

硕 士 研 究 生 读 书 报 告



题目 基于图割技术的图像和视频纹理合成

作者姓名 毛琦

作者学号 nb15032

指导教师 李启雷

学科专业 移动互联网与游戏开发技术

所在学院 软件学院

提交日期 二○ 一五 年 十二 月

Graphcut Textures: Image and Video Synthesis Using Graph Cuts

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

Qi Mao

Zhejiang University, P.R. China

2015

摘要

本文引入一个新的算法进行图像和视频纹理合成，该方法将来自样本图像（视频）的像素快转化、复制到输出图像，然后沿着最优接缝缝合在一起，以产生一个新的（通常更大）的输出图像（视频）。与其他技术相比，本文的像素块尺寸不是先验的，而是基于输入和输出的纹理偏差，利用图形切割技术（Graph cut）确定最优像素快区域。不同于动态规划，图割技术适用于任何纬度的接缝优化。除了常规的2D图像合成，我们还在进行了3D视频纹理合成。实验显示，近似偏差搜索技术（approximative offset search techniques）能与像素块尺寸优化技术很好的结合起来。对规则的、随机的、自然的图像与视频进行合成结果证实此方法可用于不同的图像的交互融合。

**关键词**：纹理合成， 基于图像的渲染， 图像和视频处理， 机器学习， 自然现象

Abstract

In this paper we introduce a new algorithm for image and video texture synthesis. In our approach, patch regions from a sample image or video are transformed and copied to the output and then stitched together along optimal seams to generate a new (and typically larger) output. In contrast to other techniques, the size of the patch is not chosen *a-priori*, but instead a *graph cut* technique is used to determine the optimal patch region for any given offset between the input and output texture. Unlike dynamic programming, our graph cut technique for seam optimization is applicable in any dimension. We specifically explore it in 2D and 3D to per- form video texture synthesis in addition to regular image synthesis. We present offset search techniques that work well in conjunction with the presented patch size optimization. We show results for synthesizing regular, random, and natural images and videos. We also demonstrate how this method can be used to interactively merge different images to generate new scenes.

**Keywords：**Texture Synthesis, Image-based Rendering, Image and Video Processing, Machine Learning, Natural Phenomenon.

1引言

在计算机图形应用中，由较小的样图生成新图被广泛认为时很重要的技术。例如为了渲染图形场景，需要基于样图的纹理合成技术生成大量逼真的纹理。类似的基于样图合成首要原因在于纹理的概念，纹理通常被定义为可以通过一个平稳随机过程建模的无限模式。本文中，我们提出一种新的方法可以从少量的训练数据生成这样的无限模式；使用纹理中的下一块样图，生成具有相似随机特点的更大图形。具体的说，我们通过复制输入像素快生成纹理。算法首先寻找合适的位置放置像素块，然后利用graph cut技术找到最适合转化到输出的像素块。我们的方法中，纹理并不限于空间（图像）纹理，还包括时空（视频）纹理。此外，我们的算法支持输出的迭代优化，通过像素块接缝的连续优化。

我们希望合成的纹理感觉上与样图纹理相似。感知上的相似通过马尔可夫虽机场MRF（Markov Random Field）变现。通过节点网格表示输出纹理，每个节点代表输入纹理中的一个像素或是领域像素。一对节点的边际概率取决于他们领域像素的相似性，从而来自相似领域的输入纹理在生成纹理中的相邻，保持输入的感知质量。纹理合成的目标可以重新表达为在整体相似性最大的条件下，确定网络节点。这个图形模型的概率问题的推论，在机器学习中时很著名的环状网络下的NP-hard（非确定性多项式难题）。因此，所有利用MRF对纹理建模的方法都需要计算近似值最优解。

通过复制像素块生成输出的纹理合成算法必须确定以下两点：

1. 输入纹理与输出纹理的相对位置（像素块偏差）；
2. 来自输入纹理的像素块（像素块接缝）。

本文首要贡献就是纹理合成的算法，该算法将问题化为一个最小成本图割问题：在MRF网格增加特殊节点，计算两个特殊节点之间的最小切割成本。这条切缝就代表着图像接缝的最优解。本文建议在每次循环中使用不同的算法寻找像素块偏差。这些算法通过大片像素块的匹配确保纹理的大比例尺寸结构。实验结果显示，我们所用的接缝优化技术的灵活性是偏差搜索算法的先决条件，两者缺一不可。

**2 相关工作**

纹理合成技术可以大致分为一下三类：

第一类是通过固定数量的参数描述纹理。Heeger和Bergen使用整个视频的颜色直方图作为纹理描述。Portilla和Simoncelli的模型中包含了各种小波特征及其关系，被认为是迄今为止最好的纹理合成参数模型。Szummer、Soatto、Wang等人提出了视频参数化描述。参数模型不能像其他本文所提到的其他模型一样适用于多种变量的纹理合成，但是参数模型是对纹理更好的概括，也更有利于内部检查与识别。因此，参数模型在纹理分析与感知过程的理解上变现优异。

