

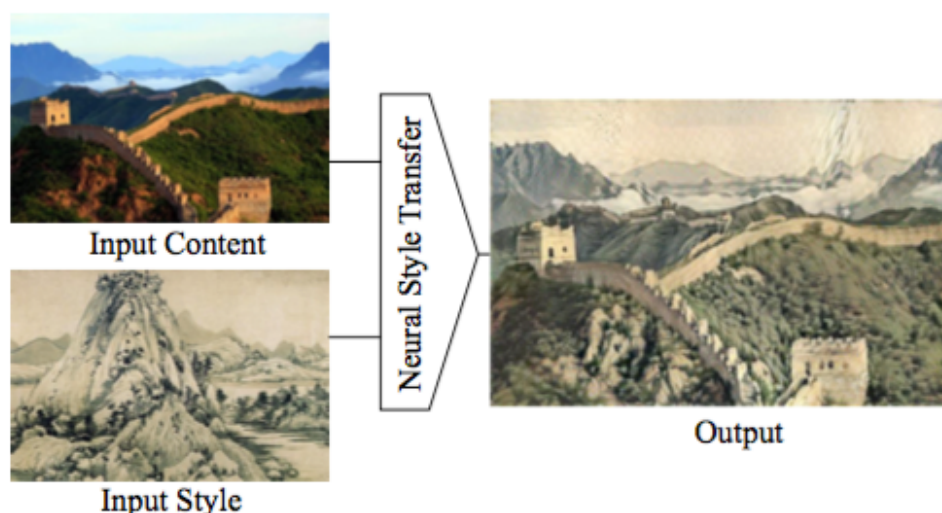
风格迁移综述

21821128 郭兆宇

一、背景

绘画是几千年来最有魅力最吸引人的艺术形式之一，不同的绘画艺术风格各异，每种风格都曾涌现出许多艺术大师，想要将同样的主题以不同的风格进行绘制需要训练有素的艺术家和大量的时间。

近年来，由深度学习所引领的人工智能技术浪潮，开始越来越广泛地应用到社会各个领域。其中人工智能与艺术的交叉碰撞，不仅在相关的技术领域和艺术领域引起了高度关注，各种图像处理软件和滤镜应用更是吸引了海量用户，在这其中很重要的技术基础就是图像风格迁移。



二、发展历史

在神经网络之前，图像风格迁移的方法有一个共同的思路：分析某一种风格的图像，给这种风格建立一个数学或者统计模型，再改变要做迁移的图像，让原图像能更好的符合建立的模型。在这些研究中，非真实感绘制（NPR）已成为计算机图形学界成熟的研究方法，然而大多数 NPR 风格化算法都是为特定的艺术风格而设计的，不能轻易地扩展到其他风格。Hertzmann 等人进一步提出了图像类比的框架，通过从提供的未分类和样式化图像的示例对中学习类似转换来执行一般化的样式转换。然而这些方法的共同缺点是它们只使用低阶图像特征，往往无法有效地捕获图像结构。

受到卷积神经网络（CNN）的启发，Gatys 等人首先研究了如何使用 CNN 在自然图像上重现著名的绘画风格，他们通过将照片的内容建模为接受过预先培训的 CNN 的特征响应，并进一步将艺术品的风格建模为摘要特征统计。实验结果表明，CNN 能够从照片中提取内容信息，并从一幅著名的艺术品中提取风格信息将两者融合。Gatys 等人的开创性著作，展示了卷积神经网络（CNN）通过分离和重组图像内容和风格来构造新图像的能力。使用 CNN 呈现不同风格的内容图像的过程称为神经风格转换（NST）。

三、无神经网络的风格转换

基于笔划的渲染

基于笔划的渲染 (SBR) 是指通过虚拟笔划在数字画布上绘制具有特定样式的图像的过程。SBR 的制作过程一般是从一张源照片开始, 通过增加笔画的构图来匹配照片, 最后制作出一种非真实感的图像, 具有特定的艺术风格。因此, 它们通常能够有效地模拟某些类型的风格 (例如油画、水彩、素描), 但是每个 SBR 算法都是为一个特定的风格设计的, 不能灵活地切换其他风格。

基于区域的渲染

基于区域的渲染是将区域分割结合起来, 使渲染能够根据区域中的内容进行自适应。通过这种方式, 不同的笔画模式可以在图像的不同语义区域中产生, Song 等人进一步提出一种基于区域的 ib-ar 算法来处理艺术风格的几何图形, 他们的算法通过用几个标准形状替换区域来创建简化的形状呈现效果。

