

硕 士 研 究 生 读 书 报 告



题目 TileGAN大规模非均匀纹理生成

作者姓名 王瀚林

作者学号 21951127

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 二○一九 年 十二 月

TileGAN: Synthesis of

Large-Scale Non-Homogeneous Textures

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Qilei Li

By

Hanlin Wang

Zhejiang University, P.R. China

2019

摘要

本文主要探讨TileGAN论文中的问题、解决方案、创新点以及研究成果的应用。内容包含对大规模非均匀纹理生成问题的描述，以及所涉及研究领域的研究现状。论文中主要提出一种基于图块组合的算法，从而生成极小边界瑕疵的大规模纹理。其次，论文提出一种用户界面使模型能够用于艺术加工。文中所述模型能够在数亿像素级别的纹理生成中起到出色的效果。

**关键词**：计算机图形学，纹理合成，生成式对抗网络

Abstract

This article mainly discusses the problems, solutions, innovations, and application of research results in TileGAN papers. The content includes the description of “synthesis of large-scale non-homogeneous textures”, as well as the research status of the related research fields. This paper proposes an algorithm based on tile combination aim to generate a large-scale plausible texture map with virtually no boundary artifacts. Secondly, the paper proposes propose a user interface to enable artistic control. The model in this paper can achieve excellent results in the texture generation of hundreds of megapixel.

**Keywords：**computer graphics, texture synthesis, generative adversarial network

**1 研究问题**

纹理合成是计算机图形学等方向的研究热点之一，基实例的纹理合成指的是对于给定的输入实例通过特定的算法生成与给定实例相似的纹理的任务。其中输入纹理的视觉特征应在生成纹理中得到再现，同时保留实例的局部与全局特征。此次阅读的TileGAN[1]一文中涉及的便是大规模的纹理合成问题。

使用传统方法往往在面对大规模的纹理合成问题时效果不佳，在没有大量参考数据时传统方法的输出会缺乏可变性，这样在大规模纹理合成时产生的结果会有重复性过高与过于均匀的问题。因此作者采用了对抗式生成网络(GAN, Generative Adversarial Network )来进行纹理的生成。现有的生成式对抗网络已经能够在纹理合成上取得很好的效果，但是现有网络的输出规模都难以直接用于大规模的纹理合成。

在将生成式对抗网络用于大规模的非均匀纹理生成的过程中，假设使用的生成式对抗网络只能生成有限分辨率的图片，那么便会产生论文中的核心问题，即如何在不产生边界瑕疵保留局部特征与全局特征的情况下，做到近乎不限规模的纹理合成。除此之外，由于目前大多数基于生成式对抗网络的纹理合成方法极少提供艺术控制，非专业人员难以将纹理合成网络用于实际用途，论文还对这一问题给出了解决方案。

**2 研究过程**

作者的方案主要分为生成模型、合成与艺术控制三个部分。其中作者主要的工作集中于其中的合成部分，这一部分运用到生成模型部分得到的模型，并设计了便于引入艺术控制的纹理合成算法。

**2.1生成模型部分**

生成模型部分作为框架中生成新图像的环节，是整个框架的基础。GAN通常包括生成器与判别器两个网络，生成器通过卷积层进行参数训练并生成图像，一个随机的潜向量经过这些卷积层不断提升分辨率最终生成全尺寸的图像。判别器网络用于评估生成的样本与训练集的匹配程度。两个网络都通过梯度计算来进行训练。

在此论文中作者初期已经尝试了使用生成式对抗网络进行纹理生成，作者在使用Karras等人的框架[2]时得到了很好的效果，Karras的论文中主要针对GAN训练不稳定与训练速度慢的问题提出了一种解决方法，在训练时先从低分辨率的图像开始训练，之后逐步增加图像的分辨率与网络层数，通过低分辨率图像的训练使网络找到大致的梯度方向，在通过增加分辨率与网络层数逐步进行细化。

