

分析式纹理合成技术及其在深度学习的应用

Analyzed Texture-synthesis Techniques and Their Applications in Deep Learning

李宏林

Honglin Li

日本山梨大学--大学院--生命情报系统系，日本山梨县甲府市武田 4 丁目 4-37，邮编：400-8510
Yamanashi University Graduate-School Life-Information System Course, Japan Yamanashi-ken Kofu-shi 4
chome 4-37, Postcodes: 400-8510.

摘要：当前国际主流的非参数和参数法分析式纹理生成技术，对于计算机视觉领域的图像纹理合成具有一定的借鉴意义。在概括总结比较分析式纹理生成技术原理、框架结构、应用发展趋势及其优缺点的基础上，分析了基于 graph cut 模型的非参数法、基于 P&S 模型的参数法两种典型纹理生成技术以及广泛应用于图像处理领域的深度学习新技术—卷积神经网络（CNN）的结构与原理，进一步讨论了以基于 CNN 的 Caffe 网络框架及在 2014 年 ImageNet 图像分类和目标识别大赛上取得优异成绩的 VGG 模型为基础的分析式纹理生成模型 VGG-19 的工作原理及其在人脑视觉分析研究方面的应用。分析结果表明：相对于普通参数法和基于 CNN 网络模型的参数法，非参数法具有更快的处理速度，可生成更高视觉质量与更多种类的目标纹理图；参数法适合作为纹理合成领域的分析研究工具；卷积神经网络应用到参数法中，可大幅缩短特征量设计与参数调整周期并提高合成效果，进一步提升了参数法作为理论分析和应用实现工具的价值。

关键词：分析式纹理合成法；非参数法纹理生成；参数法纹理生成；深度学习；卷积神经网络；VGG-19.

中图分类号： TP37 文献标识码： A

Abstract: The state-of-the-art analyzed texture synthesis techniques are divided into non-parametric and parametric methods, which contribute to the current corresponding research works on computer vision. By summarizing and comparing their principles, structures, development trends, advantages and disadvantages, a non-parametric method based on graph-cut model and a parametric method based on P&S model are analyzed in detail. In addition, the structures and principles of Convolution Neural Network (CNN) based on Deep-learning which are widely applied in image-process filed are also discussed. Finally, a new texture synthesis model VGG-19 is introduced, which is the combination of CNN-based Caffe network with VGG model that obtained high scores in the 2014 ImageNet classification and object detection competence. The VGG-19 model can be also used to analyze human visual process. The analyzed results show the facts as below. Non-parametric methods can synthesize high-quality textures of various kinds with high speed. Parametric methods are appropriate for being used as analysis tools. The parametric methods based on CNN can greatly reduce the time period of designing and adjusting feature representations and parameters and improve the synthesized results synchronously, which is proved to be valuable tools for analyzing theory and realizing applications on texture-synthesis work.

Keyword: Analyzed texture synthesis method; Non-parametric texture generation; Parametric texture generation; Deep learning; Convolutional neuron network; VGG-19.

纹理是图像的重要特征之一，纹理图是数字图像中的一个重要类别，分规则纹理图和非

作者简介：李宏林（1979-），男，福建泉州人，泉州医学高等专科学校讲师。2002 年获得中南大学自动化专业工学学士学位，2012 年获得华侨大学计算机科学与工程硕士学位，2015 年 9 月迄今在日本山梨大学攻读计算机视觉博士学位，主要从事计算机视觉、图像处理、数据挖掘等方面研究。

规则纹理图两类;前者整体图像的各部分重复均匀分布,后者图像各部分总体相似但在**大小、方向、颜色等**方面随机分布。当前纹理生成技术主要有非参数和普通参数法两大类。非参数法通过建立或改进一系列模型、方法和算法去生成纹理图像,主要有**基于滤波器采样的模型、基于像素拷贝的模型以及基于片拷贝的模型**三大类;基于滤波器采样的模型通过对样本采样纹理图的一系列不同分辨率图像的滤波结果进行采样分析生成目标纹理图,基于像素拷贝的模型通过从样本采样纹理图往输出图区域拷贝像素生成目标纹理图,基于片拷贝的模型通过拷贝纹理片生成目标纹理图。普通参数法通过在一个紧凑完备(参数数量适中并有效,不易发生过度训练)的参数模型中建立、调整参数信息描述纹理特征,并基于这些特征描述生成目标纹理图;其参数类型主要有**基于像素法、基于距离测量以及基于分析法**的统计量三类;**基于像素法**是通过建立具有与样本纹理图相同的 N 阶统计量像素的生成目标纹理图,**基于距离法**通过最小化样本纹理图和生成纹理图之间的距离差异特征量生成目标纹理图,基于分析法利用渐进式分析方法分析采样纹理和目标纹理,并通过最小化损失函数等方法逐渐缩小两者间的差距进而生成最终目标纹理图。非参数和普通参数纹理生成法主要区别在于:前者目标纹理图的建立源于一系列模型、方法、算法;后者通过对前人研究成果的分析以及对样本纹理图特征的观察和数学分析手动建立一系列参数,形成参数纹理模型,并利用该模型生成目标纹理图。非参数法的优点在于拥有更快的处理速度并生成更高视觉质量与更多种类的目标纹理图;其缺点在于无法为纹理图建立对应的特征表达,而且不适合作为一种分析研究工具。普通参数法的优点**在于可以为各类纹理图建立相应的特征表达进行推广与改进**,同时可作为分析研究工具应用于多个领域;其缺点在于建立的纹理种类相对有限,某些生成纹理效果并不理想,此外由于其参数是人工分析建立的,设计或改进参数周期很长。

