



图像艺术风格化的研究现状

摘要

图像艺术风格化作为一个正在蓬勃发展的领域被越来越多的人熟知,也引起了众多学者的研究兴趣.本文总结了图片风格化的发展现状,分析了不同风格化方法的特点,指出了目前风格化方法的缺点,总结了图片风格化的发展趋势,为进一步研究图片风格化提供了方向.

关键词

图像风格化;笔触;生成模型

中图分类号 TP391.41;TP18

文献标志码 A

1 图像艺术风格化概念

随着图片处理技术的发展,人们对图片美化有了更多的想法和要求.例如,人们希望自己拍摄的图像具有艺术大师画作的风格,如图1所示的梵高的星空等的效果.目前,PhotoShop等专业的图像处理软件,也可以用于图片的艺术风格改造.但对于普通的用户来说,PS并不是一个简单快捷的工具,而且采用人工创作必定会消耗大量的人力、时间、金钱,因此研究人员提出了风格化的概念.图像风格化又可以称为风格迁移,是指让计算机自动完成图像风格的转移,将一张具有艺术特色的图像的风格迁移到一张自然图像上,使原自然图像保留原始内容的同时具有独特的艺术风格,如卡通、漫画、油画、水彩、水墨等风格.

对于图像风格化的研究,使照片编辑、平面设计和动漫制作等领域有了重大突破,可以帮助非专业用户自由创作出带有艺术风格的作品.

2 传统图像艺术风格化方法

传统图像艺术风格化的方法可以分为3类:基于笔触的风格化、基于纹理合成的风格化和对物理过程建模的风格化.

基于笔触的方法在油画风格化中应用较多.油画中有一类风格笔触明显,可清晰看见笔触的宽度、长度和走向.Hertzmann^[1]使用多样条笔触合成图像,笔触根据原图像颜色进行选择,样条笔触方向与图像的梯度方向对齐.Zeng等^[2]提出利用图像的语义信息选择笔触和决定笔触位置方向的方法,根据区域和物体的标识等信息从笔触库中选择合适的笔触,依据方向场进行绘制.Luo等^[3]使用条件随机场优化框架,结合从输入图像中提取的法向图和获得光照估计方向后由法向图和反射率图组成的训练集,合成新的具有浮雕风格的法向图.Lu等^[4]提出了一种不同风格笔触的生成方法,处理笔触的拖尾、交叉和重叠时颜色混合的问题.

有些研究工作将图像风格化转化为纹理合成问题.Wang等^[5]由人工从参考图像样例中提取典型的笔触用于合成图像,因此每次生成的图像和用户的采集相关.算法思想是先将图像进行分割,计算每个分割块的方向场,然后使用纹理块的层次结构和图像模板合成不同分割块的方向纹理.Brooks^[6]根据参考图像中的颜色和纹理信息把

收稿日期 2017-08-28

资助项目 国家自然科学基金(61672520,61501464,61702488);北京市自然科学基金(4162056)

作者简介

邓盈盈,女,硕士生,研究方向为多媒体计算.dengyingying2017@ia.ac.cn

董未名(通信作者),男,博士,研究员,研究方向为多媒体计算.weiming.dong@ia.ac.cn

1 中国科学院自动化研究所 模式识别国家重点实验室,北京,100190

2 中国科学院大学,北京,100049



图1 梵高星空风格图像风格化效果

Fig. 1 Image stylization results by using van Gogh's starry night style

输入图像变为花窗玻璃风格的图像. Bousseau 等^[7]提出双向纹理平流方法来实现视频的水彩画,对整个视频在相反方向上提取纹理平流场,并在像素级别上优化两者的组合. Frigo 等^[8]提出了一种基于 patch 分割重构的非监督图像风格化方法. 首先,运用四叉树对原图进行分割,然后找到风格图片中与之匹配的 patch;其次,运用双线性混合消除缝线,最后进行整体的颜色迁移,最终得到风格化后的图片.