但对于大规模的纹理合成问题来说该方法输出的纹理分辨率难以达到要求，因此作者在此模型的基础上进行研究，探索生成式对抗网络的应用方法，将其用于大规模的纹理生成。

**2.2合成部分**

合成部分的算法是该论文的核心内容。论文利用预训练的生成网络来合成高质量的大规模非均匀纹理。这一过程通过生成的中间潜向量构成的图块来完成，这些图块在论文中被称为潜在场(latent field) F。如图中所示，生成器通过F来生成连贯的大规模纹理。在构建潜在场F时，为了使生成的纹理在整体上具有一致性，即保留其全局特征，需要用户随机生成或指定引导图像M来作为构建前在场F的依据。除此之外，为了防止相邻纹理间突然的纹理变化，在选择潜在场F组成部分的候选时考虑到了相邻纹理块之间兼容性。

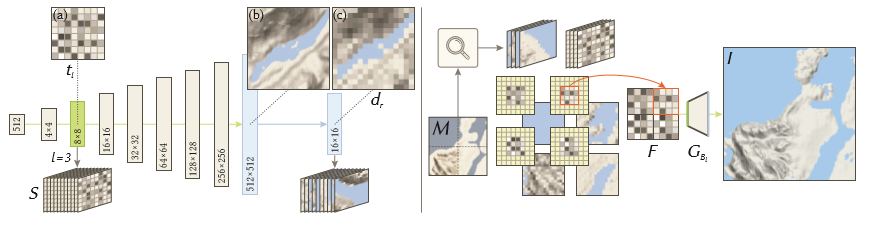


图1 合成总览

图1中左侧是生成器的部分。以大小为512的随机潜向量作为输入，在的三层潜在图块的分辨率为8×8，然后对大量的图块进行采样，对于每个潜在图块都生成相应的输出图像，并在生成器的尾部添加了降采样步骤，以此作为邻域相似度匹配的依据。右侧是合成的部分，根据降采样的结果与引导图像M中对应区域的相似性寻找图块，图块经过剪裁并组成潜在场F。剪裁步骤是为了调整结果的大小与混合程度。最终F经过生成器生成最终图像。

艺术控制部分主要是方便用户使用这一纹理合成方法，提供结果控制的核心途径就是修改潜在场F。以此为依据论文提供了绘画、复制、打乱和目标匹配等多样的操作。

本文的核心算法是合成部分的算法。过程中需要用到一个标准的DCGAN(Deep Convolution Generative Adversarial Networks)，如下：

(1)

其中z是随机的潜向量，l是具体实施潜在场合成方法的层，GAl与GBl是生成器G在第l层进行分割而成的两部分，其中生成器GAl输入z，输出在第l层的k×k图块，生成器GBl以此作为输入，并在最后一层输出最终的图像。

整体合成的步骤分为三步：仅执行一次的预处理过程、在线潜在场合成过程、在线纹理合成过程。

**2.2.1预处理过程**

预处理过程的第一步是使用标准的生成器网络生成一个巨大的纹理集S，集合中的每个样本包含两个部分：在第l层的中间张量tl与最终生成图像的降采样结果dr。第二步根据dr展现出的图像特征对S中的样本进行k-means聚类，得到各个类别的中心点ck。在之后的步骤中可以通过聚类结果来进行相似图块的快速搜索。

**2.2.2潜在场合成过程**

潜在场合成过程，论文称其是第一个将马卡洛夫随机场(MRF, Markov Random Fields)应用于GAN中潜向量的研究。考虑到后期需要引入交互编辑，论文中将潜在场的合成算法分割为潜在场的初始化与迭代优化两步。在使用中先呈现第一步的生成结果，再在后台进行优化步骤并刷新输出。

初始化步骤程对潜在场F进行初始化，对F中划分的每个部分找到S中最合适的图块进行填充，最终得到一个完整的潜在场F。其中判断图块的匹配度的依据是一个表示与引导图对应部分M匹配程度的一元能量项。