基于示例的渲染。

基于示例的渲染的目标是学习图像对之间的映射, Hertzmann 等人提出了一个名为“图像类比”的框架, 图像类比旨在以有监督的方式学习源图像和目标样式化图像之间的映射。图像模拟算法从示例训练对中学习类似的转换, 并在给定测试输入照片时创建类似的风格化结果。

四、神经风格迁移算法的分类

基于图像优化的在线神经方法 (IOB-NST)

DeepStream 尝试通过使用 IOB-IR 技术反转 CNN 表示来进行风格迁移, 通过进一步将视觉纹理建模技术与模型风格相结合, 提出了 IOB-NST 算法, 为 NST 领域的研究打下了基础。他们的基本思想是首先对相应的样式和内容图像进行建模和提取样式和内容信息, 然后将其重新组合为目标表示, 然后迭代地重建与目标表示匹配的样式化结果。一般来说, 不同的 iob-nst 算法的 iob-ir 技术基本重叠, 但它们对视觉样式建模的方式不同, 视觉样式是建立在上述两类视觉纹理建模技术之上的。

基于离线神经方法的模型优化 (MOB-NST)

IOB-NST 的效率问题成为这类神经风格迁移算法的主要局限性, MOB-NST 通过利用 MOB-IR 重建风格化结果来解决速度和计算成本问题, 对于一个或多个风格的图像, 在一组图像 IC 上优化前馈网络 G: 根据一个 G 可以产生的艺术风格的数量, MOB-NST 算法进一步分为每种风格每模型 (PSPM) MOB-NST 方法、多样式每模型 (MSPM) MOB-NST 方法和任意样式每模型 (ASPM) MOB-NST 方法。

(PSPM) MOB-NST 方法

Johnson 等人首先提出了 MOB-NST 算法, 预先训练一个前馈式特定网络, 并在测试阶段用一次前馈通过产生一个风格化的结果。网络架构设计大致遵循 Radford 等人提出的网络, 但对于残余块体以及分段跨步式进行改动。

(MSPM) MOB-NST 方法

尽管上述 PSPM 方法可以比以前的 IOB-NST 方法更快地生成两个数量级的风格化图像，但是必须为每个特定的风格图像训练单独的生成网络，这是非常耗时和不灵活的。但许多绘画都有相似的笔触，只是在色彩调色板上有所不同。因此，John 等人提出了 MSPM，进一步将多种样式合并到一个模型中，从而提高了 PSPM 的灵活性。

(ASPM) MOB-NST 方法

ASPM-MOB-NST 旨在建立单一的可训练模型，以传递任意的艺术风格，同样还有两种类型的 ASPM，一种是基于 MRF 的非参数纹理建模，另一种是基于总和统计的参数纹理建模。

改进和扩展

自 NST 算法出现以来，也有一些研究致力于通过控制感知因素来改进当前的 NST 算法。此外，上述所有的 NST 方法都是针对一般静止图像进行设计的，不适用于专门类型的图像和视频。因此，各种后续研究旨在将一般 NST 算法扩展到这些特定类型的图像，甚至扩展到艺术图像样式（例如音频样式）之外。

语义样式转换：给定一对内容相似的样式和内容图像，语义样式转换的目标是在样式和内容之间建立语义对应关系，将每个样式区域映射到相应的语义相似的内容区域。然后将每个样式区域中的样式传输到语义相似的内容区域。

实例/样式传输：实例样式转换是建立在实例分割的基础上的，其目的是使图像中只有一个用户指定的对象具有样式。挑战主要在于风格化对象与非风格化背景之间的转换。Castillo 等人通过为平滑和反混叠边界像素添加额外的基于 MRF 的损失来解决此问题。

涂鸦风格转换：涂鸦风格转换是利用 NST 将草图转化为精美的艺术品的研究，该方法简单地抛弃了内容丢失项，采用涂鸦作为分割图进行语义样式转换。

立体风格转换：在 AR/VR 需求的驱动下，Chen 等人提出立体图像的立体 NST 算法，该算法可以为不同的视图生成更一致的笔画。

五、实验与结果

本综述挑选文章《Perceptual Losses for Real-Time Style Transfer and Super-Resolution》参考代码 https://github.com/Heumi/Fast_Multi_Style_Transfer-tensorflow 进行了复现，文章作者对样式传输网络的简单修改，即引入条件实例规范化，使其能够学习多种样式。并且证明了这种方法的灵活性，在定性和收敛性方面都可以与单用途类型的传输网络相媲美，并且证明了嵌入的空间表现允许人们以之前未观察到的新颖方式任意组合艺术风格，实验结果如下图。