基于建模的方式也被应用于图像风格化. Baxter 等^[9]利用毛笔书写时的变形数据,对毛笔的绘制过程进行建模. Wang 等^[10]对彩色墨水的扩散进行建模,将水和墨的扩散分开考虑,使用 KM 理论处理像素混合颜色,而且考虑了纸的特性和重力因素. Zhang 等^[11]使用马尔科夫随机场将图像风格传递过程规划为一个整体优化问题,并使用置信传播算法进行求解. Wang 等^[12]通过联合使用颜色传递、基于突出度的多层次绘制、手颤模拟以及湿画法模拟等技术将图像合成为水彩画的效果,该方法仍然无法自由控制水彩的具体风格. Kim 等^[13]提出了创建双目立体数字画作的方法. Chen 等^[14]构建了可模拟画刷、油彩和画布之间交互的实时绘画系统,方便用户自由创作数字艺术作品.

3 基于生成模型的图像艺术风格化

近些年来,“人工智能”一词越来越频繁地出现在大众面前.众所周知,人工智能就是让机器完成原来只有人类才能完成的任务.人工智能的核心是算法,而深度学习是形成算法的重要过程.深度学习的目的是模拟大脑神经感知外部世界,因此对于深度学习的研究重点一直放在人工神经网络方面.而神经网络的研究使得语音、图像处理领域有了长足发展.一些学者将神经网络运用到图像风格化上来,用卷积的方法提取图像在神经网络各个层的内容和风

格信息.目前基于生成模型的图像艺术风格化方法是主要的研究热点. Gatys 等^[15]利用卷积神经网络提取图像不同层级的特征,使用低层次响应描述图像的风格,使用高层次响应描述图像的内容,使用梯度下降调整输入响应,在特定层次获得特定的响应,经多次迭代之后,输入响应即为特定风格和内容的图像,其效果如图 2 所示.该方法的缺点是速度很慢. Johnson 等^[16]提出了基于感知损失函数的实时风格转换方法,通过使用感知损失替代像素损失,使用预训练的 VGG 模型简化损失计算,并增加一个转换网络(网络结构如图 3 所示),从而直接生成风格化图像,大大提高了训练速度,缩短图片产生的时间. Li 等^[17]提出了马尔科夫生成对抗网络用于训练生成网络(网络结构如图 4 所示),对抗生成网络是最近比较火的网络结构之一.生成对抗网络包含 2 个部分,对抗网络尽可能判断出输入的图片是生成网络产生的还是一张真实的图片,生成网络使产生的结果尽可能逼真,令对抗网络信以为真.2 个网络相互竞争,使两者都达到最好的效果.对抗网络用于辨别从合成图像取样神经碎片是否为真,对于每一个神经碎片它输出分类为 s (1 表示为真),然后最小化纹理损失公式,通过损失反转使纹理损失最小.再由生成网络将原图转换成一张风格化后的图片,从而实现高效纹理合成图像艺术风格化. Frigo 等^[18]将复杂的图像风格概念分解为局部纹理迁移和全局颜色迁移,以此构建了一个无监督的图像风格化方法. Ulyanov 等^[19]提出了一种前馈卷积神经网络,产生大量任意尺寸的相同纹理,将给定照片的艺术风格迁移到任意图片上,该方法引入一个用于风格转移的生成网络,大大减少迭代的次数,提高了训练速度,缩短图片生成的时间. Liao 等^[20]提出了一种在具有相似语义结构和较大外观差异的图像之间进行视觉属性迁移的方法.该方法利用从深度卷积神经网