迭代优化过程会一直执行，直到潜在场的能量E低于人工设定的阈值。阈值确保在优化过程中保留足够多的视觉特征。在每次优化中，我们随机取F中的一个图块，然后寻找能够让能量最小的候选图块，目标是找到使潜在场的加权一元项能量Em与二元项能量En之和E最小。

Em指的是候选图块Fi对应的dr与对应区域M的视觉相似度的总和。在论文中使用欧氏距离来衡量视觉相似度。En指的是4个相邻的潜在图块按如下公式进行计算得到的加权相异项。

(2)

其中包含三个主要项,DV通过计算相邻图块重叠区域对应dr之间的欧氏距离代表视觉相异性，DL计算相邻图块重叠区域的张量tl的欧氏距离得到潜在张量之间的相异性，DC用来表示相邻块的在预处理中的聚类结果是否匹配，匹配则为0，反之为1。具体的衡量过程是先用Em筛选出10个最相似的候选，再用En找出符合要求的结果。

可以看出，这种衡量方式首先考虑了潜在场中的每个图块与引导图像M中对应区域的相似性，之后也考虑了每个图块与其相邻区域之间的相似性。

在组合潜在图块时，仅让相邻图块足够相似还是会在图块连接处出现不连贯的现象，为了保障块之间的连贯性，文中在组合潜在场F时将候选图块剪裁至2×2或4×4的大小，然后在候选图块的搜索与优化时仍使用完整的对应视觉信息dr与引导图像M进行相似度的比对，提升了整体图像的连贯性。

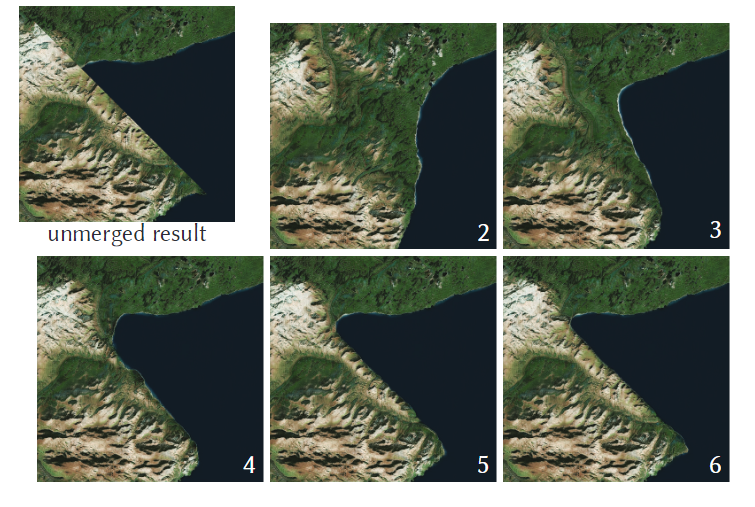


图2 在不同层进行潜在场合并的效果

图中为对分割生成器的潜在层l的选择产生的不同的效果差异。

**2.2.3纹理合成过程**

最后纹理合成过程，这一步骤将潜在场F输入生成器GBl产生最终的完整图像，最终生成大规模的图像。论文中采用l=3, n=9的过程，即使用总共9层的生成器，在第3层进行潜在场的合成。由于使用的是卷积网络，这一方法可以有效地用于局部纹理更新，并且便于大规模纹理的快速生成。

总体来说，论文的重点在于将生成器分割为两部分，并在分割层进行潜在张量的拼接匹配，生成器中的这一过程用到的基于马尔可夫随机场的算法是整个合成过程的核心。对抗式生成网络的结构并非重点，文中对对抗生成网络的训练过程进行了部分省略。