深度学习(Deep learning)是近年来兴起的新技术[1],广泛应用于图像处理及计算机视觉领域;卷积神经网络(Convolution neutron network, 以下简称 CNN)是基于深度学习的一项重要应用,是一种基于大数据自动学习的端到端参数模型。利用 CNN 自动训练生成的特征表达可以推广应用到其他各种数据集,显著节省了人工设计特征的时间周期;同时,由于其在**图像分类和目标识别**领域方面的准确率已经大幅度超越了传统机器学习方法,基于 CNN 的参数模型纹理生成方法正逐渐取代普通参数法,其合成纹理质量也正逐步向非参数法逼近。

为此,在概括总结并对比分析基于 graph cut 模型的非参数法和基于 P&S 模型的参数法两个典型的国际主流分析式纹理合成技术以及深度学习卷积神经网络(CNN)的原理、框架结构、应用发展趋势和优缺点的基础上,进一步讨论了结合基于 CNN 的 Caffe 网络和 VGG 结构的 VGG-19 纹理合成模型,通过对其结构框架、特征量建立与优化等过程原理的分析,可以发现深度学习 CNN 网络的引入有助于参数法纹理合成模型的建立与优化,能有效提高参数法纹理合成效果并缩短改进周期。

1 非参数纹理生成法 Graph-cut

Graph-cut 是一种**著名的非参数纹理生成方法**[2],基于该方法建立的纹理图效果迄今依然超越大量普通参数法甚至 CNN 参数法。该方法运用了片拷贝的纹理生成法,可**实现纹理图像生成、纹理图像优化、不同种类图像合成以及视频合成**等多种应用;其运行过程主要包含两大核心步骤,选择片放置位置与获取优化块:1)利用随机法或匹配法在输出纹理图区域**选择放置采样纹理片的位置范围**(等于或小于采样纹理区域的矩形区域);2)利用 graph-cut 算法确定采样纹理片的优化提取块(一般为小于采样区域的不规则形状),拷贝该优化块到目标纹理图的放置区域。其原理如图 1 所示:

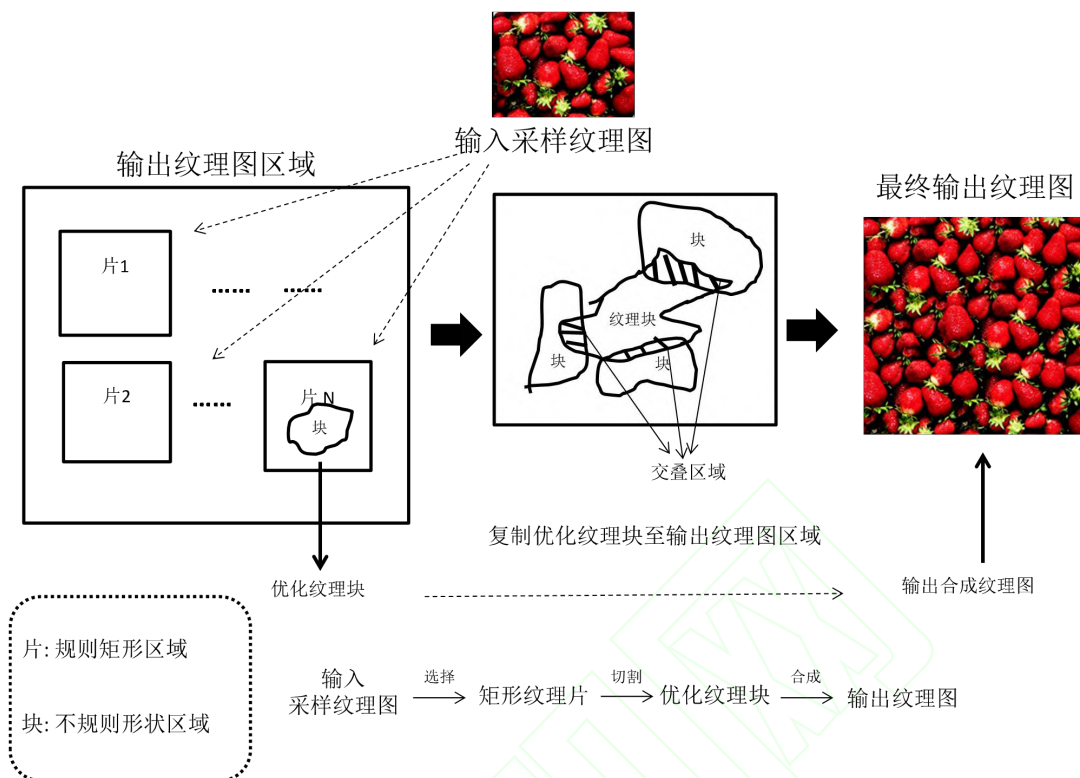


图 1 Graph-cut 算法纹理合成基本原理

Fig 1 Graph-cut algorithm texture synthesis principle

1.1 选择片放置位置

片放置方法主要有随机放置和匹配放置两类: 随机放置指将整个输入采样纹理片每次随机放置在输出目标纹理图区域的任意一个位置, 这种方法耗时最短, 主要适用于生成随机纹理图; 匹配放置分两种方式: 第一种是整片匹配放置, 该方法每次都放入整个输入采样纹理片的最优块 (通过计算重叠区域的最小代价生成), 适用于合成规则性纹理图; 第二种是子片匹配放置, 该方法每次放入输入采样纹理一个局部区域片 (远小于采样纹理片) 的最优块, 适合合成随机纹理图和纹理视频。