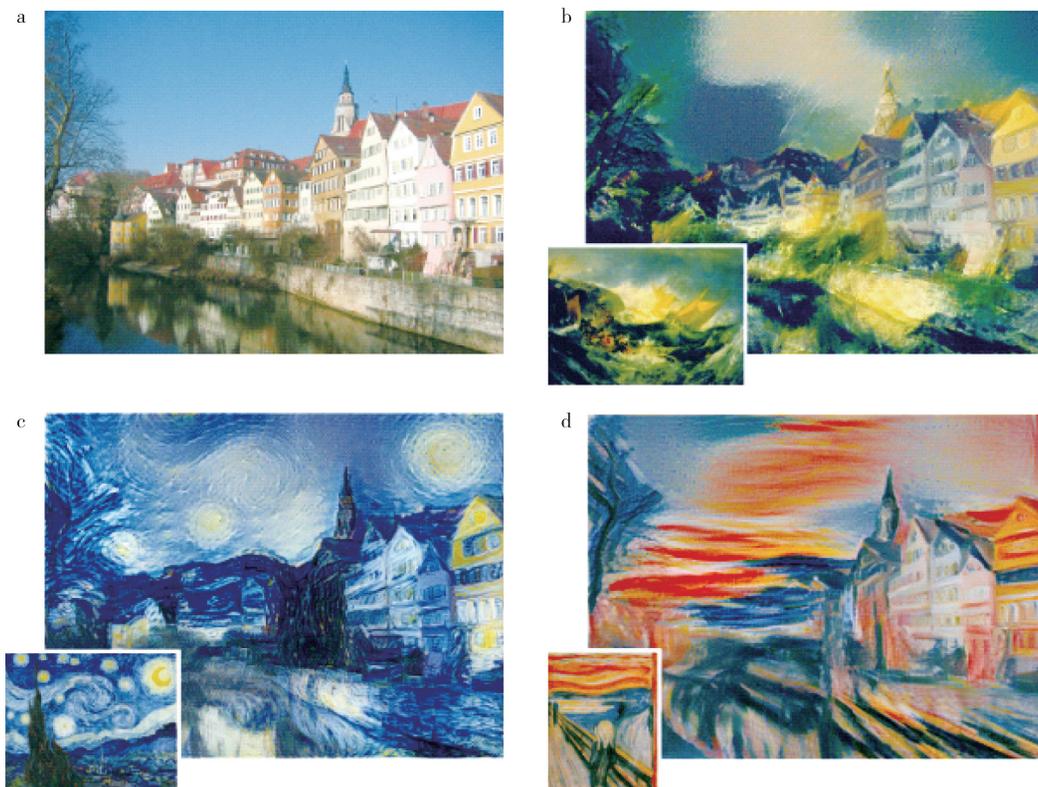


图2 Gatys等^[15]的风格化结果

Fig. 2 Image stylization results using method presented by Gatys et al^[15]

络提取的特征在2张输入图像之间找到语义上有意义的稠密对应关系来构建新的“图像类比”方法,进而使用一种由粗到细的策略来建立最近邻域并生成风格化结果.Zhu等^[21]提出了一种在非配对图像集之间进行转换的方法.该方法通过引进 Cycle Consistency 的概念,在原本的生成式对抗网络增加一个 generator,将产生的目标图像再转回原本的源领域,并且限制转换回来的图像要与源图像越接近越好,如此便可在目标域产生对应的图像.然而,该方法仍然需要在2个域分享类似的语义内容才能成功.Chen等^[22]给图像风格提供了一个明显的表示,使得网络在训练好之后可以从内容中完全分离出样式.该方法不仅可以同时训练多个共享自编码的风格,还可以在不改变自编码的情况下,增量学习一个新的风格.Dumoulin等^[23]在文献[16]的基础上提出了更灵活的风格化模型,提出了条件实例正则化的方法来混合多种不同的画风,从而学习到风格的表示.这种风格表示学习的方法更有助于发现新的风格.

4 人像艺术风格化

人像艺术风格化是图像风格化中的一个重要研

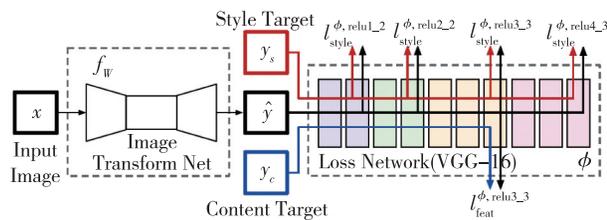
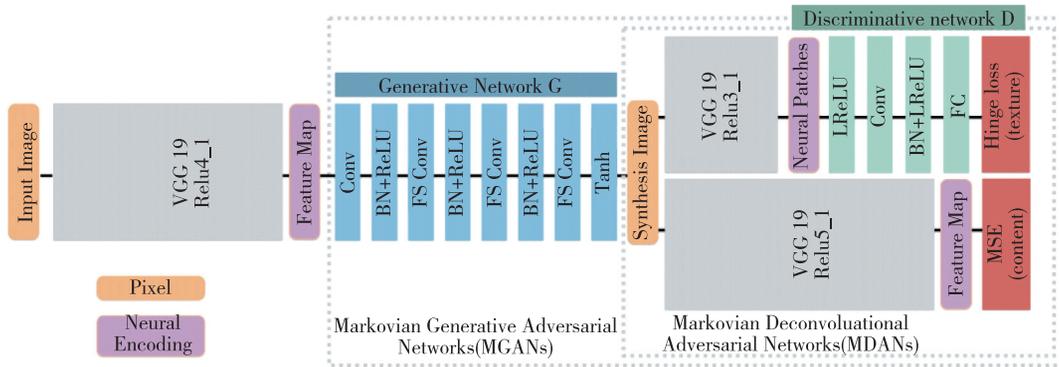


