

❖ 特约专栏 ❖

基于对抗生成网络的纹理合成方法

余思泉^{1,2}, 韩志², 唐延东^{1,2}, 吴成东¹

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110000;
2. 中国科学院沈阳自动化研究所 机器人学国家重点实验室, 辽宁 沈阳 110000)

摘要: 纹理合成是计算机图形学、计算机视觉和图像处理领域的研究热点之一。传统的纹理合成方法往往通过提取有效的特征样式或统计量并在该特征信息的约束下生成随机图像来实现。对抗生成网络作为一种较新的深度网络形式, 通过生成器和判别器的对抗训练能够随机生成与观测数据具有相同分布的新数据。鉴于此, 提出了一种基于对抗生成网络的纹理合成方法。该算法的优点是不需要经过多次迭代就能够生成更真实纹理图像, 且生成图像在视觉上与观测纹理图像一致的同时具有一定随机性。一系列针对随机纹理和结构性纹理的合成实验验证了该算法的有效性。

关键词: 纹理合成; 深度学习; 生成模型; 对抗生成网络

中图分类号: TP391 文献标志码: A DOI: 10.3788/IRLA201847.0203005

Texture synthesis method based on generative adversarial networks

Yu Siqian^{1,2}, Han Zhi², Tang Yandong^{1,2}, Wu Chengdong¹

(1. School of Information Science and Engineering, Northeastern University, Shenyang 110000, China;
2. State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110000, China)

Abstract: Texture synthesis is a hot research topic in the fields of computer graphics, vision, and image processing. Traditional texture synthesis methods are generally achieved by extracting effective feature patterns or statistics and generating random images under the constraint of the feature information. Generative adversarial networks (GANs) is a new type of deep network. It can randomly generate new data of the same distribution as the observed data by training generator and discriminator in an adversarial learning mechanism. Inspired by this point, a texture synthesis method based on GANs was proposed. The advantage of the algorithm was that it could generate more realistic texture images without iteration; the generated images were visually consistent with the observed texture image and also had randomness. A series of experiments for random texture and structured texture synthesis verify the effectiveness of the proposed algorithm.

Key words: texture synthesis; deep learning; generative model; generative adversarial networks

收稿日期: 2017-09-05; 修订日期: 2017-10-05

基金项目: 国家自然科学基金(61773367, 61303168); 中国科学院青年创新促进会(2016183)

作者简介: 余思泉(1988-), 男, 博士生, 主要从事图像中层表达、深度学习等方面的研究。Email: yusiquan@sia.cn

导师简介: 唐延东(1962-), 男, 研究员, 博士生导师, 博士, 主要从事图像处理、目标识别方面的研究。Email: ytang@sia.cn

0 引言

纹理合成是指利用人工合成的方法生成视觉上具有真实感的纹理图像^[1]。作为一种重要的数字图像修补方法,多年来一直是计算机视觉中一个重要的研究课题。现有纹理合成方法大致可分两大类:参数化方法和非参数化方法。其中,参数化方法有纹理样图法,非参数化方法有基于马尔可夫随机场的方法和基于深度学习的纹理合成方法。

在数字图像中,具有模式相似的图像被定义为纹理图像。由于纹理图像的基本特征是重复性和周期性。所以,基于样本块的合成方法是纹理合成的最基本方法。Criminisi 等人^[2]于 2003 年提出了一种比较样本块之间特征差异的纹理合成方法。该方法从全局特征特征出发分析图像,对于由小样本块合成较大尺寸的纹理图像,具有较好的合成效果。Kwatra 等人^[3]将 Graph-Cut 方法引入到纹理合成中,并利用这个原理最小化合成图像块之间的不一致性,取得良好的合成效果。该方法的缺点是合成速度慢。Lefebvre 等人^[4]提出了在像素级上的可控制的并行纹理合成方法,并将该算法在 GPU 上实现,极大的提高了纹理合成速度。张伟伟等^[5]提出了一种自适应选取样本块大小的纹理合成方法。该方法主要利用灰度共生矩阵提取待修复图像的纹理特征,通过比较相关纹理特征和最优样本块的大小的相互关系,自动求取样本块尺寸,进而对传统算法进行了改进。