其放置过程分初始化与优化两个阶段。初始化阶段主要将采样纹理片的最优块逐步填充到空的输出纹理图区域中, 分以下三个步骤: 1) 将第一个采样纹理片放置在输入纹理区的随机位置; 2) 选择第二次片放置位置 (与上一次放入的部分或全部交叠), 运行 graph-cut 算法获取放入片的优化块后放入该交叠位置; 3) 重复第二个步骤直至填满整个输出图区域。优化阶段是对已填充完毕的输出纹理图再次填充采样输入纹理信息达到修正改进目的, 分以下两个步骤: 1) 利用代价函数计算出已填满的输出纹理图的最大误差块, 在该范围内放入输入纹理片的优化块区域; 2) 重复上一个步骤, 直至对最终结果满意结束迭代获得最终优化输出纹理填充图。

1.2 获取优化块

获取优化块的核心方法是运用 Graph-cut 算法, 该算法是基于图的最小代价路径切取方法, 是在 Dynamic program 算法[3]上的进一步改进。Dynamic program 算法用于选择两个片的边界交叠区域的优化块, 其原理是利用该区域的相邻配对像素点分属新旧片区的色差值之和来计算最小代价位置, 其步骤为: 1) 对交叠区域逐行扫描相邻像素对, 计算像素对属于新旧区域时的色差值的和, 记为代价值; 2) 记录每一行的最小代价值点位置, 继续扫描下一行; 3) 重复上一步骤直至扫描整块交叠区域结束, 连接每行的最小代价值位置点, 生成

最优块的边界线。

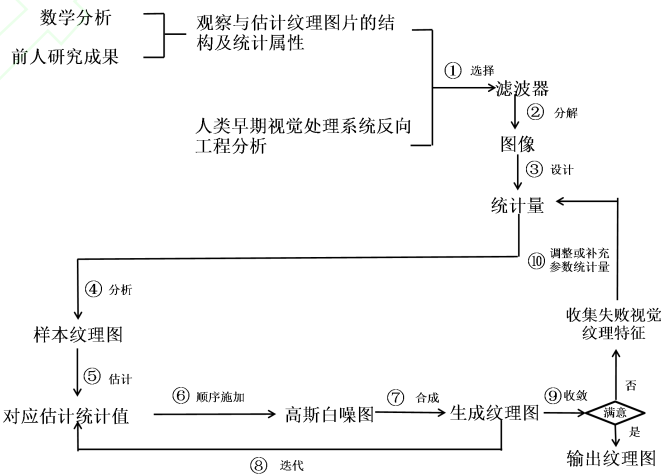
Graph-cut 算法在 Dynamic program 算法的基础上从处理两个边界域交叠区进一步扩展到处理多个边界域或包围域交叠区，通过在旧优化块切割线（边界线）的相邻像素节点间插入块节点，建立新的连接弧并重新计算生成最小代价切割线（优化块边界线）。该算法主要步骤是：1）在旧优化块边界线的左右相邻两像素节点间插入新的块节点；2）在每个块节点和新片之间建立一条新的连接弧，计算该块节点与相邻两像素节点及新片之间三条连接弧的代价值；3）存在三种新的切割方式，分别是（A）从块节点与新片区的连结弧切入、（B）从三条连接弧以外的区域切入、（C）从块节点与相邻像素节点之间切入，分别对应保留、取代、移除旧块切割线。

1.3 改进与扩展应用

Graph-cut 算法还引入三种方法优化合成效果：运用梯度计算法用于优化代价函数以更好的检测边缘，运用羽化和多分辨率处理法隐藏或移除视觉假象，运用快速傅里叶法以加速基于平方差和的交叠区域代价计算算法。此外，该算法还实现了多种扩展应用：对输入采样纹理进行翻转、镜像、缩放变换处理以增加生成纹理的多样性，与用户交互式方法结合实现不同种类图像合并以及从二维纹理图像生成扩展到三维纹理视频合成。但是该算法往往难以很好地合成具有明显边界结构特征的纹理，尤其是对那些独立性很强、排列紧密并具有明显边界的色彩模式单元[4]。为了更好的保持纹理的边界结构特征，国内研究人员先后提出了基于边界图的纹理合成法[4]、基于不规则块的纹理合成法[5]与基于 Graph Cut 的快速纹理合成算法[6]对其加以改进，提高合成效果并缩短合成时间。

