

# 一种自然场景图像的目标材质视觉特征映射算法

李策<sup>1</sup> 贾盛泽<sup>1</sup> 曲延云<sup>2</sup>

**摘要** 针对自然场景图像目标材质视觉特征映射中,尚存在特征提取困难、图像无对应标签等问题,本文提出了一种自然场景图像的目标材质视觉特征映射算法.首先,从图像中获取能表征材质视觉重要特征的反射层图像;然后,对获取的反射层图像进行前景、背景分割,得到目标图像;最后,利用循环生成对抗网络对材质视觉特征进行无监督学习,获得对图像目标材质视觉特征空间的高阶表达,实现了目标材质视觉特征的映射.实验结果表明,所提算法能够有效地获取自然场景图像目标的材质视觉特征,并进行材质视觉特征映射;与同类算法相比,具有更好的主、客观效果.

**关键词** 自然场景图像, 材质视觉特征, 特征映射, 循环生成对抗网络

**引用格式** 李策, 贾盛泽, 曲延云. 一种自然场景图像的目标材质视觉特征映射算法. 自动化学报, 2019, 45(6): 1198–1206

**DOI** 10.16383/j.aas.c180618

## A Material Visual Features Mapping Algorithm With Natural Scene Image Objects

LI Ce<sup>1</sup> JIA Sheng-Ze<sup>1</sup> QU Yan-Yun<sup>2</sup>

**Abstract** The natural scene image object material visual feature mapping is still a challenging task, such as the difficulties of feature extraction and images without corresponding labels. A material visual features mapping algorithm with natural scene image object is proposed in this paper. First, the proposed algorithm extracts the reflection layer image that can represent the essential features of the material vision, then it divides the acquired reflection layer image into the segmentation of foreground and background to obtain the object image. Finally, the cycle-GAN (Generative adversarial network) is used for the unsupervised learning of the material visual features, and the visual features of image object is obtained. The high level expression of material visual feature space can achieve the mapping of visual characteristics of object material. The experiment results show that the algorithm can obtain the visual features of the natural scene image object and carry on the material visual feature mapping much better. Compared with the existing methods for material visual feature mapping, the proposed algorithm achieves better subjective and objective effects.

**Key words** Natural scene image, material visual feature, feature mapping, cycle-GAN

**Citation** Li Ce, Jia Sheng-Ze, Qu Yan-Yun. A material visual features mapping algorithm with natural scene image objects. *Acta Automatica Sinica*, 2019, 45(6): 1198–1206

材质视觉特征是一种基于视觉机制的高阶综合特征,描述了图像或图像区域对应目标的材质信息,从中可以获取目标的材质属性及其性质<sup>[1]</sup>.日常生活中,我们会遇到各类的材质,如木质、陶瓷、金属等,这些材质信息具有广泛的物理性和功能性表达.通过视觉系统得到的材质视觉特征,能够帮助我们

在接触物体之前,决定我们以何种接触方法更合适.同时,该方法应用于机器人领域,自主机器人可以通过视觉系统获取到物体的材质特征<sup>[2-3]</sup>,进一步决定抓取目标物体的发力控制.如果材质易碎,需小心接触;若材质密度较高,则考虑使用更大的力量搬动.因此,研究获取材质视觉特征空间具有重要的理论价值和广泛的应用前景.

近些年,随着映射学习技术的兴起,吸引了越来越多的研究人员开始从特征映射的角度着手,获取图像目标的材质特征空间信息,进而完成不同图像间目标材质视觉特征映射.传统的材质视觉特征映射方法是这类问题看做图像的滤波问题,而不进行材质视觉特征的显式分析.Khan等<sup>[4]</sup>利用简单的启发式算法,从图像中获取图像的近似形状和光照,并利用这些信息进行材质视觉特征映射.Boyadzhiev等<sup>[5]</sup>提出了几种滤波器,能够改变图像的性质,如光照和光泽度等.传统的方法虽然也能取

收稿日期 2018-09-18 录用日期 2018-10-26  
Manuscript received September 18, 2018; accepted October 26, 2018

国家自然科学基金(61866022, 61876161), 甘肃省基础研究创新群体(1506RJIA031)资助

Supported by National Natural Science Foundation of China (61866022, 61876161) and Gansu Province Basic Research Innovation Group Project (1506RJIA031)

本文责任编辑 张军平

Recommended by Associate Editor ZHANG Jun-Ping

1. 兰州理工大学电气工程与信息工程学院 兰州 730050 2. 厦门大学信息科学与技术学院计算机科学系 厦门 361005

1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050 2. College of Information Science and Engineering, Department of Computer Science, Xiamen University, Xiamen 361005

得一定的效果,使得映射后的图像具有逼真的效果,但以上传统方法都未能对图像的材质视觉属性进行有效分析。

随着深度学习技术在各种计算机视觉任务中不断取得成功,深度卷积神经网络也被用来获取材质视觉特征。Liu等<sup>[6]</sup>使用卷积神经网络对材质特征进行了识别与分类。Li<sup>[7]</sup>基于卷积神经网络和集成学习进行了材质的识别与分割。Zheng等<sup>[8]</sup>利用基于径向基函数的卷积神经网络进行超宽带探地雷达目标的材质识别。Tang等<sup>[9]</sup>提出了一种深度朗伯网络,从单点光源的方向获取反射材质的属性,并利用高斯受限的玻尔兹曼机模拟材质表面的反射率和方位。Richter等<sup>[10]</sup>使用随机森林从数据库中提取物体表面的材质块,以获取具有均匀反射率的物体形状。Zhou等<sup>[11]</sup>利用卷积神经网络获取图像中两个像素间的相对材质性质,之后进行约束优化,以解决整幅图像的反射率问题。Narihira等<sup>[12]</sup>提出了一种直接从图像中获取反射率和阴影的卷积神经网络。Kulkarni等<sup>[13]</sup>在单个图像中使用变分自动编码器将视角、光照和其他固有特征(如形状、纹理等)分离出来。Rematas等<sup>[14]</sup>使用了两个独立的卷积神经网络,从材质和光照的特征组合中,将两种特征单独分离出来。Liu等<sup>[15]</sup>认为图像的形成过程,是图像的重要特征(如形状、材质、光照等)相互影响和融合的结果。该算法通过模拟这一融合过程,使用卷积神经网络,将包括图像材质信息在内的特征信息分离了出来,进而完成图像材质视觉特征的映射。