随着研究的深入,马尔可夫随机场法(Markov Random Field, MRF) 被引入到图像纹理合成的问题中。MRF 主要用于纹理图像建模,其基本思想是纹理图像都具有局部统计相似特征,局部区域的纹理特征由其邻域决定,且跟其具体位置无关。Zhu 等人^[6]将纹理图像统计量, MRF 模型和最大熵准则相结合,提出了 FRAME 模型。该模型比一般 MRF 纹理模型具有更强的描述能力。Kwatra 等人^[7]基于 MRF 的迭代优化纹理合成方法,它将纹理图像视为一个整体,通过定义纹理的能量函数并用最大期望算法实现纹理合成。该方法可以很好地保持纹理图像的全局结构,但有时会因为局部最优而导致合成图像局部重叠区域的不一致性。Urs 等人^[8]提出一种通过建立二维 MRF 模型和邻域搜索法合成纹理。

MRF 模型的优点是可以很好的表征和描述纹理的局部特征,缺点是很难体现具有一定结构的纹理特征的完整性和连贯性。这可能会导致合成纹理出现结构性错误,相邻块的边界处产生不连续现象,使得生成的图像缺乏真实性。

近年来,深度学习理论的不断发展和完善,涌现出了许多基于卷积神经网络的纹理合成的新方法。Xie^[9]和 Lu^[10]利用卷积神经网络和纹理图像的 FRAME 模型相结合合成纹理图像。传统的 FRAME 模型的构建是利用固定的滤波器。上述两篇文献利用在大数据集上训练的卷积神经网络模型的卷积层代替固定滤波器构建 FRAME 模型,生成更加真实纹理图像。Gatys^[11]和 Schreier^[12]利用训练好用于识别的卷积神经网络的特征图构建 Gram 矩阵作为纹理图像特征,这是纹理图像的一种统计特征,通过最小化目标图像和生成图像特征之间的差别以及梯度下降法最终实现生成图像与目标图像具有相同的统计特征,进而达到纹理合成的目的。

综上所述,纹理合成问题对合成结果有如下基本要求^[13-14]:

- (1) 新生成的纹理样本在视觉上类似于样本纹理,在结构上要求连续平滑;
- (2) 生成纹理图像应具有随机性,即在结构上生成样本与输入样本应在视觉上保持一致,同时又不能存在明显的重复感。

此外,由于现存大部分方法合成一张图像都需要经过采样或多次迭代优化,导致往往需要很长时间才能生成一幅图像。这也是纹理合成问题研究的一个重点问题。

自 2014 年以来,对抗神经网络这种训练生成模型的方法在许多图像生成问题上取得了成功。针对纹理合成所存在的问题,文中提出了一种基于全卷积对抗生成网络的纹理合成算法。该算法对传统纹理合成方法有如下改进:

- (1) 在满足感官上相近的同时具有一定随机性,生成的图像是一张新的图像而非原图像的采样和拼接;
- (2) 合成效率高。当网络训练好后,只需要输入随机噪声向量就可以生成纹理图像,而不用像其他深度学习算法那样需要多次迭代;
- (3) 利用全卷积对抗生成神经网络实现纹理合

成。由于现有对抗生成网络中有下采样和上采样层的存在,这就使得当网络生成较大尺寸的图像时难以收敛。全卷积网络的优点是能够训练所有网络参数,有利于训练过程收敛。

1 对抗生成网络生成纹理图像原理

1.1 网络结构

对抗生成网络启发自博弈论中的二人零和博弈,该模型由 Goodfellow 于 2014 年首先提出^[15]。一个典型的对抗生成网络由生成器 G 和判别器 D 两部分组成。其中,生成器 G 的作用是学习样本数据的分布,将一个输入的随机噪声 z 映射成为一张图像。判别器 D 作用是区分输入的图像是真实图像还是生成器生成的伪图像。整个网络的优化过程是一个二元极小极大博弈问题,即生成网络的训练目标生成更加真实的图像,判别网络的训练目标是尽