2 参数法 P&S 模型纹理生成法

基于样本的纹理合成技术是一种重要的参数法纹理生成技术[7]，它是一种基于给定的小区域纹理样本并按照表面的几何形状拼合生成整个图像纹理的方法，其生成纹理在视觉上是相似而连续的[8]，代表性模型是 2000 年国外学者提出 P&S 模型[9]，该模型迄今仍是普通参数法纹理生成的重要基础模型。P&S 模型是一种基于分析合成的模型，它通过观察前人设计的参数并运用数学推理和实验分析设计及改进参数，利用最大熵方法采集样本图像像素密度值，运用实践遍历性方法（对图像空间信息取平均值）估算样本图像的参数值，生成可推广的参数纹理生成模型。该模型利用一系列小波特征量和它们的相互关系建立边缘、系数相关性、系数大小、跨尺度相位四类统计量参数，运用梯度投影法将这些特征统计量参数先后施加在高斯白噪图上渐进式合成纹理，并迭代这一过程直至收敛得到最终纹理图。其原理如图 2 所示。



P&S 模型参数对应的**基础函数特征与运用顺序梯度投影算法**合成纹理图的方法决定了该模型利用可控金字塔滤波器[10]而非高斯滤波器分解图像特征；可控金字塔滤波器方法具有**良好的图像重建特性、平移不变性和旋转不变性**，包含了方向波段和缩放比例（金字塔层次）两套系数。P&S 模型以小波多分辨率原理为基础，运用可控金字塔滤波器将源样本纹理图分解成多个不同尺度纹理图像，进而设计了四类参数统计量：边缘统计量包含了**偏差、峰度、尺寸、像素深度差异等信息**，系数相关性统计量对应周期性、全局性的方向纹理结构，系数大小统计量体现了纹理边角信息，跨尺度相位统计量则反映了纹理立体阴影细节。

但是 P&S 参数模型方法具有如下缺点：1）对于纹理图，无法区分直线和曲线轮廓，无法捕捉**线条末端粗细度**，无法形成封闭式轮廓，无法区分某些样例中的线条与边缘；2）对于非纹理图，只能捕捉到局部结构信息而无法得到图中不同目标的整体空间联系；3）虽然该模型在处理对应样本库中的数百种纹理样本总能控制迭代次数 50 次左右，但在理论上还无法确定参数模型何时收敛；4）由于参数是手动设计的，所以无法从理论上验证是否还有更优秀的参数集存在。P&S 参数模型方法主要适用于**单种纹理样本合成**，国内研究者后续提出**基于相关性原理**的多样图纹理合成方法以增加生成纹理种类满足更多需求[11]，结合图象细节特征的**全局优化纹理合成方法**也被提出用于更好的解决基于样本的纹理合成中纹理图像不连续、纹理结构或特征容易断裂等情况[12]。

3 卷积神经网络

模式识别、机器学习和深度学习技术长期以来广泛应用于计算机视觉方面，在 2012 年的基于 ImageNet 图像数据库的 ILSVRC 大赛中，以基于深度学习的 CNN 方法设计的 AlexNet 网在图像分类比赛上获得了第一名，其 Top-5 准确率（五次选择有一次正确即纳入正确分类）超出了基于传统机器学习的第二名 SVM 方法达到 10% 以上，此后 2012 年至 2015 年的图像分类冠军**均为基于 CNN 方法的模型**。CNN 网络主要包含**卷积层、池化层和全连接层**三种层次结构，使用**损失函数来评估系统性能**，利用**梯度下降算法**优化卷积层和全连接层连接权值；其所处理的输入图像一般具有宽度、高度和通道数三种属性，比如 ImageNet 的 224×224×3 的标准输入图像即指像素宽度、高度均为 224 以及包含 RGB 共 3 个通道的输入图像。

与传统神经网络相比，CNN 的不同点在于：1）主要用途是将**输入的传感信息**转为有效的**特征表达**，这些表达目前**在大尺度图像识别领域已经超越了人工特征**；2）用卷积层和池化层代替了大部分全连接层，应用**局部连接、参数共享和特征图方法**，在大幅度降低参数数量的同时尽可能保有图像特征；3）通过**前向卷积与池化进程**将输入图像分解为**特征图表达**，再通过**反向传播进程调用梯度下降算法优化卷积层和全连接层连接权值**，迭代前向和反向进程直至收敛完成网络训练生成最优参数组合。与传统机器学习方法相比，CNN 的不同点在于：1）SVM、KNN、Boosting 等传统机器学习方法都是**浅层网络结构**，网络深度远低于 CNN；2）传统机器学习方法一般使用**人工设计特征**，比如用于**目标检测的 SIFT 特征**、用于人脸识别的 LBP 特征与用于行人检测的 HOG 特征，其设计周期长效率低，CNN 通过自动从大数据分析学习自动优化调整特征参数，大大缩短了特征设计周期；3）传统机器学习方法的**特征提取和分类阶段是相互独立先后进行**，CNN 是一种端到端系统（前向输入卷积与反向输出梯度下降优化过程迭代进行至收敛到最优），其特征提取和分类过程相互作用同步优化，是全局性的统一过程，两者不可分割。

3.1 卷积层

CNN 卷积层**处理输入图像生成输出特征图**，其 3 个重要特性分别是**局部连接、参数共享与特征图建立**。局部连接原则借鉴了人眼视觉神经成像原理，即每个神经关注的只是图像的某个局部位置而非全图，最后再合成所有神经的局部观察结果形成最终成像结果图；参数共享原则令多个神经共享一套参数，该原则适用于有一定规律的图像（比如纹理图），对非规则图像应适当放宽条件；**特征图建立**则是通过**增加特征图数量增强图像表达能力**（一张特