基于以上分析,目前已有方法虽然在获取图像目标材质视觉特征,进行特征映射的问题上取得了一定的效果,但仍存在以下问题:1)已有方法中所使用的图像数据,均为依赖计算机图形学所生成的图像,而对于无标签的自然场景图像,仍然无法有效获取相应的材质视觉特征,并进行特征映射;2)由于材质视觉特征的自身属性,光照信息在进行材质视觉特征映射过程中,往往会带来一定的影响。例如,不同的材质对光的反射效果不同,不同的光照角度会有不同的材质视觉感受等;3)目标与背景的相似性,在材质视觉特征映射过程中,同样也会带来一定的干扰,导致不能准确地获取目标材质视觉特征空间。针对以上问题,本文提出了一种基于循环生成对抗网络(Cycle-generative adversarial network, Cycle-GAN)<sup>[16]</sup>的自然场景图像目标材质视觉特征映射算法。

生成对抗网络(Generative adversarial network, GAN)<sup>[17-18]</sup>由一个生成器和一个判别器构成,通过对抗学习的方式来训练,已被广泛应用在人脸表情识别<sup>[19]</sup>、图像识别<sup>[20]</sup>等诸多领域。而循环生成对抗网络是以生成对抗网络为基础,使用两个生

成对抗网络组成循环生成结构,在特征映射网络的学习训练过程中,采用自然场景下的无标签图像作为训练数据,形成无监督学习方式。同时,在前期处理过程中加入了相应的处理方法,有效避免了光照信息和背景相似性的影响。最终,完成自然场景下不同目标间的材质视觉特征映射。

## 1 本文所提算法

物体的材质视觉特征和物体的颜色、亮度、纹理、空间位置一样,是物体的视觉固有属性<sup>[21]</sup>。材质特征可以简单地看作是物体表面各可视特征属性的结合,这些特征属性包括物体表面的纹理、色彩、光滑度、透明度、折射率等,正是这些特征的有机融合,构成了材质视觉特征这一高阶综合特征。

本文提出了一种基于循环生成对抗网络的自然场景图像目标材质视觉特征映射算法,主要包括两个部分:图像前期预处理和构建材质视觉特征空间的映射网络,所提算法整体框架如图1所示。本文所提算法可对源域 $X$ 和目标域 $Y$ 中的图像同步实施操作。首先,基于Retinex理论<sup>[22]</sup>,建立光照-反射模型,将人眼观测到的原始图像认为是光照层图像和反射层图像的乘积,进行光照层信息估计,从原始图像中获取反射层图像,能够更好地表征材质视觉特征,避免了光照信息在材质视觉特征映射过程中的影响。然后,对获取到的反射层图像进行前景、背景分割,获取目标图像,以减少背景相似性的影响。以上两部分均属于对图像的前期预处理过程,目的是减少材质视觉特征在映射过程中所受到的干扰。最后,构建循环生成对抗网络,利用深度多层卷积的方式对材质特征进行无监督学习,得到对图像目标材质视觉特征空间的高阶表达,建立源域图像 $X$ 与目标域图像 $Y$ 之间关于材质视觉特征空间的映射网络。

### 1.1 基于Retinex理论的反射层图像获取

材质视觉特征可以被认为是决定光照如何与物体表面相互作用的属性。物体表面反射的灯光颜色、反射率、透明度和光泽度等都是构成物体表面材质的参数。不同的材质,其物理组成是不同的,这就导致其在不同的光照角度下,具有不同的反射效果,若直接进行材质特征的提取或识别,光照信息就会具有较大的干扰性<sup>[23-24]</sup>。

本文使用自然场景图像作为原始数据,其所在场景的光照信息一般都较为复杂,光照信息往往会在材质视觉特征空间的提取过程中带来较大的干扰,进而导致材质视觉特征在映射过程中会出现较大的偏差。因此,本文在前期的预处理过程中,对自然场景图像进行了处理,即选择基于色彩恒常性理论发