图3 基于感知损失函数的实时风格转换结构^[16]

Fig. 3 Real-time style transform network based on perceptual loss function^[16]

究问题,尽管 Gatys等^[15]的算法已可以进行通用图像的风格转移,但它还不适用于头像的风格转移.由于空间约束不强,直接应用 Gatys等^[15]的方法可能会使人物头部变形.对于这种类型的风格转移它是不可接受的.由于人的视觉系统对脸部细微的不正常都很敏感,自动的肖像绘画是一项巨大的挑战.很多研究工作采用统计学习模型人脸风格化的规则.Chen等^[24]采用非均匀的马尔可夫随机场描述简笔线条和真实人脸图像的统计关系,并使用非均匀无参采样技术进行求解.生成卡通画分为2个步骤,先生成描述线条的矢量图,然后使用已有的笔触合成

图4 Li等^[17]提出的马尔科夫生成对抗网络用于训练生成网络Fig. 4 Markov generative adversarial network to train generative network presented by Li et al^[17]

最后的图像,缺点是不同人像上的笔触单一,且头发部分并未实现自动化以及缺乏纹理信息.Chen等^[25]提出了基于五官拼贴的人脸卡通生成方法,其结果非常依赖于人脸五官轮廓线条的精度,且笔触渲染方法也很难推广到其他风格的人脸卡通画框架.Tseng等^[26]关注风格化人脸的夸张效果并同时保证生成人脸和输入人脸的相似性.Selim等^[27]提出了头像绘画技术,当前的技术不能很好捕捉画作的纹理,导致面部结构变形.基于图像类比限制了他们的领域的适用性.此技术使用一个卷积神经网络基于新颖的空间约束的方法.这种方法在转移绘画风格的同时可以保持面部结构的完整性,其风格化结果如图5所示.Zhao等^[28]提出了一个基于样本的从照片生成肖像绘画的方法,采用的方法主要是从预先由艺术家绘制的肖像模板中传递笔触.Wang等^[29]同样从样本中迁移笔触,但他们学习了未绘制的样本图像和笔触属性之间的关系.Zhang等^[30]提出了一个数据驱动的人像卡通化生成框架,给定一幅人脸图像,可生成多种风格的卡通头像.卡通头像生成分为2个阶段:人脸解析和卡通化生成.人脸解析包括人脸检测、人脸配准、性别检测、眼镜检测等.利用人脸检测和人脸配准技术,定位人脸图像中的器官,提取器官并对齐.对输入图像中的每个真实器官,从建立的数据库中优化选择相应的卡通器官,并通过学习艺术家拼贴卡通头像的策略对其进行优化调整,从而组合生成卡通头像.

近年来人脸素描合成也得到了较多关注,比较常用的思想是利用马尔可夫随机场的平滑性将训练集中的图像和素描对分成重叠的图像块进行合成,这些基于图像块的方法在人脸素描中得到广泛应用得益于其局部人脸特征的表达能力^[31].Wang等^[32]

采用马尔可夫随机场进行建模,把真实人脸的小块作为隐节点,把可能的素描块作为显节点.Li等^[33]通过融合从特定训练集中获取的引导图像内容进行卡通画人脸合成,但其使用的映射规则是否有效和能否取得高质量的合成结果无法论证.Wang等^[34]同样采用马尔可夫随机场从训练集中选择最合适的近邻以合成目标图像块.Zhang等^[35]从图像块中学习特征词典并在搜索过程中将图像块替换为一组稀疏参数表示.Zhang等^[36]建了只通过训练单一模板来合成人脸素描的框架,使用多特征优化模型来选择备选图像块.基于马尔可夫随机场的方法的缺点是其忽略了人脸图像的全局信息,如描述人脸五官关系的几何信息,从而导致很多全局性夸张化的相关信息难以在合成图像中体现.使用加权平均的方法来处理图像块重叠区域的机制会导致结果中出现重叠变形或者过度平滑的现象.另外,该方法需要较长的计算时间和较大的存储空间,这也限制了其实际应用.