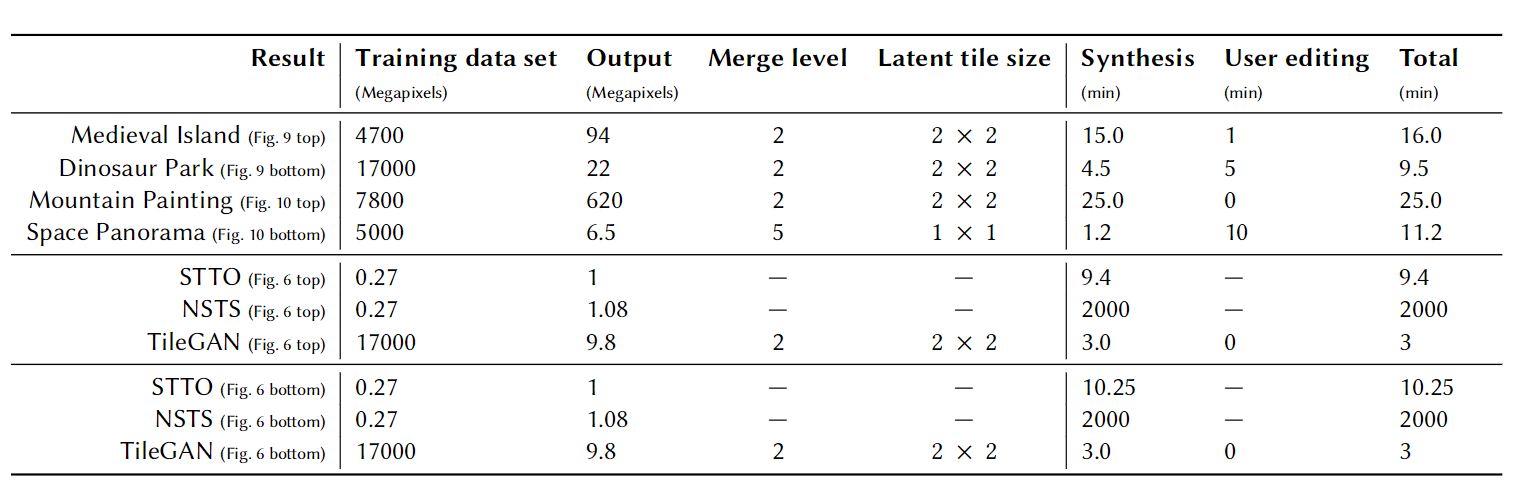
**2.3艺术控制部分**

在艺术控制方面，得益于网络的卷积结构与潜在场合成算法的步骤划分，艺术控制界面能够有较好的用户体验，在合成与局部编辑的速度上的表现都十分优秀。用户的编辑功能主要通过修改潜在场实现，论文中给出的编辑方法的核心有两种,一种是直接对潜在场中的图块进行修改，用户可以从图块库中选择替换图块。第二种是对引导图像的修改，用户将图像作为笔刷，替换绘制区域的引导图像，重新在局部进行合成算法，最终的绘制结果与周边风格相似，形成良好过渡。

**2.4 最终效果**

论文的最后，作者将TileGAN方法与当前先进的纹理生成算法进行了对比。

表1 TileGAN方法与SOTA方法的数值对比



表中展示了TileGAN方法与当前最先的进纹理合成方法数值上的对比，其中包含代表非神经网络算法的自调节纹理优化(STTO, self-tuning texture optimization)[4]与代表神经网络算法的非平稳纹理合成(NSTS, non-stationary texture synthesis)[5],对比结果可以展示出TileGAN方法在生成大规模的非均匀纹理时的时间优势，有效缩短了纹理合成与用户编辑操作的耗时，并且与其余两种方法不同，包含用户编辑功能的运算耗时。