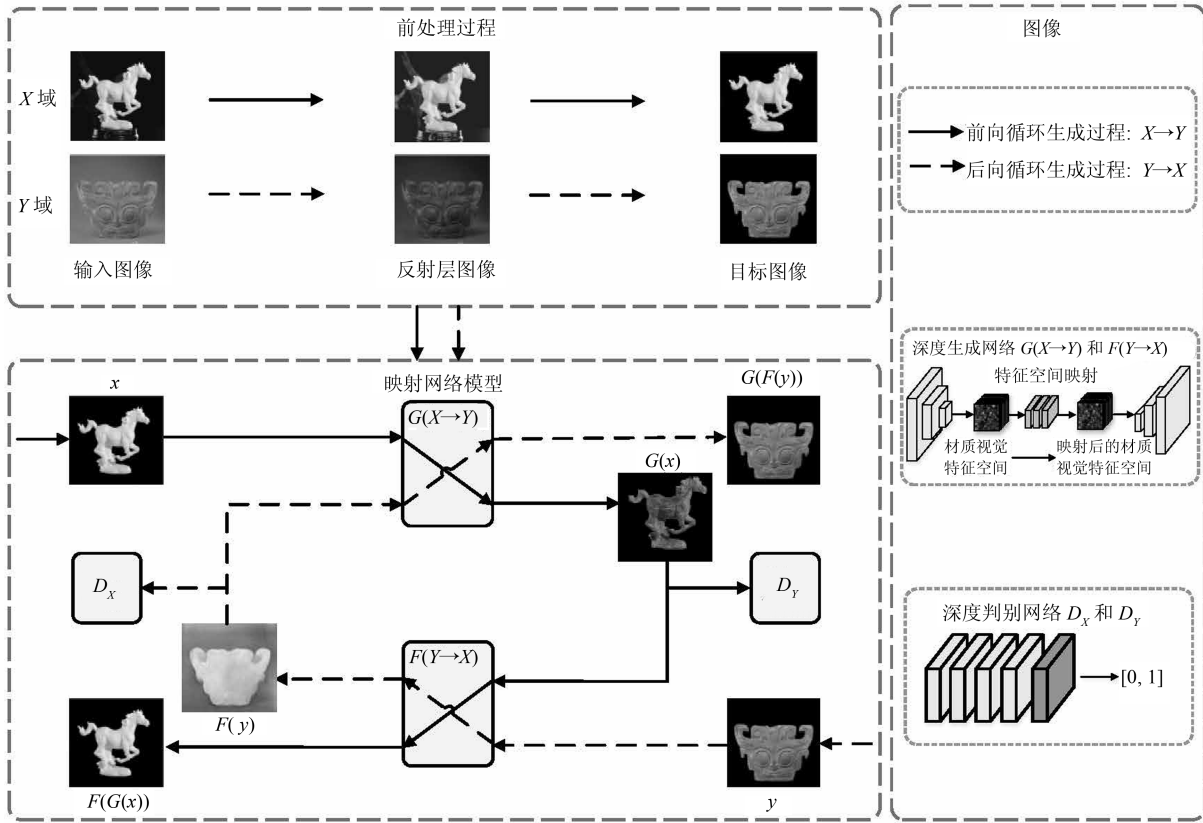


图 1 本文所提算法框架

Fig. 1 The framework of the proposed algorithm

展而来的 Retinex 理论<sup>[22]</sup> 建立光照-反射模型, 对光照层信息进行估计, 从而分离图像的光照信息, 解决光照干扰的问题, 具体步骤如下:

1) 建立光照-反射模型. 采用 Land 的视网膜理论 (Retinex 理论)<sup>[22]</sup> 建立光照-反射模型. 该理论认为图像的颜色是由入射光照射到目标物体, 并经由物体表面不同的反射系数反射到图像采集传感系统形成. 已知图像在位置  $(x, y)$  处的采样值  $I(x, y)$ , 对影响图像采集过程中的光照信息  $L(x, y)$  进行估计, 从而计算得到反射层图像. Retinex 理论<sup>[22]</sup> 应用在图像处理中, 将人眼观测到的原始图像认为是光照层图像和反射层图像的乘积:

$$I(x, y) = L(x, y)R(x, y) \quad (1)$$

转换为对数域描述:

$$\ln(I(x, y)) = \ln(L(x, y)) + \ln(R(x, y)) \quad (2)$$

其中,  $I(x, y)$  为在位置  $(x, y)$  的图像像素值,  $L(x, y)$  表示光照信息,  $R(x, y)$  表示图像中表征材质视觉重要特征的反射层图像信息.

2) 获取光照层信息. 在进行光照层信息估计时, 使用双边滤波<sup>[25]</sup> 得到光照层信息. 双边滤波作为一

种非线性滤波器, 具有边缘保护和降噪平滑的效果, 采用加权平均的方法, 用周边像素亮度值的加权平均代表某个像素的强度. 同时, 双边滤波的权重不仅考虑了像素的欧氏距离, 还考虑了像素范围域中的辐射差异, 因而选用双边滤波可以更好地对原始图像的光照层信息进行计算:

$$L(x, y) = BF(I; \sigma_s, \sigma_r) = \frac{1}{W} \sum_{q \in \Omega} G_{\sigma_s}(|p - q|) G_{\sigma_r}(|I_p - I_q|) I_q \quad (3)$$

其中,  $BF(I; \sigma_s, \sigma_r)$  表示双边滤波过程及其参数,  $I$  为滤波图像, 参数  $\sigma_s$  和  $\sigma_r$  定义了双边滤波的尺寸大小;  $I_p$  为  $p$  点像素值,  $I_q$  为  $q$  点像素值;  $\Omega$  表示整幅图像的范围;  $W$  是归一化权重:

$$W = \sum_{q \in \Omega} G_{\sigma_s}(|p - q|) G_{\sigma_r}(|I_p - I_q|) \quad (4)$$

$G_{\sigma_s}(x)$  确定了与图像像素点位置相关的权重;  $G_{\sigma_r}(x)$  确定了与图像像素值大小相关的权重.

3) 获取反射层图像. 将图像从 RGB 颜色空间转换到 HSV 颜色空间, 可以得到色度 (Hue)、饱和度 (Saturation) 以及亮度 (Value) 分量. 相对于

RGB 颜色特征空间, HSV 颜色特征空间能够直观地表达色彩的明暗、色调以及鲜艳程度, 而光照对图像的影响则主要集中在亮度分量的变化上. 基于以上分析, 本文在亮度分量上建立光照-反射模型, 获取反射层图像, 保持色度以及饱和度分量信息不变.