征图对应源图像的某一方面特征)以弥补前两个特性造成的图像信息损失。以神经通过 3×3 的滤波器去观察 6×6 分辨率的输入图像为例:设滤波器卷积横向纵向滑动步长值均为 2,共需移动 9 次方能处理完整幅图像(即用 9 个神经方能观测整幅图像):若采用全连接方式观测,共需 $6 \times 6 \times 9$ 个连接权值(每个像素对应一个连接权值);运用局部连接原则,设每个神经只观测该图像的 3×3 局部区域,则只需 $3 \times 3 \times 9$ 个;运用参数共享原则,令 9 个神经均使用同一套参数,则进一步缩减为只需 $3 \times 3 \times 1$ 个;建立 10 张特征图表达图像,则需 $3 \times 3 \times 10$ 个;最终参数数量(连接权值数量)缩减为最初的 27.8%,上述特性在处理大尺度分辨率图像时,缩减参数幅度更为明显。

每个卷积层包含的参数数量由输入特征图数量和输出特征图数量共同决定。假设某卷积层接收的输入信息为 64 张特征图,输出信息是 128 张特征图,使用的是 3×3 的滤波器(不考虑偏差值时,则每一套连接权值参数为 $3 \times 3 = 9$ 个);每个滤波器用 64 套不同参数对 64 张输入特征图进行卷积,每层需 128 个滤波器生成 128 张输出特征图,因此该卷积层所包含的连接权值参数个数一共是 $64 \times 3 \times 3 \times 128 = 73728$ 个。

3.2 池化层和全连接层

池化层采用了基于图像多分辨率处理原理的金字塔式图像处理方法,它利用向下采样技术缩小输入特征图的大小进而减少连接权值参数数量;一般采用 2×2 大小(3×3 及其以上大小的池化层会造成信息损失过大)步长 2(不重叠)的池化层,该池化层可以将输入特征图的宽度和高度分别缩小为原来的二分之一。池化层主要有平均值、最大值、随机值池化层三类,目前以最大值池化层使用最为频繁, 2×2 大小的最大值池化层就是取池化窗口所套入的邻域空间的四个像素值的最大值作为该邻域空间的输出值。池化层除了具有缩小特征图大小特性以外,还保持了图像的平移、旋转和缩放不变性,其多分辨率处理方式令整个 CNN 网络能更好的识别图像的不同区域。

最大连接层一般放置于 CNN 网络的尾部,用于将卷积层和池化层处理分解出的图像各种特征表达合并输出作为最终分类层的输入信息;但由于 CNN 网络的连接权值参数主要产生在全连接层,为了减少计算代价,全连接层正逐渐被移除或转换成卷积层。

4 Caffe 框架和 VGG 模型

4.1 Caffe 框架

Caffe 是 CNN 的重要框架之一[13],包含了当前主流的深度学习算法和一系列参考模型(如 Alexnet 网);该框架将模型表达和实际应用分离,支持 C++ 和 Python 语言,广泛应用于图像分类、目标识别、语义特征学习等领域。Caffe 采用名为 Blobs 的 4 维数组存储数据(图像坐标、权值和偏差),利用谷歌协议缓冲区存储 CNN 网络模型,调用 CUDA 核进行 GPU 卷积运算,包含了前向进程(接收输入数据生成输出数据)和后向进程(利用随机梯度下降算法优化连接权值),并不断迭代前向后向进程直至收敛得到最优参数网络。

4.2 VGG 模型

VGG 模型[14]是建立在 Caffe 框架下的一个实用型 CNN 模型,在 2014 年 ImageNet 比赛中的图像分类和目标识别方面分别取得了第二名和第一名。相对于 2012 年比赛冠军(11×11 滤波器,步长为 4)和 2013 年冠军(7×7 滤波器,步长为 2),VGG 网的主要改进在于采用了更小的滤波器(3×3 滤波器,步长为 1)和更深的层数(19 个权值层)。由于小型滤波器能够表达更多输入图像特征进而使得决策函数更具区分性,同时还能减少参数数量,VGG 模型使用 3 个堆叠的 3×3 滤波器卷积层取代 1 个 7×7 滤波器卷积层(其卷积输出结果图像大小相等,缺点是由于层数增加导致计算代价增大)。一共有 6 个模型

(A-LRN/VGG-11/VGG-13/VGG-16/VGG-16*/VGG-19)参与了研究分析,前四个主要用于实验测试对比以及参数初始化,最后两个 VGG-16* 和 VGG-19 分别用于目标识别和图像分类。VGG-19 模型一共包含 16 个卷积层、5 个池化层和 3 个全连接层以及 1 个 soft-max 分类层,

其中含连接权值的卷积层和全连接层一共 19 个（VGG-19 名字源于此），其结构示意图如图 3 所示（图中各层详细示意图标注于主框架图下方）。由于深层网络的梯度具有不稳定性，一般先用随机连接权值初始化浅层网络，再将训练后的浅层网络的各层连接权值作为深层网络对应层的初始化权值（该深层网络的其余未赋值层连接权值使用随机初始化值），最后使用梯度下降算法自动迭代训练优化连接权值得到最终应用模型。VGG-19 模型的前 4 层卷积层和后 3 层全连接层的初始化权值来自训练后的 VGG-11 模型的对层，剩余 12 层则采用随机初始化值。