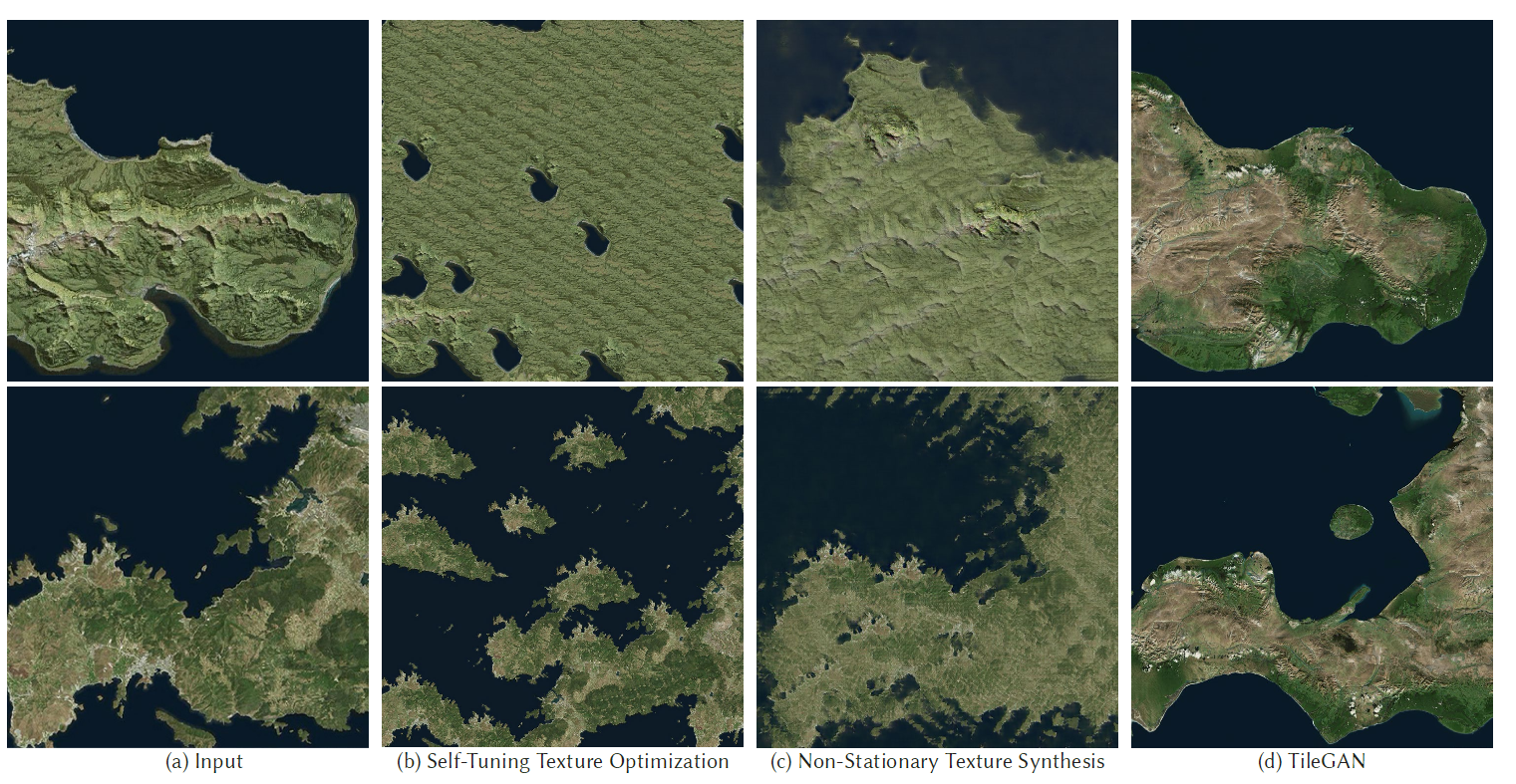


图3 TileGAN方法与SOTA方法的合成结果对比

图3中展示的图片与表1中所展示的方法相对应，展示了在相同输入的条件下不同方法在生成纹理上的视觉效果。可以看到论文中所用方法相较其余方法具有明显的优势，在生成结果上质量很高，且仅有少量瑕疵。