## 1.2 目标图像获取

已有的材质视觉特征映射算法, 所使用的图像数据均为使用计算机图形学的方式所生成图像, 这类图像数据的特点除每一幅图像都有相应的标签数据外, 其背景环境也相对简单. 本文所使用的图像数据均为自然场景图像, 背景往往相对复杂, 其目标与背景也具有一定的相似性. 这些干扰因素, 同光照信息一样, 都会在材质视觉特征的映射过程中带来一定的影响. 为解决这个问题, 本文在进行网络训练前, 对获取的反射层图像进行前景、背景分割, 获取目标图像. 同时, 考虑到本文所使用的数据图像均来自于互联网, 整体数据量较小, 为满足深度学习对训练数据的要求, 在网络的训练阶段, 还需对获取的目标图像采取一系列措施来扩充训练数据库, 具体步骤如下:

1) 前景、背景分割. 采用 GraphCut 分割算法<sup>[26]</sup> 对获取的反射层图像进行前景、背景分割, 获取目标图像, 以避免背景因素在材质视觉特征映射过程中的干扰.

2) 数据集扩充. 从互联网中采集自然场景图像来建立数据库, 并对获取的目标图像进行翻转、随机裁剪等操作来扩充数据集, 得到最终的训练数据集, 整体的预处理过程如图 2 所示.

## 1.3 材质视觉特征映射网络

本文所提出的算法, 与已有材质视觉特征映射算法最大的区别在于, 其使用的原始数据为依据计算机图形学所生成的图像, 而本文算法使用自然场景下无标签图像作为原始数据, 导致一般的深度学习网络无法有效获取材质视觉特征空间. 为解决上述问题, 本文采用循环生成对抗网络<sup>[16]</sup>, 在生成对抗网络<sup>[17-18]</sup> 的基础上, 使用两个生成对抗网络构成循环生成机制, 以强化材质视觉特征学习能力. 在训练过程中, 不使用标签信息对网络进行指导, 从而

构成无监督学习模型. 采用从自然场景图像中获取的反射层目标图像作为训练数据, 得到对目标材质视觉特征空间的高阶表达, 建立源域图像  $X$  与目标域图像  $Y$  间关于材质视觉特征空间的映射网络, 将源域图像  $X$  的材质视觉特征映射到目标域图像  $Y$  的材质视觉特征上, 使目标域图像  $Y$  具有源域图像  $X$  的材质视觉特征信息. 整体的网络结构包括两个生成网络  $G(\cdot): x \rightarrow y$  和  $F(\cdot): y \rightarrow x$  以及两个判别网络  $D_X$  和  $D_Y$ . 其中, 生成网络  $G(\cdot)$  和  $F(\cdot)$  具有相同的网络结构, 以学习构建材质视觉特征映射网络, 得到材质视觉特征映射后的图像. 同样, 两个判别网络  $D_X$  和  $D_Y$  也具有相同的网络结构, 以判别生成图像与真实图像.

### 1.3.1 判别网络

判别网络用来对输入的图像进行判别分类, 得到正确的判别概率. 整个判别网络结构如图 3 所示, 具体包括一个输入层、4 个卷积层、1 个全连接层和一个输出层. 为降低输入与初始化对参数更新的影响, 加快训练速度, 在判别网络中加入批标准化处理 (BatchNorm)<sup>[27]</sup>, 并进行 LeakyReLU 激活函数操作. 输入层将生成图像和真实图像分别送入判别网络中, 利用卷积神经网络对输入信息进行多层感知, 得到表征输入信息属性的高阶特征, 经过全连接层后连接到输出层, 通过激活函数 Sigmoid 进行分类判别.

### 1.3.2 生成网络

生成网络用来完成对材质视觉特征的提取, 并完成特征空间的映射过程. 整个生成网络结构如图 4 所示, 具体包括 4 个卷积层、2 个去卷积层、9 个残差网络块 (ResNet)<sup>[28]</sup>. 在生成网络的构造过程中, 从视觉感知的角度出发, 考虑图像目标材质视觉特征空间的提取和映射过程, 将生成网络分为编码、映射和解码三大主要组成部分.

首先, 编码部分利用卷积神经网络对材质特征空间进行逐层抽象化, 并结合 ReLU 激活函数对相关特征信息进行融合和特征提取, 得到自然场景图像中材质视觉特征空间的高阶表达. 其次, 映射部分使用残差网络块<sup>[28]</sup> 建立特征空间的映射网络, 用来加强网络层间的联系, 避免特征信息在映射过程中丢失. 最后, 利用去卷积网络<sup>[29]</sup> 对抽象特征进行感

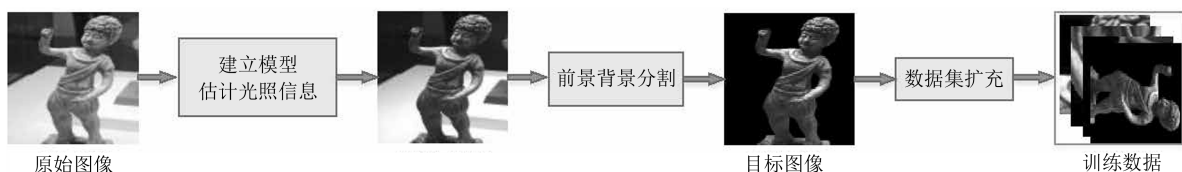


图 2 反射层图像及目标图像获取

Fig. 2 Extract the reflection layer and object images

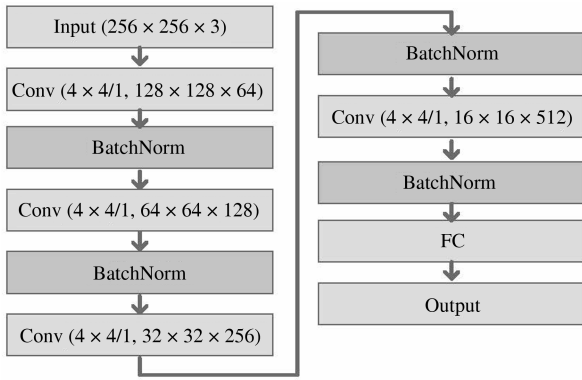


图 3 判别网络结构

Fig. 3 The structure of discriminator

知分解, 通过 ReLU 激活函数融合特征得到特征映射后的图像。

### 1.3.3 损失函数

在训练循环生成对抗网络时, 为指导网络的学

习, 保障网络输入量与输出量间的关系, 可结合网络结构的特殊性以及图像目标材质视觉特征的独特性, 将损失函数设定为:

