架构图

本文ESMR方法的目标函数如式(3):

$$L(G) = \lambda(\|x^T - x^T\|_1) + L(G, D) + \|Y^T - Y^T\|_1 \quad (3)$$

其中, X^T 是生成的分解图像, X^T 是输入的转置,从式(3)中可以看到分解和掩模预测网络在生成更清晰的输出和预测方面相互补充,互相加强。

1.5 完整训练模型

根据上述分析,本文的训练模型完整流程图如图3。

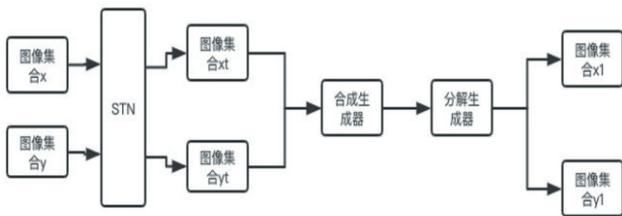


图3 完整框架流程

1) 在合生成器,分解生成器和相对STN模型的输出上进行像素重建L1损失函数:

$$D_{L1}(G_c) = E(x, y, c)[\|c - \hat{c}\|_1] \quad (4)$$

$$D_{L1}(G_{dec}) = E(x, y)[\|(x^T, y^T) - G_{dec}(c)\|_1] \quad (5)$$

$$D_{L1}(STN) = E(x, y)[\|(x^c, y^c) - (x^T, y^T)\|_1] \quad (6)$$

其中, $(x^T, y^T) = STN(x, y)$, $c = G_c(x^T, y^T)$ 。此外, (x^c, y^c) 是在合成场景中对应于合成图像c的转置输入对象。此外, $D_{L1}(G_{dec})$ 表示自适应约束惩罚分解图像与输入图像的偏差。

2) 计算交叉熵掩模预测损失 $L_{CE}(G_{dec}^M)$, 到为生成的合成的每个像素分配标签图像。

3) 组合和分解网络的条件GAN损失函数为:

$$L_{cGAN}(G_c, D_c) = E_{(x,y,c)}[\log D_c(x^c, y^c, c) + E(x, y)[1 - \log D_c(x^c, y^c, c)]] \quad (7)$$

$$L_{cGAN}(G_{dec}, D_{dec}) = E_{(x,y)}[\log D_{dec}(x^c, y^c, \hat{c}) + \log D_{dec}(x^c, y^c, \hat{c})] \quad (8)$$

2 试验和结果分析

本文使用Shapenet数据集^[3],这个数据集包含2D和3D图像,涵盖了55个常见对象类别,约51300个图像模型。本文重点研究2个组合任务:(1)桌子旁边的椅子。(2)电脑旁边的瓶子。在合成过程中主要处理以下4个问题:空间布局,相对缩放,遮挡。在该试验中,主要目标是预测正确的遮挡像素以及2对象的相对缩放。合成桌子和椅子:本文手工制作了一张来自Shapenet椅子和桌子的合成图像的集合,用作配对和非配对训练方案中的真实关联组合C。输入输出集中的桌子和椅子可以在[-180°, 180°]的范围内以10°为步长构成随机方位角。给定随机视角和输入中任意椅子,本文算法在与桌子一致的视角中合成椅子。合成电脑和瓶子:利用100张数据进行训练,再分别准备90个和80个联合实例的测试集,主要的对比算法有基于进化的协方差算法^[10],传统的cGAN没有加入EMSR的cGAN。训练集如图5所示。



图5 测试数据集

对于试验结果,本文随机征集了100位不同年龄段的视力良好的人群对本文生成的合成图像进行评估,让他们判断合成的图片是否有违和感,以比较本文算法在不同情况下的性能。表1总结了评估结果,可以看到,采用了本文算法cGAN合成图片的质量明显高于其他几种算法,对于次优的无EMSR的cGAN,本文算法的逼真程度也有4%以上的提高。

表1 评估结果对比

输入	测试数 数据量	进化协方 差算法/%	传统 cGAN/%	无EMSR的 cGAN/%	本文 算法
椅子和桌子	90	71.22	67.22	72.38	76.39
瓶子和电脑	80	71.38	61.34	65.29	76.22