**3研究创新点**

研究的创新点主要分为以下几个方面：

第一点是作者将生成式对抗网络运用在了大规模纹理合成问题上，并解决了其中的关键问题。根据作者对于研究背景的描述，在纹理合成领域主要的分离合成方法主要分为非参数纹理合成与参数纹理合成，有部分研究将生成式对抗网络用于解决纹理合成问题，但仍然存在缺乏可扩展性与可变性方面的问题。与作者所用方法最接近的案例使用对抗生成网络用于基于实例的地形编辑[3]。

第二点在算法的创新上，论文中作者面对大规模输出的问题，引入马尔可夫随机场，提出了一种更加灵活的拼接算法。通过此算法进行潜在张量的匹配与更新，使得该纹理生成算法可以用于生成近乎不限规模的图像，而又保持整体与局部上的纹理特征，在同类算法中具有较大优势。

第三点，作者在提出纹理合成算法的同时提供了类似于笔刷的艺术控制界面，提供艺术控制在同类论文中很少出现，并且在作者的演示中拥有较好的效果，算法的分步骤设计使得在实际的合成与编辑中相较其余先进方法有较快的速度。艺术控制方法的引入使得这一纹理合成算法更容易应用于实际场景。

论文中提到，当前的方法还存在一定的缺陷。首先作者测试得到的最大输出约为160亿像素，在常规的艺术加工场景下难以承受更大输出的内存消耗，需要对合成与查看过程需要对内存管理过程进行设计与优化。

如果因为某种原因选择了与临近图块差异较大的图块来构成潜在场，或遇到比较难的引导图像时，输出结果质量会有所下降。有时会产生较为明显的块型结构，这一问题在较大的潜在图块上更为显著，但是可以通过调整生成器的分割层与调整潜在图块的大小进行改善。极少情况生成的纹理中会存在瑕疵，并且在周边潜在图块质量不佳时更加明显。

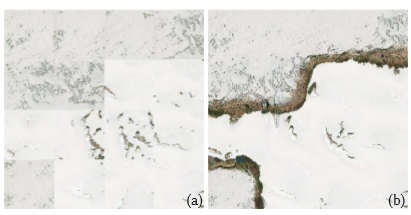


图4 产生不合理纹理的情况

在使用视觉上相似但实际不同的图块进行重复拼接时有可能会产生不同颜色的可见边界，论文中举例为北极冰与陆地冰，在模型进行纹理生成时会在两种图块位置的中间添加绿地。作者推测这一问题产生的原因可能是在训练集中缺少这两种图块的过渡数据导致的。

**4 成果应用**

论文中的算法主要用于大规模的纹理生成，可以用于高分辨率的艺术创作以及高精度大规模地图纹理的生成。算法对艺术控制的添加十分友好，可以进一步扩展作为更加专业的图像处理工具来使用。同样可以用于像游戏、影视等规模较大且需要艺术控制功能的场景。

论文中具体的通过分割生成式对抗网络的生成器，在分割层进行潜在张量操作算法也可以进一步用于其他的研究与应用领域。

参考文献

1. Frühstück, Anna, Alhashim I , Wonka P . TileGAN: Synthesis of Large-Scale Non-Homogeneous Textures[J]. ACM Transactions on Graphics, 2019.
2. Karras T , Aila T , Laine S , et al. Progressive Growing of GANs for Improved Quality, Stability, and Variation[J]. 2017.
3. Guérin, éric, Digne J , Galin, éric, et al. Interactive example-based terrain authoring with conditional generative adversarial networks[J]. ACM Transactions on Graphics, 2017, 36(6):1-13.
4. Kaspar A, Neubert B, Lischinski D, et al. Self tuning texture optimization[C]//Computer Graphics Forum. 2015, 34(2): 349-359.
5. Yang Z , Zhen Z , Xiang B , et al. Non-stationary texture synthesis by adversarial expansion[J]. ACM Transactions on Graphics, 2018, 37(4):1-13.