量判别生成图像为伪图像。利用一种对抗机制进行训练,当训练生成器时,固定判别器参数。当训练判别器时,固定生成器参数。训练过程交替迭代这两个操作,生成器和判别器的训练目标都是使对方错误最大化。最终,生成器会生成更接近真实样本数据的分布,判别器也将具备更好的判别能力。一般地,当 D 的判别能力提升到一定程度,并且无法正确判别数据来源时,可以认为这个生成器 G 已经学到了真实的数据分布。

文中构造了一个生成器和判别器均为五层全卷积神经网络的对抗生成网络,网络结构如图 1 所示。其中,生成器用于合成纹理图像,判别器输入为生成器的输出和数据库中某一合成的纹理样本的图像,用于比较二者差别并对整个网络参数进行更新。当网络训练好后,生成器输出图像将会与样本图像在视觉上一致。

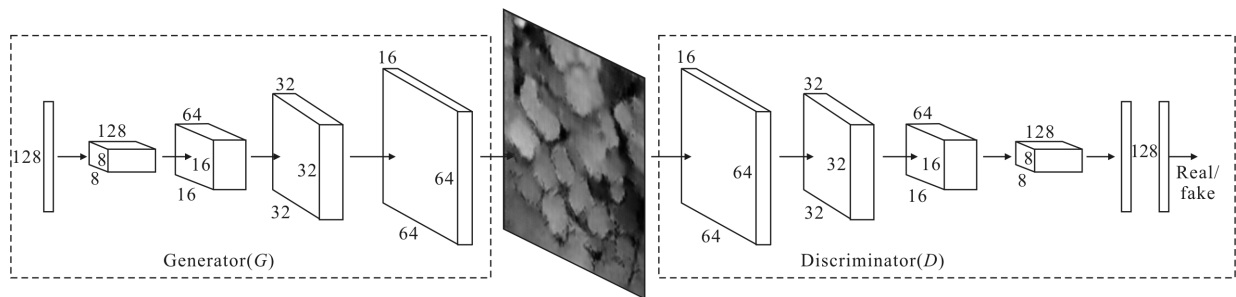


图 1 网络结构

Fig.1 Network architecture

1.2 网络训练过程

对抗生成网络的训练过程是一个交替优化生成器和判别器的过程。首先,固定生成器 G ,训练判别器 D 。此时,网络等价于一个二分类模型,模型损失函数为交叉熵,定义如下:

$$V(G, D) = E_{x \sim p_{\text{data}}} [\log D(x)] + E_{z \sim p_G} [\log(1 - D(z))] \quad (1)$$

式中: x 为对输入数据分布的采样; z 为对于给定数据分布的采样; $E(\cdot)$ 为期望。通过损失函数可知,与一般二分类问题不同,判别器的输入数据来源于真实数据和生成数据。给定生成器 G ,优化 D ,优化目标为使公式(1)最大化。

在连续空间上,上述损失函数定义为:

$$V(G, D) = E_{x \sim p_{\text{data}}} + E_{z \sim p_G} \quad (2)$$

其中,

$$E_{x \sim p_{\text{data}}} = \int_x P_{\text{data}}(x) \log D(x) dx$$

$$E_{z \sim p_G} = \int_z P_G(z) \log(1 - D(z)) dz$$

生成器 G 的最优解为:

$$D^*(x) = \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_G(z)} \quad (3)$$

由于当判别器输入来源于真实数据时, $D(x)$ 应趋近于 1,当判别器输入来源于生成器时, $D(G(z))$ 应趋近于 0。而此时, G 的目标是使 $D(G(x))$ 趋近于 1。

由此可知,生成器的损失函数为:

$$G^* = \operatorname{argmin}_G \max_D V(G, D) \quad (4)$$

这样, GAN 的优化过程可以描述为一个极小-极大优化问题。当 $P_G(z)=P_{\text{data}}(x)$ 时, G 取得最优值。式中 $V(G, D)$ 衡量了生成数据与真实数据之间概率分布的差异, 判别器的训练目标是最大化这种差异, 此时判别器能够最大化判别数据来源于真实数据还是生成数据。生成器训练目标是 minimized 这个最大差异, 使生成数据的概率分布尽可能的与真实数据的概率分布一致。