3 结语

本文提出了一种新的组合条件生成对抗网络模型用来解决条件图像生成中的物体组成问题,该算法利用分解算法对合成结果进行监督,充分考虑到输入对象的相对仿射和视点变换(除了像素遮挡

或生成),从而生成逼真的合成图像。通过定性试验合成和实际数据集的合成结果进行了判断,证明了本文算法能综合考虑2个物体合成时候的相对交互,合成的图片更加自然和谐,本文的算法对图像合成领域具有一定的参考价值。

参考文献:

- [1] 吴昊,徐丹.数字图像合成技术综述[J].中国图象图形学报,2018,17(11):1333-1346.
- [2] 刘春晓,朱建锋.无颜色失真的无缝图像合成算法[C]//中国工业与应用数学学会几何设计与计算专业委员会.第五届全国几何设计与计算学术会议论文集.2011.
- [3] 蔡柏林,吴家俊,张磊,等.多视角图像标定与图像合成[J].应用光学,2019,40(1):15-21.
- [4] 王攀,杨平利,黄少华.小通信开销的Direct Send并行图像合成方法[J].计算机研究与发展,2018,55(4):215-222.
- [5] BADKI A, GALLO O, KAUTZ J, et al. Computational zoom: a framework for post-capture image composition[J].Acm Transactions on Graphics,2017,36(4):1-14.
- [6] HENSMAN P, AIZAWA K. cGAN-based manga colorization using a single training image[EB/OL].(2017-06-21)[2019-11-01].<https://arxiv.org/pdf/1706.06918.pdf>.
- [7] SENARAS C, NIAZI M K K, SAHINER B, et al. Optimized generation of high-resolution phantom images using cGAN: Application to quantification of Ki67 breast cancer images[J].Plos One,2018,13(5):e0196846.
- [8] RAMPONI G, PROTOPAPAS P, BRAMBILLA M, et al. T-CGAN: Conditional generative adversarial network for data augmentation in noisy time series with irregular sampling[EB/OL]. (2018-12-19)[2019-11-15]. <http://www.doc88.com/p-6931788330527.html>.
- [9] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[EB/OL]. (2016-02-06)[2019-12-01]. <https://arxiv.org/pdf/1506.02025.pdf>.
- [10] SHEN H L, ZHANG H G, LI Z N. Detection of STN-LCD defects based on image registration[J]. Opto-Electronic Engineering,2008,35(9):60-65.

(责任编辑:蒋召雪)

(上接第49页)

参考文献:

- [1] 《求是》评论员.勇于担当作为 奋力开创宣传思想工作新局面[EB/OL].(2018-09-03)[2020-02-06].<http://theory.people.com.cn/n1/2018/0903/c40531-30268888.html>.
- [2] 郭艳,李长福.地方特色文化语境下皖北地区城市景观设计探索[J].皖西学院学报,2019,35(3):88-93.
- [3] 胡嫔.论楚汉文化艺术品质的承转与出新[J].求索,2012(10):251-252+233.
- [4] 李怡然,房鑫.安徽江淮地区汉墓文化因素分析[J].文物鉴定与鉴赏,2016(6):92-95.
- [5] 陈琳,陈丽丽.淮河文化的成因与特色[J].江苏地方志,2007(1):43-46.
- [6] 杨瑞峰,费晓萍.楚文化中凤纹样演变及其创新核心探索[J].美术教育研究,2018(23):28-29.
- [7] 高介华.“楚辞”中透射出的建筑艺术光辉—文学“幻想”,楚乡土建筑艺术的全息折射[J].华中建筑,1998(2):30-40.
- [8] 周天游.论汉代文化的基本特征[J].社会科学战线,2007(2):253-255.
- [9] 周玉佳,吴伟东,马之路.淮南地区楚风汉韵建筑设计研究与传承[J].淮阴工学院学报,2018,27(3):57-62.
- [10] 李泽厚.美的历程[M],天津:天津社会科学院出版社.2001:116
- [11] 李敏,王逸玮,王祖成.汉代楼阁建筑形式初探[J].西安建筑科技大学学报(自然科学版),2017,49(4):503-507.

(责任编辑:曲继鹏)