$$L(G, F, D_X, D_Y) = \alpha \cdot (L_{GAN}(G, D_Y, X, Y) + L'_{GAN}(F, D_X, Y, X)) + \beta \cdot L_{cyc}(G, F) + \gamma \cdot L_{perceptual} \quad (5)$$

其中,  $\alpha$ 、 $\beta$ 、 $\gamma$  分别表示三项损失所占的权重, 在本文中, 分别设定  $\alpha = 1$ ,  $\beta = 8$ ,  $\gamma = 5$ .  $L_{perceptual}(\cdot)$  为感知损失<sup>[30]</sup>, 其结构如图 5 所示。

$L_{GAN}(G, D_Y, X, Y)$  和  $L'_{GAN}(F, D_X, Y, X)$  分别表示生成网络  $G(\cdot): X \rightarrow Y$  及其判别网络  $D_Y$  和生成网络  $F(\cdot): Y \rightarrow X$  及其判别网络  $D_X$ , 在传统生成对抗网络下的损失函数:

$$L_{GAN}(G, D_Y, X, Y) = E_{y \sim p_{data}(y)}[\log D_Y(y)] + E_{x \sim p_{data}(x)}[\log(1 - D_Y(G(x)))] \quad (6)$$

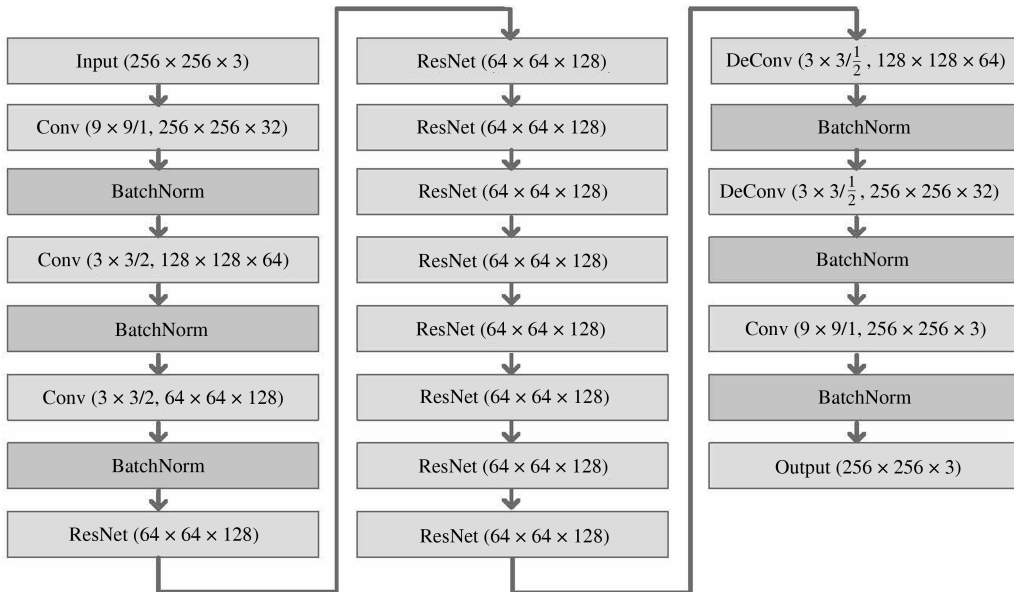


图 4 生成网络结构

Fig. 4 The generator structure

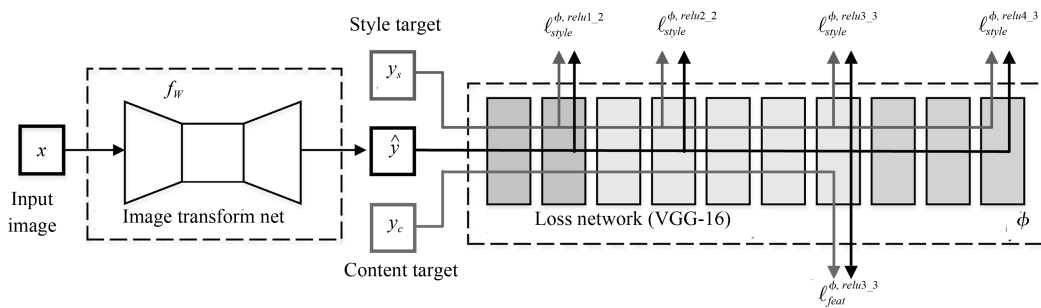


图 5 感知损失结构<sup>[30]</sup>

Fig. 5 The perceptual loss structure<sup>[30]</sup>

$$L'_{GAN}(F, D_X, Y, X) = E_{x \sim p_{data}(x)}[\log D_X(x)] + E_{y \sim p_{data}(y)}[\log(1 - D_X(F(y)))] \quad (7)$$

$L_{cyc}(G, F)$  表示循环生成对抗网络结构中, 前向循环过程  $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$  的损失和后向循环过程  $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$  的损失, 其结构如图 6 所示, 具体表示为:

$$L_{cyc}(G, F) = E_{x \sim p_{data}(x)}[\|F(G(x)) - x\|_1] + E_{y \sim p_{data}(y)}[\|G(F(y)) - y\|_1] \quad (8)$$