对抗生成网络收敛的关键是需要同时训练生成网络 G 和判别网络 D 。在实际训练中, 为了防止梯度消失, 通常对 D 的参数更新 k 次再对 G 的参数更新 1 次。

2 实验

2.1 实验参数

该小节主要从定性角度验证提出的算法的有效性, 实验所用图像均来自网络。由于纹理图像数量较少, 而训练对抗生成网络需要大量样本, 文中首先对数据集进行了预处理。具体的方法是在原图像上进行对比度变化, 旋转, 亮度变化等。目的是扩大训练库样本数量。数据库中的图像如图 2 所示。

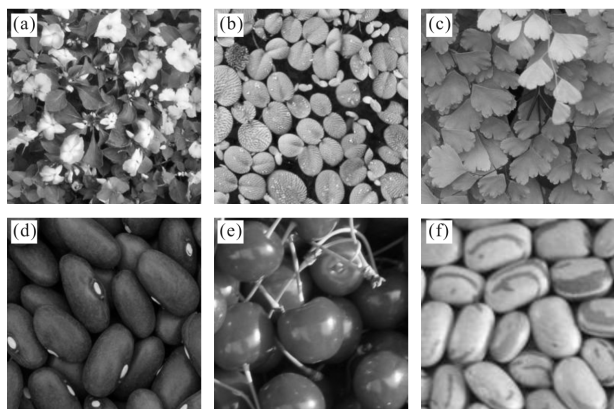


图 2 数据库部分图像

Fig.2 Some images from database

在实验过程中, 生成器为一个五层的反卷积神经网络, 每层均为 5×5 的卷积核, 卷积核的数量分别为 128、64、32、16 和 3。判别器由四层的卷积层, 一层全链接层和 softmax 层组成, 各卷积层的卷积核大小同样均为 5×5 , 卷积核的数量分别为 16、32、64、128。此外, 在卷积过程中, 每层卷积层的激活函数为 Relu。所有实验的实验环境为主频 3.6 GHz, 内存

26 GB, GPU 为 NVIDIA Titan X。网络采用 Tensorflow 框架编程实现。

2.2 合成结果

通常评价生成的纹理图像的方法是通过视觉上对生成样本进行评价。纹理图像可分类为随机纹理图像和结构性纹理图像, 文中分别生成这两种纹理以验证算法的普适性。实验中, 生成网络输入为一个 128 维的在 0~1 之间均匀分布的随机采样, 生成器输出为一幅 128×128 的图像。判别器输入为 128×128 的生成图像和数据库图像, 输出为对输入图像真伪的判别。

图 3 为随机纹理合成结果, 图 4 为结构性纹理合成结果。其中, 左侧一列为输入的训练图, 右侧 4 幅为合成图像。从实验结果可以看出, 文中提出的算法在两种纹理生成的问题上都取得了很好的效果。在结构性纹理的合成中, 从视觉上生成图像具有真实图像的基本结构, 同时又具备了训练图像细节。