第二类纹理合成的方法是非参数的，该类方法通过一系列的样图对纹理建模。1997年率先利用该技术从一系列的多尺度滤波相应的生成纹理。使用矢量化加速处理速度。这些技术的共同点是每次只生成一个像素。

第三种，也就是最近出现的方法是，通过拷贝输入的整个像素块来生成纹理。Ashikmin使用基于像素的像素块拷贝技术，有利于像素块的统一。该技术在合成多变量纹理的应用中有着很大的优势，但是不想参数模型那样，它只能提供有限的纹理分析信息。

纵观不同的合成技术，纹理经常被描述为MRF。我们使用图割技术优化MRF。采用Boykov提出的图割算法，该算法特别适合纹理合成中的成本函数。

**3 图割技术**

我们通过复制样图中的不规则像素块到输出图像中合成纹理。像素块复制过程分为两步：（1）选择一个矩形纹理块；（2）通过图割算法在矩形纹理块中选出最优的不规则区域用于合成。

在介绍图割技术之前，我们先了解Efros和Freeman合成纹理所用的image quilting算法。该算法将样图中的小像素块复制到输出图像中。每一个像素块是随机复制的，接下来像素块都要与前一个像素块部分重叠，重叠区域的宽度基本上为4或8个像素。使用动态规划法选择重叠区域的最小成本路径。

新旧路径之间的像素匹配质量通过颜色差异比较。在Graph Cut算法执行过程中，最佳路径是由许多相关的像素点结合而成，像素点的生成是从每一对像素点中选择而来，最简单的匹配标准是通过比较相邻像素点的颜色误差。设s和t是重叠区域中的两个相邻像素点，A（s）和B（s）分别代表像素点s在旧纹理块上的像素颜色值，通过定义匹配质量M来计算颜色误差权值，具体的公式如下：

*M*(*s*,*t*,**A**,**B**) = ∥**A**(*s*)−**B**(*s*)∥+∥**A**(*t*)−**B**(*t*)∥

把纹理块重叠区域的每一个像素当作一个节点，相邻像素点s和t节点的弧代表匹配质量权值M。同时添加两个节点分别代表旧纹理块A和新纹理块B，某些像素点与这两个节点之间的匹配质量权值是无限高的，这些就是约束弧，约束有些像素必须来自特定纹理块。

**3.1 旧接缝**

上面的例子并没有完全展示出纹理合成中图割技术的功能。假设输出纹理中已经存在几块像素块，我们想添加一个新的像素块，当我们在放置像素块的时候，可以通过图割技术的弧计算旧像素块边界的接缝。并将这些旧接缝值计入新的图割问题，因此来确定哪些像素块应该覆盖这些旧接缝。

这些弧的权值就是创建旧接缝时所得到的匹配权值。如果接缝节点和新纹理块之间的弧被切断，输出纹理中留有旧接缝；没被切断，意味着该节点被新像素改写，最终权重计算不计入内；如果接缝节点和相邻像素之间的弧被切断，那么相同位置上新的接缝和权重被引入。

当且仅当接缝节点处的一条弧被包含在最小切里的时候，接缝权值与最小图切的等式成立。这条弧的权重是新接缝权重，如果没有弧被切断，那么接缝移除，权值归零。如果M满足三角不等式，意味着从接缝节点中选两条弧的权值总比选择其中之一大，因此至多选择一条弧。

**3.2 环绕的区域**

目前为止，我们讨论的都是沿着已有像素块边界重叠的新像素块。实际上在纹理合成中，在已合成区域之中添加新像素块被限制来自现有像素块。这些约束反映在从边界像素到A的弧，我们还放在一个来自内部像素到节点B的约束弧，来确保至少一个像素来自B。红色线显示了如何将所得图切割形成一个封闭环，来决定了最优不规则形状的区域。

**4 像素块的放置和匹配**

我们采用三种不同的方法筛选像素块：（1）随机放置；（2）整体匹配；（3）子像素块匹配的这些算法中，我们规定路径选择未使用的偏差。此外，对于后两个基于匹配的算法，我们首先寻找当前纹理中需要很多改进的区域。使用接缝成本量化图像特定区域的误差，并拾取具有最大误差的区域。一旦我们选择了错误区域，像素块的选择算法只能挑选完全覆盖该区域的。当纹理正在进行初始化，即当它不完全覆盖输出纹理的像素块，为了不同的目的选择不同的错误区域，以便输出纹理包含初始化和未初始化的成分，保证了纹理被一定量的延伸且延伸部分与已初始化部分一致。