5 目前图像风格化存在问题

以上的风格化方法主要存在以下几个问题.首先,艺术风格化方法大多是基于单张图像,比如从单张图像中选择典型的笔触,在彩色玻璃中从单张图像中寻找颜色和纹理最近的分割块.然而,单张图像往往无法代表一种风格,使用单张图像来描述一类图像的共性是不合适的.其次,风格化中较少考虑图像的语义信息.大多数的方法是根据图像中提取的低层特征比如分割块的方向场决定笔触的位置,很少尝试对图像的语义信息进行建模以及考虑图像中的内容信息.再次,图像中的重要部分和非重要部分缺乏区分.风格化在一定程度上对图形进行了抽象,比

图5 基于空间约束的人脸风格化结果^[27]Fig.5 Spatial constraint-based face stylization results^[27]

如剪影、素描、优化等,需要考虑图像中不同部分的重要性.文献[30]做了这方面的尝试,在皮肤、人脸上使用比较小的笔触,在其他不太重要的风景上使用大的笔触,这是对图像中不同内容重要性进行了区分的例子.最后,以上方法皆没有对风格化的结果如何进行分析,结果好坏全靠人的主观评断.

6 结论

通过分析现有风格化的方法,发现对目标风格图片的笔触、纹理、颜色等风格信息的提取是图像风格化的关键,而人工神经网络的引入使人们对图像风格化的研究有了新的方向:卷积神经网络可以用来提取图片的语义信息,更好地表达内容以及风格特征;生成对抗网络的引入可以提高风格化的速度.因此,在未来的关于风格化的研究中,可以不断尝试使用人工神经网络进行风格化处理,更多地考虑图片的语义信息以及不同部分的重要性.

参考文献

References

- [1] Hertzmann A. Painterly rendering with curved brush strokes of multiple sizes [C] // Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, 1998:453-460
- [2] Zeng K, Zhao M T, Xiong C M, et al. From image parsing to painterly rendering [J]. ACM Transactions on Graphics, 2009, 29(1):2
- [3] Luo W, Lu Z, Wang X G, et al. Synthesizing oil painting surface geometry from a single photograph [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2012:885-892
- [4] Lu J, Barnes C, Diverdi S, et al. Realbrush: Painting with examples of physical media [J]. ACM Transactions on Graphics, 2013, 32(4):117
- [5] Wang B, Wang W P, Yang H P, et al. Efficient example-based painting and synthesis of 2D directional texture [J]. IEEE Transactions on Visualization and Computer Graphics, 2004, 10(3):266-277
- [6] Brooks S. Image-based stained glass [J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(6):1547-1558
- [7] Bousseau A, Neyret F, Thollot J, et al. Video water colorization using bidirectional texture advection [J]. ACM Transactions on Graphics, 2007, 26(3):104
- [8] Frigo O, Sabater N, Delon J, et al. Split and match: Example-based adaptive patch sampling for unsupervised style transfer [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2016:2338-2351
- [9] Baxter W, Govindaraju N. Simple data-driven modeling of brushes [C] // ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, 2010:135-142
- [10] Wang C M, Wang R J. Image-based color ink diffusion rendering [J]. IEEE Transactions on Visualization and Computer Graphics, 2007, 13(2):235-246
- [11] Zhang W, Cao C, Chen S F, et al. Style transfer via image component analysis [J]. IEEE Transactions on Multimedia, 2013, 15(7):1594-1601
- [12] Wang M Y, Wang B, Fei Y, et al. Towards photo watercolorization with artistic verisimilitude [J]. IEEE Transactions Visualization and Computer Graphics, 2014, 20(10):1451-1460
- [13] Kim Y, Winnemoller H, Lee S. WYSIWYG stereo painting