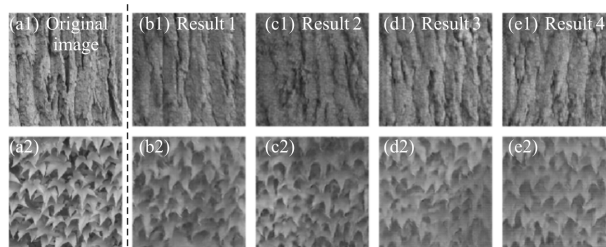


图 3 随机纹理合成结果

Fig.3 Synthesis results of random texture

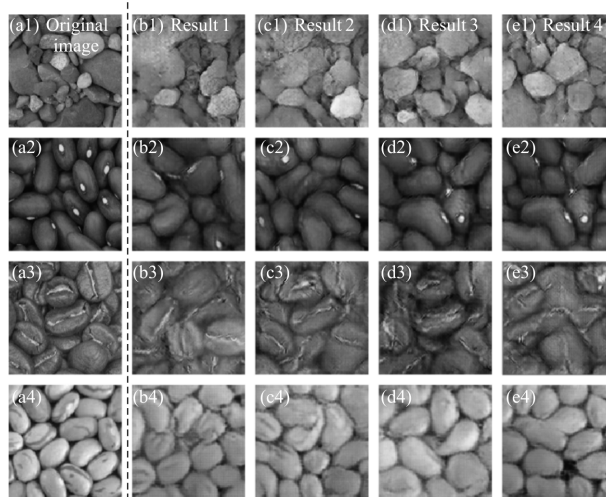


图 4 结构性纹理合成结果

Fig.4 Synthesis results of structured texture

2.3 合成结果与迭代次数关系

图 5 为合成结果随网络训练迭代次数增加的变

化情况。从结果可以看出,在网络训练的初始阶段,合成结果变化较为剧烈,随着网络的收敛,合成结果变化逐渐变得不明显,最后趋于稳定,得到了与原图像在视觉上相似,并具有一定随机性的纹理图像。

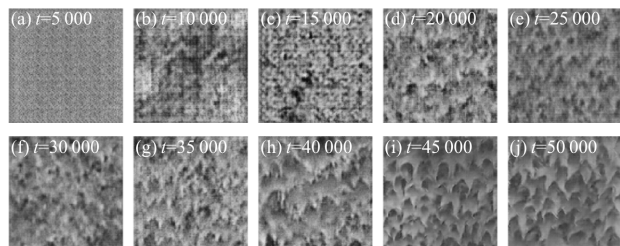


图 5 合成过程

Fig.5 Synthesis process

另外,还对不同迭代次数生成不同的纹理图像之间的差异进行了比较。计算差别率的公式定义为:

$$\text{Rate} = \frac{\sum |I_1 - I_2|}{\sum I_1} \quad (5)$$

式中: I_1 和 I_2 分别为原图像或不同训练迭代次数生成的图像。实验结果如表 1 所示,通过分析表 1 可以得出结论,对于不同迭代次数,文中提出的算法生成纹理图像相互间差异明显,具有一定的随机性。

表 1 不同合成结果的差别率

Tab.1 Difference rate of different synthesis results

No.	1	2	3	4	5
Difference rate	23.91%	29.33%	21.69%	24.71%	22.81%
1	0	5.94%	10.53%	15.63%	10.77%
2	21.82%	0	29.84%	27.52%	27.31%
3	16.39%	23.55%	0	28.84%	13.96%
4	17.86%	10.63%	22.17%	0	16.87%
5	15.56%	16.66%	16.65%	13.24%	0

2.4 与其他方法比较

将文中的纹理合成方法与其他合成方法进行比较,比较结果如图 6 所示。从实验结果可以看出,与参考文献[9,11]方法相比,文中提出方法具有更真实的合成结果。另外,还对各算法的合成时间进行了比较,比较结果如表 2 所示。从表中可以看出,虽然本文提出方法训练时间较长,但不同于其他两种方法是在训练过程中合成纹理,一旦合成其他种类的纹理需要重新训练。文中方法网络训练好后合成过

程无需训练,合成时间也明显少于其他两种方法。通过对合成图像质量和合成时间两方面比较,验证了算法的优越性。

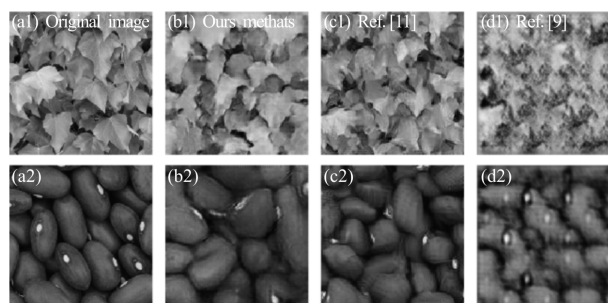


图 6 与其他方法比较

Fig.6 Comparison with other methods

表 2 合成时间比较

Tab.2 Comparison of synthesis time

Methods	Training time	Testing time/s
CNN+Graw ^[11]	About 15 min	-
CNN+FRAME ^[9]	About 20 min	-
Our methods	About 10 h	0.2