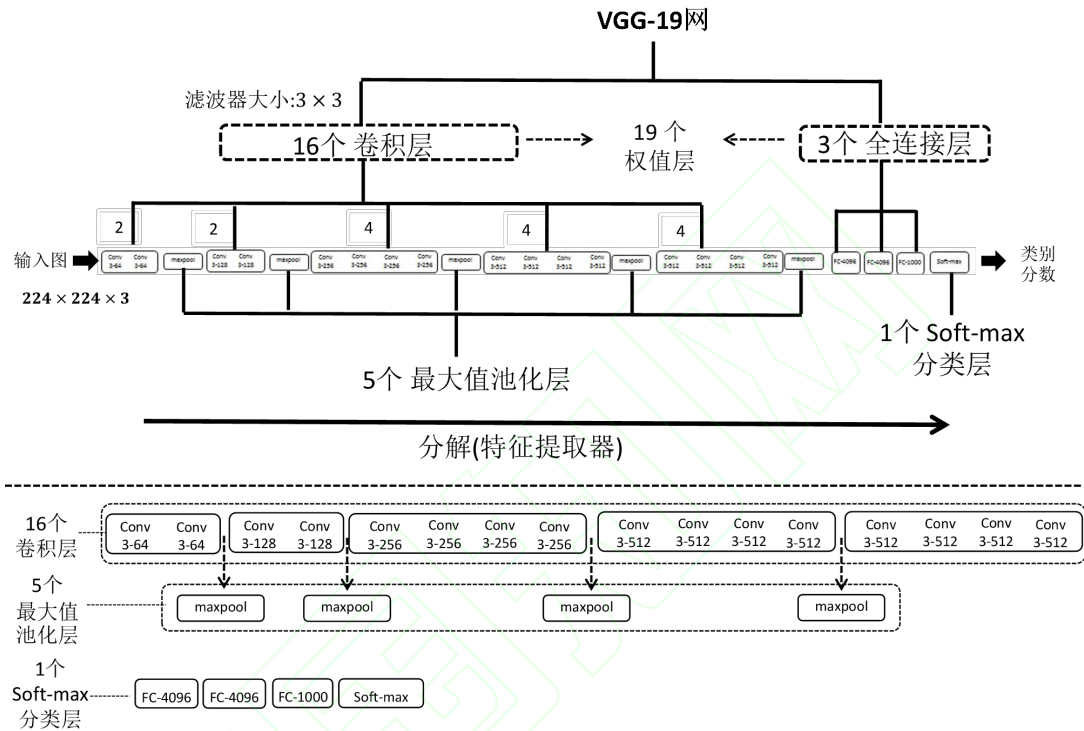


图 3 VGG-19 网络结构图
Fig 3 VGG-19 net architecture

利用大数据集训练完毕的 VGG 网络除了可直接运用于图像分类和目标识别外，还可推广应用到其他小数据集建立特征表达，并结合传统机器学习分类器实现对小数据集的图像分类和目标识别。

5 基于 VGG-19 的纹理生成模型

VGG-19 纹理生成模型[15]的结构是在 VGG-19 模型的基础上移除了最后 3 个全连接层以及将 5 个最大值池化层转为平均值池化层，其工作原理源于 P&S 模型，是一个基于 CNN 的参数法纹理生成模型。VGG-19 模型建立了一种新型特征表达（关系矩阵）来表示纹理，可用于合成纹理图；此外，该模型在不同卷积层生成的阶段性纹理还可用于探测分析人脑视觉成像的不同阶段。

关系矩阵纹理特征表达建立步骤如下：假定 VGG-19 模型的第 l 层含有 N_l 个滤波器生成 N_l 个输出特征图（每张输出特征图是对应的输入特征图经由相应滤波器卷积过每一个位置后生成所有信息的加权和），设每个特征图的大小为 M_l （例如，大小为 6×6 的特征图的大小为 36），则第 l 层的所有特征图的集合可以表示为矩阵 $F^l \in R^{N_l \times M_l}$ ； F_j^l 是矩阵 F^l 的第 j 行

行向量，对应的是第 l 层的第 j 个特征图（该矩阵的行向量为特征 $G_{i,j}^l = \sum_k F_{i,k}^l F_{j,k}^l$ 图对应矩形像素图的向量化值，例如将 6×6 的特征图的 36 个值转为 36 个元素值构成的行向量）； $F_{j,k}^l$ 指代的是输入特征图经由滤波器卷积过第 k 个位置生成的输出特征图的相应局部信息值（第 j 行行向量的部分信息）； $F_{i,k}^l \times F_{j,k}^l$ 则对应第 l 层的第 i 个特征图与第 j 个特征图在滤波器滑动到第 k 个位置时产生的局部信息值的点积； $G_{i,j}^l = \sum_k F_{i,k}^l F_{j,k}^l$ 为 i 特征图和 j 特征图在所有位置的卷积信息值点积的加权和，其最终结果是一段局部行向量； G^l 则是第 l 层所有输出特征图两两之间在所有位置卷积信息值点积加权和所构成的矩阵，其中 $G_{i,j}^l$ 是该矩阵第 i 行第 j 列的元素；最终生成的 $\{G^1, G^2, \dots, G^L\}$ 是整个 CNN 网络产生的所有层的关系矩阵的集合，该集合即为相应纹理图的特征表达。上述特征表达形成步骤原理如图 4 所示：