- with usability enhancements [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2014, 20 (7): 957-969
- [14] Chen Z L, Kim B, Ito D, et al. Wetbrush: GPU-based 3D painting simulation at the bristle level [J]. *ACM Transactions on Graphics*, 2015, 34(6): 200:1-200:11
- [15] Gatys L A, Ecker A S, Bethge M. Image style transfer using convolutional neural networks [C] // *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 2414-2423
- [16] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution [C] // *European Conference on Computer Vision*, 2016: 694-711
- [17] Li C, Wand M. Precomputed real-time texture synthesis with Markovian generative adversarial networks [C] // *European Conference on Computer Vision*, 2016: 702-716
- [18] Frigo O, Sabater N, Delon J, et al. Split and match: Example-based adaptive patch sampling for unsupervised style transfer [C] // *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 553-561
- [19] Ulyanov D, Lebedev V, Vedaldi A, et al. Texture networks: Feed-forward synthesis of textures and stylized images [C] // *International Conference on Machine Learning*, 2016: 1349-1357
- [20] Liao J, Yao Y, Yuan L, et al. Visual attribute transfer through deep image analogy [J]. *ACM Transactions on Graphics*, 2017, 36(4): 120:1-120:15
- [21] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [J]. *arXiv e-print*, 2017, arXiv:1703.10593
- [22] Chen D D, Yuan L, Liao J, et al. StyleBank: An explicit representation for neural image style transfer [C] // *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1897-1906
- [23] Dumoulin V, Shlens J, Kudlur M. A learned representation for artistic style [J]. *arXiv e-print*, 2017, arXiv:1610.07629
- [24] Chen H, Zhang N N, Liang L, et al. PicToon: A personalized image-based cartoon system [C] // *ACM International Conference on Multimedia*, 2002: 171-178
- [25] Chen H, Liu Z, Rose C, et al. Example-based composite sketching of human portraits [C] // *International Symposium on Non-photorealistic Animation and Rendering*, 2004: 95-153
- [26] Tseng C C, Lien J J J. Colored exaggerative caricature creation using inter and intra-correlations of feature shapes and positions [J]. *Image and Vision Computing*, 2012, 30(1): 15-25
- [27] Selim A, Elgharib M, Doyle L. Painting style transfer for head portraits using convolutional neural networks [J]. *ACM Transactions on Graphics*, 2016, 35(4): 129
- [28] Zhao M T, Zhu S C. Portrait painting using active templates [C] // *ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering*, 2011: 117-124
- [29] Wang T H, Collomosse J P, Hunter A, et al. Learnable stroke models for example-based portrait painting [C] // *British Machine Vision Conference*, 2013: 36
- [30] Zhang Y, Dong W M, Ma C Y, et al. Data-driven synthesis of cartoon faces using different styles [J]. *IEEE Transactions on Image Processing*, 2017, 26(1): 464-478
- [31] Wang N N, Tao D C, Gao X B, et al. A comprehensive survey to face hallucination [J]. *International Journal of Computer Vision*, 2014, 106(1): 9-30
- [32] Wang X G, Tang X O. Face photo-sketch synthesis and recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(11): 1955-1967
- [33] Li H L, Liu G H, Ngan K N. Guided face cartoon synthesis [J]. *IEEE Transactions on Multimedia*, 2011, 13(6): 1230-1239
- [34] Wang N N, Tao D C, Gao X B, et al. Transductive face sketch-photo synthesis [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, 24(9): 1364-1376
- [35] Zhang S C, Gao X B, Wang N N, et al. Face sketch synthesis via sparse representation-based greedy search [J]. *IEEE Transactions on Image Processing*, 2015, 24(8): 2466-2477
- [36] Zhang S C, Gao X B, Wang N N, et al. Robust face sketch style synthesis [J]. *IEEE Transactions on Image Processing*, 2016, 25(1): 220-232

A survey of image artistic stylization

DENG Yingying^{1,2} TANG Fan^{1,2} DONG Weiming¹

¹ National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

² University of Chinese Academy of Sciences, Beijing 100049

Abstract In recent years, image artistic style transfer has become a prosperous research field. More and more activities in this field have been promoted by scientific challenges and industrial needs, so image artistic style transfer is worthy of researching. In this paper, we analyze the present situation of image artistic style transfer, the characteristics of different style transfer methods, the shortcomings of the current style transfer methods and the development trend of image style transfer. Finally, we provide the direction for further style transfer research.

Key words image stylization; stroke; generative model