## 2 实验结果与分析

本文自行建立数据库, 从互联网中采集了四种具有不同材质视觉特征的自然场景图像, 分别为青白瓷器、陶器、玉器以及青铜器, 每类图像为 500 幅. 为验证所提算法的有效性, 在所构建数据集上与当前一些主流的特征映射算法 Gatys<sup>[31]</sup>、Li<sup>[30]</sup>、VAT<sup>[32]</sup> 进行了主、客观实验对比. 其中, Gatys<sup>[31]</sup> 使用卷积神经网络的方式将整个特征映射过程分为了内容重建与特征重建, Li<sup>[30]</sup> 则建立了感知损失网络, 获取整个特征重建所需的损失函数, 而 VAT<sup>[32]</sup> 则把图像块匹配和特征重建从图像领域扩充到了特征领域. 本文实验依赖于 Pytorch0.4, 计算平台 CPU 为 Intel(R) Core(TM) i7-6700, GPU 为 NVIDIA GTX1070 8GB 显存, 使用 Ubuntu16.04 系统. 在训练网络的过程中, 进行 5000 次迭代, 学习率设定为 0.002.

### 2.1 主观结果分析

根据从互联网中采集到的四种具有不同材质视觉特征的自然场景图像, 本文一共设计了四类材质视觉特征映射, 分别为从青白瓷映射到青铜器, 陶器映射到青白瓷, 青铜器映射到玉器以及从陶器

映射到青铜器. 图 7 所示是本文算法与对比算法 Gatys<sup>[31]</sup>、Li<sup>[30]</sup>、VAT<sup>[32]</sup> 的主观实验效果, 从图中可以看出, 对比算法在对自然场景图像目标进行材质视觉特征映射时, 均存在一定程度上的颜色偏差或目标偏离, 而本文算法在保持图像目标完整性的基础上, 具有良好的材质视觉特征映射效果, 并在各类材质视觉特征中都具有一定的效果, 可以准确地映射得到材质视觉特征.

### 2.2 客观结果分析

本文采用自然场景图像作为原始数据, 在进行材质视觉特征映射后, 由于得到的图像在自然界中并不真实存在, 所以基于有参考图像的客观评价标准在本文中并不适用. 为解决这个问题, 本文对映射生成的图像采用客观质量评价指标进行比较, 分别选取局部信息熵、平均梯度以及两种近年来具有代表性意义的无参考图像质量评价方法 IL-NIQE<sup>[33]</sup> 和 MEON<sup>[34]</sup> 对本文算法和对比算法进行客观评价. 其中, 局部信息熵通常用来衡量图像中信息的丰富程度, 平均梯度可以用来表示图像的相对清晰度. IL-NIQE<sup>[33]</sup> 是一种完全无参考图像质量评价方法, 预先使用大量的优质自然场景图像, 建立依据 NSS 特征的 MVG 模型, 在将待测图像的质量, 而 MEON<sup>[34]</sup> 结合深度学习技术, 使用一个失真判别网络和一个质量判别网络来实现无参考图像的质量评价.

表 1 为本文算法与对比算法在平均梯度和局部信息熵上的对比结果, 表 2 为在图像质量评价方法 IL-NIQE<sup>[33]</sup> 和 MEON<sup>[34]</sup> 上的对比结果. 其中, mAP 表示四类材质视觉特征映射客观评价标准的平均值, 避免使用单一种类进行评价所带来的误差. 实验结果表明, 本文算法与对比算法相比, 具有更高

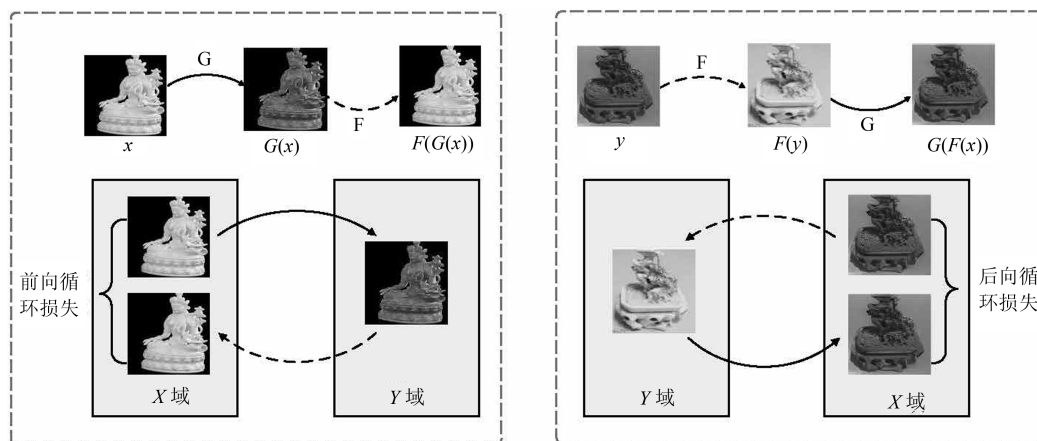


图 6 循环损失结构

Fig. 6 The cycle loss structure



图 7 主观实验结果对比图

Fig. 7 The comparison of subjective experiment results

的局部信息熵和平均梯度, 更低的 IL-NIQE<sup>[33]</sup> 和 MEON<sup>[34]</sup> 值, 说明进行材质视觉特征映射后图像的信息丰富程度、清晰度相对更高, 具有更好的图像质量.

主、客观实验结果表明, 本文所提算法能够有效的获取自然场景图像目标材质视觉特征, 抑制无用信息的干扰和影响, 较为真实地映射得到材质视觉

特征信息. 在完成特征映射的同时, 可以有效地保持场景的原有结构, 所得结果较同类特征映射算法具有更好的主、客观效果.