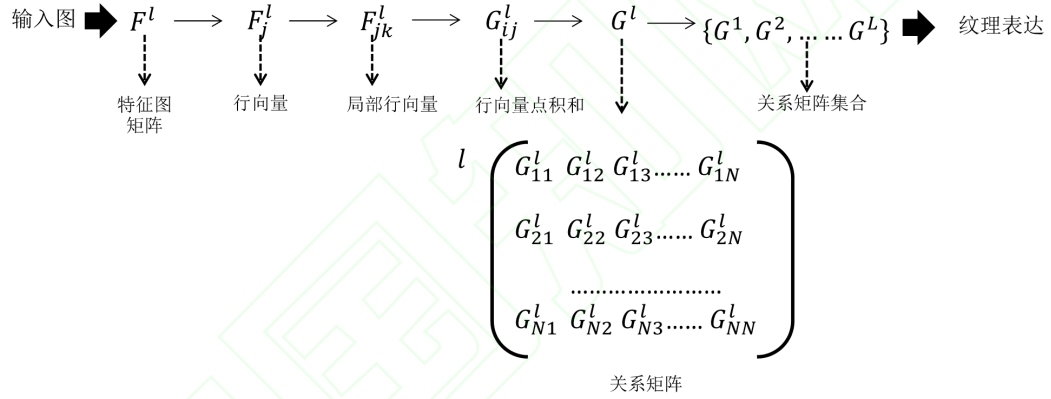


图 4 纹理特征表达形成步骤原理图

Fig 4 Texture representation generation principle

VGG-19 纹理生成模型包含两个 VGG-19 网络，第一个用于生成源纹理图的关系矩阵集表达式，第二个则用于将白噪声图（色谱由白光构成符合高斯分布的适用于初始化的图像）逐层

合成为与源纹理图类似的纹理图像。 $E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{i,j}^l - T_{i,j}^l)^2$ 作为该模型第 l 层的损失

函数（ $G_{i,j}^l$ 和 $T_{i,j}^l$ 分别对应白噪声图与源纹理图的关系矩阵元素）， $L(\vec{x}_t, \vec{x}_g) = \sum_{l=0}^L w_l E_l$ 是整个

网络的总损失函数。VGG-19 利用 BP 算法从最高层开始后向传播计算出整个网络的梯度值（损失函数相对于生成纹理图的导数，每层梯度值是上层传播下来的与层计算的梯度值之和），该梯度值被作为梯度下降优化算法 L-BFGS 的输入值，运行 L-BFGS 算法[16]优化整个模型的连接权值，训练得到新的 VGG-19 模型，将源纹理图和白噪声图输入新模型再次计算损失函数值；重复迭代上述步骤，直至得到的损失函数值达到最小，结束训练生成最终 VGG 纹理合成模型，该模型可将输入白噪声图逐层合成至与源纹理图类似的纹理图，该模型框架如图 5 所示。