3 结论

文中针对纹理合成要求合成结果应具备一定随机性和视觉上与真实纹理一致两个方面,设计了一种基于对抗生成网络的纹理合成算法。该算法主要由一个生成器和一个判别器组成。生成器主要负责合成纹理图像,判别器的作用是不断驱动生成器合成与给定图像在视觉上相同的图像。整个网络利用生成器与判别器相互促进交替训练的方式进行训练。通过合成随机纹理和结构性纹理验证了算法的有效性。

利用对抗生成网络进行纹理合成是一种有效的纹理合成算法,该算法所合成的图像保证了视觉上与输入纹理相似的同时,又能具有一定随机性。与传统算法相比,文中提出的纹理合成算法不只是对给定图像的简单采样和拼接,而是生成了新的纹理图像。结合纹理合成问题的本质,设计一个网络合成多种不同类别的纹理图像将是未来有意义的研究方向。

参考文献:

- [1] Huang Jun, Li Feng, Gui Yan, et al. Surfaces texture synthesis based on texel distribution [J]. *Journal of Chinese Computer Systems*, 2016, 37(10): 2361–2365. (in Chinese)
- [2] Criminisi A, Perez P, Toyama K. Region filling and object removal by exemplar-based image inpainting [J]. *IEEE Transactions*, 2004, 13(9): 1200–1212.
- [3] Kwatra V, Essa I, Turk G, et al. Graphcut textures: Image and video synthesis using graph cuts [C]//ACM Transactions on Graphics, 2003: 277–286.
- [4] Lefebvre S, Hoppe H. Parallel controllable texture synthesis [C]//ACM Transactions on Graphics, 2005, 24(3): 777–786.
- [5] Zhang Weiwei, He Kai, Meng Chunzhi. Texture synthesis method by adaptive selecting size of patches [J]. *Computer Engineering and Applications*, 2012, 48(17): 170–173. (in Chinese)
- [6] Song C Z, Wu Y, Mumford D. Filters, random fields and maximum entropy (FRAME): towards a unified theory for texture modeling [J]. *International Journal of Computer Vision*, 1998, 27(2): 107–126.
- [7] Kwatra V, Essa I, Bobick A, et al. Texture optimization for example-based synthesis[C]//ACM Transactions on Graphics, 2005: 795–802.
- [8] Urs R D, Costa J P D, Germain C. Maximum-likelihood based synthesis of volumetric textures from a 2D sample[J]. *IEEE Transactions on Image Processing*, 2014, 23(4): 1820–1830.
- [9] Xie J, Hu W, Zhu S C, et al. Learning sparse FRAME models for natural image patterns [J]. *International Journal of Computer Vision*, 2015, 114(2–3):1–22.
- [10] Lu Y, Zhu S C, Wu Y N. Learning FRAME models using CNN filters[J]. *Computer Science*, 2015, arxiv: 1509.08379v3.
- [11] Gatys L A, Ecker A S, Bethge M. Texture synthesis using convolutional neural networks[J]. *Febs Letters*, 2015, 70(1): 51–55.
- [12] Schreiber S, Geldenhuys J, Villiers H D. Texture synthesis using convolutional neural networks with long-range consistency and spectral constraints [C]//Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference. IEEE, 2017:1–6.
- [13] Xiao Chunxia, Huang Zhiyong, Nie Yongwei, et al. Global texture optimization incorporating with image detail [J]. *Chinese Journal of Computers*, 2009, 32 (6): 1196–1205. (in Chinese)
- [14] Tang Ying, Lin Qifeng, Xiao Tingzhe, et al. GPU-based texture synthesis with preserved structures [J]. *Computer Science*, 2016, 43(4): 299–302. (in Chinese)
- [15] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C]//International Conference on Neural Information Processing Systems, 2014: 2672–2680.