### 3 结论

本文提出了一种自然场景图像的目标材质视觉特征映射算法, 运用获取的目标反射层图像, 通过在

表1 平均梯度与局部信息熵对比结果

Table 1 Comparison in terms of both the average gradient and the local information entropy

对比方法	青白瓷映射到青铜器		陶器映射到青白瓷		青铜器映射到玉器		陶器映射到青铜器		mAP	
	平均梯度	局部信息熵	平均梯度	局部信息熵	平均梯度	局部信息熵	平均梯度	局部信息熵	平均梯度	局部信息熵
Gatys <sup>[31]</sup>	0.020	3.470	0.030	3.884	0.024	3.457	0.022	3.937	0.024	3.687
Li <sup>[30]</sup>	0.017	4.425	0.030	4.134	0.023	3.906	0.042	<b>4.956</b>	0.028	4.356
VAT <sup>[32]</sup>	0.025	4.279	0.043	4.391	<b>0.031</b>	4.368	0.101	4.639	0.050	4.419
Ours	<b>0.042</b>	<b>4.796</b>	<b>0.056</b>	<b>4.554</b>	0.029	<b>4.652</b>	<b>0.117</b>	4.462	<b>0.061</b>	<b>4.616</b>

表2 IL-QINE 与 MEON 对比结果

Table 2 Comparison in terms of both the IL-QINE and the MEON

对比方法	青白瓷映射到青铜器		陶器映射到青白瓷		青铜器映射到玉器		陶器映射到青铜器		mAP	
	IL-QINE	MEON	IL-QINE	MEON	IL-QINE	MEON	IL-QINE	MEON	IL-QINE	MEON
Gatys <sup>[31]</sup>	61.946	50.513	69.141	25.403	54.519	22.457	49.441	38.496	58.762	34.217
Li <sup>[30]</sup>	57.231	42.357	52.973	40.510	48.956	55.490	43.268	42.449	50.607	45.201
VAT <sup>[32]</sup>	52.932	<b>16.013</b>	50.618	51.271	50.280	45.226	45.257	29.625	49.771	35.534
Ours	<b>46.689</b>	17.309	<b>49.372</b>	<b>6.709</b>	<b>47.817</b>	<b>12.917</b>	<b>43.021</b>	<b>14.955</b>	<b>46.725</b>	<b>12.973</b>

生成对抗网络上加入循环结构, 构成循环生成对抗网络来获取材质视觉特征空间, 继而完成特征映射, 得到材质视觉特征映射后的图像. 主、客观对比实验表明, 所提算法在保持原有图像目标不变的基础上, 能够有效地完成材质视觉特征映射, 与同类算法相比, 具有更好的图像质量效果.

## References

- Fleming R W. Material perception. *Annual Review of Vision Science*, 2017, **3**(1): 365–388
- Zhong Xun-Gao, Xu Min, Zhong Xun-Yu, Peng Xia-Fu. Multimodal features deep learning for robotic potential grasp recognition. *Acta Automatica Sinica*, 2016, **42**(7): 1022–1029  
(仲训泉, 徐敏, 仲训昱, 彭侠夫. 基于多模特征深度学习的机器人抓取判别方法. *自动化学报*, 2016, **42**(7): 1022–1029)
- Jia Bing-Xi, Liu Shan, Zhang Kai-Xiang, Chen Jian. Survey on robot visual servo control: Vision system and control strategies. *Acta Automatica Sinica*, 2015, **41**(5): 861–873  
(贾丙西, 刘山, 张凯祥, 陈剑. 机器人视觉伺服研究进展: 视觉系统与控制策略. *自动化学报*, 2015, **41**(5): 861–873)
- Khan E A, Reinhard E, Fleming R W, Bülthoff H H. Image-based material editing. *ACM Transactions on Graphics*, 2006, **25**(3): 654–663
- Boyadzhiiev I, Bala K, Paris S, Adelson E. Band-sifting decomposition for image-based material editing. *ACM Transactions on Graphics*, 2015, **34**(5): 163–179
- Liu Hao, Li Zhe, Shi Jing, Xin Min-Si, Cai Hong-Xing, Gao Xue, Tan Yong. Study on classification and recognition of materials based on convolutional neural network. *Laser & Infrared*, 2017, **47**(8): 1024–1028  
(刘昊, 李哲, 石晶, 辛敏思, 蔡红星, 高雪, 谭勇. 基于卷积神经网络的材质分类识别研究. *激光与红外*, 2017, **47**(8): 1024–1028)
- Li Wan-Wan. Ensemble Learning for Material Recognition and Segmentation with Convolutional Neural Networks [Master thesis], Beijing Jiaotong University, China, 2018  
(李婉婉. 基于卷积神经网络和集成学习的材质识别和分割方法研究 [硕士学位论文], 北京交通大学, 中国, 2018)
- Zheng Jun-Ting, Li Jian, Li Jian-Xun. The application of RBF neural network in material recognition ultra wideband ground penetrating radar. *Journal of Shanghai Jiaotong University*, 2006, **40**(1): 98–102  
(郑军庭, 李建, 李建勋. 径向基函数神经网络在超宽带探地雷达目标材质识别中的应用. *上海交通大学学报*, 2006, **40**(1): 98–102)
- Tang Y C, Salakhutdinov R, Hinton G. Deep lambertian networks. In: *Proceedings of the 2012 International Conference on Machine Learning*. Edinburgh, Scotland: ACM, 2012. 1623–1630
- Richter S R, Roth S. Discriminative shape from shading in uncalibrated illumination. In: *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA: IEEE, 2015. 1128–1136
- Zhou T H, Krahenbuhl P, Efros A A. Learning data-driven reflectance priors for intrinsic image decomposition. In: *Proceedings of the 2015 IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015. 3469–3477
- Narihira T, Maire M, Yu S X. Direct intrinsics: learning albedo-shading decomposition by convolutional regression. In: *Proceedings of the 2015 IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015. 2992–3001
- Kulkarni T D, Whitney W F, Kohli P, Tenenbaum J. Deep convolutional inverse graphics network. In: *Proceedings of the 2015 Annual Conference on Neural Information Processing Systems*. Montreal, Canada: MIT Press, 2015. 2539–2547
- Rematas K, Ritschel T, Fritz M, Gavves E, Tuytelaars T. Deep reflectance maps. In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016. 4508–4516
- Liu G, Ceylan D, Yumer E, Yang J M, Lien J M. Material editing using a physically based rendering network. In: *Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE, 2017. 2261–2269