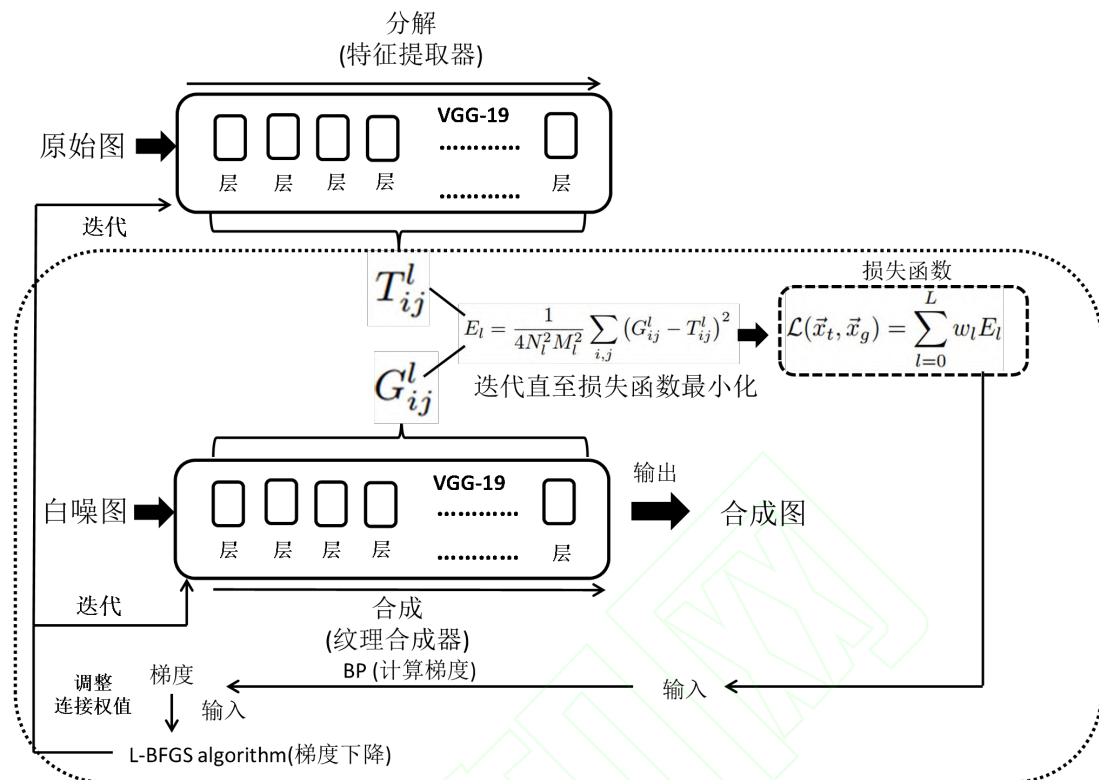


图 5 VGG-19 纹理生成模型

Fig 5 VGG-19 texture generation model

VGG-19 纹理模型建立的纹理图的效果比 P&S 模型及 Caffe 结果更好[15]，但依然弱于 Graph-cut 方法。该模型工作原理类似于人脑视觉成像过程（由低级信息合成中级信息直至最终形成高级成像信息，逐层清晰化），在每个层次上合成的纹理图可以用于探测视觉成像系统，根据不同信号反馈分析模型和成像系统的层与层之间的联系。此外，由于所采用的特征表达为关系矩阵集，该模型生成非纹理图时保留了图像中的目标信息而丢失相关目标之间的全局联系信息。

6 结束语

纹理作为数字图像的一个重要基础特征广泛应用于计算机视觉等图像处理研究领域，分析式纹理合成方法是重要的纹理合成技术之一。以基于 Graph-cut 模型的非参数法、基于 P&S 模型的参数法和基于 CNN 的 VGG-19 网络模型为研究对象，综合回顾并对比分析三者的框架结构、过程原理和实现步骤，并讨论了非参数法、参数法的应用发展趋势和优缺点。通过深入剖析 VGG-19 纹理合成模型原理，得出了引入深度学习卷积神经网络新兴技术的参数法能有效缩短参数设计改进周期并显著提高合成效果的结论。此外，还进行了大量合成式纹理主流与新兴技术的调研，对于计算机视觉领域的图像纹理合成研究工作有一定借鉴意义，下一步工作将调研分析信息应用于相关研究课题项目上，并进一步论证其在相关领域方面的实际功效。

参考文献

- [1] Johnson J, Karpathy A. Convolutional Neural Networks for Visual Recognition[EB/OL]. <http://cs231n.github.io/convolutional-networks>,2015.
- [2] Kwatra V, Schödl A, Essa I, et al. Graphcut textures: image and video synthesis using graph

- cuts[J]. ACM Transactions on Graphics (ToG), 2003, 22(3): 277-286.
- [3] Efros A A, Freeman W T. Image quilting for texture synthesis and transfer[C]. USA, Proceedings of the 28th annual conference on Computer graphics and interactive techniques, 2001: 341-346.
- [4] 杨刚, 王文成, 吴恩华; 基于边界图的纹理合成方法[J]. 计算机研究与发展, 2005, 42 (12): 2118-2125.
- [5] 熊昌镇, 黄静, 齐东旭; 基于不规则块的纹理合成方法[J]. 计算机研究与发展, 2007, 44 (4): 701-706.
- [6] 邹昆, 韩国强, 李闻, 张潇元; 基于 Graph Cut 的快速纹理合成算法[J]. 计算机辅助设计与图形学学报, 2008, 20 (5): 652-658.
- [7] 朱文浩, 魏宝刚; 基于样本的纹理合成技术综述[J]. 中国图象图形学报, 2008, 13 (11): 2063-2068.
- [8] 徐晓刚, 鲍虎军, 马利庄; 纹理合成技术研究[J]. 计算机研究与发展, 2002, 39 (11): 1405-1410.
- [9] Portilla J, Simoncelli EP. A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients[J]. Computer Vision, 2000, 40(1): 49-70.
- [10] Simoncelli EP, Freeman WT. The steerable pyramid: a flexible architecture for multi-scale derivative computation[C]. International Conference on Image Processing, 1995, 3: 444-447.
- [11] 徐晓刚, 鲍虎军, 马利庄; 基于相关性原理的多样图纹理合成方法[J]. 自然科学进展, 2002, 12 (6): 665-668.
- [12] 肖春霞, 黄志勇, 聂勇伟, 刘梦, 何发智; 结合图象细节特征的全局优化纹理合成[J]. 计算机学报, 2009, 32 (6): 1196-1205.
- [13] Ia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[C]. USA, MM '14 Proceedings of the 22nd ACM international conference on Multimedia, 2014: 675-678.
- [14] Karen S, Andrew Z. Very Deep Convolutional Networks for Large-Scale Image Recognition. [EB/OL] http://www.robots.ox.ac.uk/~vgg/research/very_deep/2015-4-10.
- [15] Gatys LA, Ecker AS, Bethge M. Texture Synthesis Using Convolutional Neural Networks[C]. Canada, NIPS'15 Proceedings of the 28th International Conference on Neural Information Processing Systems, 2015, 70(1): 51-5.
- [16] Zhu C, Byrd R H, Lu P, et al. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization[J]. Acm Transactions on Mathematical Software, 1997, 23(4): 550-560.