六、应用

社会交往

NST 之所以在学术界和业界引起关注，主要原因是它在一些社交网站（如 facebook 和 twitter）中很受欢迎。另外一些移动应用程序（如 prisma）是第一个将 NST 算法作为服务提供的工业应用程序。由于其高的造型质量，Prisma 获得了巨大的成功，并正在世界各地流行，一些提供相同服务的其他应用程序相继出现并开始提供商业利益。在这些工业应用程序的帮助下，人们可以创作自己的艺术作品，并在 Twitter 和 Facebook 上与他人分享自己的艺术作品，这是一种新的社交形式。

用户辅助创建工具

NST 的另一个用途是使其充当用户辅助创建工具，NST 作为画家和设计师的创作工具，可以使画家更方便地创作出特定风格的艺术作品，尤其是在创作电脑制作的艺术作品时。此外，使用 NST 算法，为时装设计师生成风格化的时装元素和各种风格的建筑师生成风格化的 CAD 绘图是简单易操作的。

娱乐应用的生产工具

一些娱乐应用程序，如电影、动画和游戏，是 NST 最常用的应用形式。例如，创建动画通常需要每秒 8 到 24 帧的绘制帧，如果能应用 NST 自动将实况动作视频风格化为动画风格，则生产成本将大大降低。同样，NST 在制作一些电影和电脑游戏时可以显著节省时间和成本。

参考文献

- [1]. Dumoulin V , Shlens J , Kudlur M . A Learned Representation For Artistic Style[J]. 2016.
- [2]. Johnson J , Alahi A , Fei-Fei L . Perceptual Losses for Real-Time Style Transfer and Super-Resolution[J]. 2016.
- [3]. L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2414–2423.
- [4]. O. Frigo, N. Sabater, J. Delon, and P. Hellier, "Split and match: Example-based adaptive patch sampling for unsupervised style transfer," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 553–561.
- [5]. M. Elad and P. Milanfar, "Style transfer via texture synthesis," IEEE Transactions on Image Processing, vol. 26, no. 5, pp. 2338–2351, 2017.
- [6]. L. A. Gatys, A. S. Ecker, and M. Bethge, "A neural algorithm of artistic style," ArXiv e-prints, Aug. 2015.
- [7]. I. Prisma Labs, "Prisma: Turn memories into art using artificial intelligence," 2016. [Online]. Available: <http://prisma-ai.com>
- [8]. A. Semmo, T. Isenberg, and J. Do'llner, "Neural style transfer: A paradigm shift for image-based artistic rendering?" in Proceedings of the Symposium on Non-Photorealistic Animation and Rendering. ACM, 2017, pp. 5:1–5:13.
- [10]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [11]. G. Berger and R. Memisevic, "Incorporating long-range consistency in cnn-based texture generation," in International Conference on Learning Representations, 2017.
- [12]. A. Dosovitskiy and T. Brox, "Inverting visual representations with convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4829–4837.
- [13]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
- [14]. Y. Li, N. Wang, J. Liu, and X. Hou, "Demystifying neural style transfer," in Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 2017, pp. 2230–2236. [Online]. Available: <https://doi.org/10.24963/ijcai.2017/310>
- [15]. J. Ye, Z. Feng, Y. Jing, and M. Song, "Finer-net: Cascaded human parsing with hierarchical granularity," in Proceedings of the IEEE International Conference on Multimedia and Expo. IEEE, 2018, pp.
- [16]. E. Risser, P. Wilmot, and C. Barnes, "Stable and controllable neural texture synthesis and style transfer using histogram losses," 1–6.
- [18]. ArXiv e-prints, Jan. 2017.
- [19]. S. Li, X. Xu, L. Nie, and T.-S. Chua, "Laplacian-steered neural style transfer," in Proceedings of the 2017 ACM on Multimedia Conference. ACM, 2017, pp. 1716–1724.
- [22]. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," ArXiv e-prints, Nov. 2015.