- 16 Zhu J Y, Park T, Isola P, Efros A A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 2223–2232
- 17 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 2014 Annual Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014. 2672–2680
- 18 Wang Kun-Feng, Gou Chao, Duan Yan-Jie, Lin Yi-Lun, Zheng Xin-Hu, Wang Fei-Yue. Generative adversarial networks: the state of the art and beyond. *Acta Automatica Sinica*, 2017, **43**(3): 321–332  
(王坤峰, 苟超, 段艳杰, 林懿伦, 郑心湖, 王飞跃. 基于多模特征深度学习的机器人抓取判别方法. *自动化学报*, 2017, **43**(3): 321–332)
- 19 Yao Nai-Ming, Guo Qing-Pei, Qiao Feng-Chun, Chen Hui, Wang Hong-An. Robust facial expression recognition with generative adversarial networks. *Acta Automatica Sinica*, 2018, **44**(5): 865–877  
(姚乃明, 郭清沛, 乔逢春, 陈辉, 王宏安. 基于生成式对抗网络的鲁棒人脸表情识别. *自动化学报*, 2018, **44**(5): 865–877)
- 20 Tang Xian-Lun, Du Yi-Ming, Liu Yu-Wei, Li Jia-Xin, Ma Yi-Wei. Image recognition with conditional deep convolutional generative adversarial networks. *Acta Automatica Sinica*, 2018, **44**(5): 855–864  
(唐贤伦, 杜一铭, 刘雨微, 李佳歆, 马艺玮. 基于条件深度卷积生成对抗网络的图像识别方法. *自动化学报*, 2018, **44**(5): 855–864)
- 21 Zhu Y, Zhang Z Y. Research on users' product material perception. In: Proceedings of the 2010 IEEE International Conference on Computer-aided Industrial Design & Conceptual Design. Wenzhou, China: IEEE, 2010. 1277–1280
- 22 Land E H, McCann J J. Lightness and retinex theory. *Journal of the Optical Society of America*, 1971, **61**(1): 1–11
- 23 Fleming R W. Visual perception of materials and their properties. *Vision Research*, 2014, **94**(1): 62–75
- 24 Zhang H, Dana K, Nishino K. Reflectance hashing for material recognition. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 3071–3080
- 25 Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of the 1998 IEEE International Conference on Computer Vision. Bombay, India: IEEE, 1998. 839–846
- 26 Ben S M, Mitiche A, Ben A I. Multiregion image segmentation by parametric kernel graph cuts. *IEEE Transactions on Image Processing*, 2011, **20**(2): 545–557
- 27 Ioffe S, Szegedy C. Batch normalization: Accelerating deep-network training by reducing internal covariate shift. In: Proceedings of the 2015 International Conference on Machine Learning. Lille, France: ACM, 2015. 448–456
- 28 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016. 770–778
- 29 Springenberg J T, Dosovitskiy A, Brox T, Riedmiller M. Striving for simplicity: The all convolutional net. In: Proceedings of the 2015 International Conference on Learning Representations. San Diego, CA, USA: IEEE, 2015. 1–14
- 30 Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution. In: Proceedings of the 2016 European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 694–711
- 31 Gatys L A, Ecker A S, Bethge M. Image style transfer using convolutional neural networks. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016. 2414–2423
- 32 Liao J, Yao Y, Yuan L, Hua G, Kang S B. Visual attribute transfer through deep image analogy. *ACM Transactions on Graphics*, 2017, **36**(4): 1–15
- 33 Zhang L, Zhang L, Bovik A C. A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, 2015, **24**(8): 2579–2591
- 34 Ma K D, Liu W T, Zhang K, Duanmu Z F, Wang Z, Zuo W M. End-to-end blind image quality assessment using deep neural networks. *IEEE Transactions on Image Processing*, 2018, **27**(3): 1202–1213



**李 策** 工学博士, 兰州理工大学电气工程与信息工程学院教授. 主要研究方向为计算视觉与模式识别, 智能机器人, 图像处理及应用. 本文通信作者.

E-mail: xjtulice@gmail.com

(**LI Ce** Ph.D., professor in the College of Electrical and Information Engineering, Lanzhou University of Technology.

His research interest covers computer vision and pattern recognition, intelligent robot, image processing and application. Corresponding author of this paper.)



**贾盛泽** 兰州理工大学硕士研究生. 主要研究方向为计算机视觉与图像处理.

E-mail: jiasz0607@163.com

(**JIA Sheng-Ze** Master student in the College of Electrical and Information Engineering, Lanzhou University of Technology. His research interest covers computer vision and image processing.)



**曲延云** 工学博士, 厦门大学信息科学与技术学院计算机科学系教授. 主要研究方向为模式识别, 计算机视觉和机器学习. E-mail: yyqu@xmu.edu.cn

(**QU Yan-Yun** Ph.D., professor in the Department of Computer Science, the College of Information Science and Engineering, Xiamen University. Her

research interest covers pattern recognition, computer vision and machine learning.)