下面详细介绍这三种算法：

随机放置：把新的纹理块转化到随机偏移位置，通过graph cut算法挑选一块区域放到输出中，重复这个过程，这是最快的合成方法，对于随机纹理效果很好。

整体匹配：由于概率函数生成可能的变换，并从中随机挑选新纹理；利用重叠区域的面积对SSD权重归一化处理。

子像素块匹配：从输出纹理中选择一小块纹理，在输入中找与之相匹配的子纹理块。

**5 图像合成**

我们应用该技术在图像和视频纹理合成产生规则的，结构化的和随机的纹理以及合成自然图像。在本文中坚果和键盘的图像，我们用整体的匹配作为我们像素块选择算法，而该子块匹配用于生成图像。图像的运算是相当快的，主要是由于基于FFT的搜索，这里主要介绍的所有图像合成结果用了5分钟运行。百合图像花了5分钟，因为它最初产生为1280\*1024的尺寸。

添加像素块：我们的算法依赖于将输入接插适当和确定有效接缝。上文讨论在输出区域添加像素块的可能性，实际上还有其他的变换，像是旋转、缩放、放射或是投影。这些变换在图片纹理合成中的应用提供了更多灵活性和多样性的输出。但是但这些变换增加时，搜索成本也在提高。因此，我们限定变换的数目，生成输入的变化版本后开始合成。将变换后的图像并列放到一副单独的图像中，该图像中只有包含特定变换版本的部分参与图切割算法。

在上文中，我们利用旋转和镜像变换减少在橄榄图像合成的重要性。缩放允许不同尺寸纹理元素混合在一起。缩放的一个有趣的应用是生成图像深层透视。我们可以约束不同部分输出的纹理从输入纹理的不同尺度复制。单调改变输出图像的尺度，生成能够描绘立体的图像。

交互式合并与混合：在这种应用中，我们挑选了一系列的图像，并将它们结合起来，形成一个的输出图像，正如前文解释的纹理合成，如果我们限制输出图像的某些像素的来自一个特定的像素块，则图切割算法发现最好接缝经过不受约束的像素。这里所说的像素块就是我们想要结合起来的。用于合并两个这样的图像，用户首先指定了源图像在输出中的位置并以交互方式建立约束像素，切割算法然后找到图像之间的最佳接缝。

这是一个强大的结合不相似的图像的方法。注意，等式的成本函数生成的接缝要经过边缘。我们的结果表明两种接缝存在于合成的输出图像。此文件的扉页上旗帜通过组合花和叶交互生成：用户需要将花朵放置在树叶背景，约束输出的一些像素来从花内，图割算法半自动计算花和树叶之间的适当接缝。

**6 视频合成**

图形切割技术其中一个主要的优势：允许直接扩展到视频合成。我们认为一个视频序列是其中一个时间轴的3D集合体系。在视频的情况下，像素块就是整个视频的3D空间－时间块，可以在任何地方放置三维卷。因此，图像纹理合成的两个步骤，放置像素块和接缝生成，在这里是需要的。

与二维纹理合成类似，视频像素块的选择必须基于视频的类型。一些视频序列只显示时间的平稳性，而另一些则是在空间以及时间都具有平稳性。对于那些只显示时间的平稳性的视频，搜索三个纬度（时间和空间）的像素块是不必要的，我们可以限制仅在时间域进行搜索，对于时间和空间上平稳的视频，则需要在所有三个纬度进行搜索。

为了更找到更好的时空相对偏移像素块，我们首先发现了成对图像比是一个很好的转化。然后计算在一个窗口内最佳接缝周围有限数量的接缝转化。其结果是等效于每个像素转化的时间，而不是整个图像的过度时间。所得接缝然后可以重复，以形成一个视频回路。对于每一个序列，我们计算在一个最佳接缝60帧时空窗口最好的转化。例如瀑布、草、池塘、喷泉和海滩。前人的瀑布序列结果存在间歇模糊，我们生成的序列，没有任何可感知的人工痕迹。

随机时间偏差，对于非常短的视频，循环会导致非常明显的周期性，在这种情况下，可以通过应用一系列输入纹理像素块合成视频。我们的实际应用主要在火、烟火、海洋和烟雾四个方面。

视频的空间扩展：在空间和时间的变换的情况下，还可以增加视频序列帧的大小。例如，河流序列的空间分辨率成功地从170X116提高到210X160.通过使用时间约束，还能够将此扩大的视频序列进行循环。视频的合成的运行时间在5分钟至1小时之间，取决于视频的大小和搜索采用的方法（寻找单一的时间偏差比对时空搜索的速度更快）。FFT应用很大程度上提高了工作效率。

**7 总结**

我们已经证实了新算法可用于图像和视频的合成，图形切割算法在计算区域接缝和决定放置像素块上的应用很理想。本文展示了各种合成的例子，包括结构化和随机的图像与视频纹理的合成。扩展的功能允许输入像素块的改变，可应用于两种非同源图像的交互融合。该技术的显著改善表现在：

（1）不限制生成接缝区域的形状；

（2）考虑到旧接缝；

（3）容易推广创造接缝；

（4）易于添加约束。

参考文献

[1] Exploring Photo Collections in 3D. Noah Snavely. Steven M.Seitz Richard Szeliski. IEEE Transactions on Visualization and Computer Graphics.