深度特征融合的头发属性转移方法

谢志峰[1,2], 苏旭[1], 刘思维[1], 张桂菘[1], 马利庄[2,3]

[1] (上海大学影视工程系 上海 200072)
[2] (上海电影特效工程技术研究中心 上海 200072)
[3] (上海交通大学计算机科学与工程系 上海 200240)
(zhifeng_xie@shu.edu.cn)

摘 要: 针对现有的属性转移方法无法有效地转移头发属性这一问题, 提出一种深度特征融合的头发属性转移方法. 该方法包括特征提取、属性向量获取和图像合成 3 个子网络. 首先从特征提取网络中提取原图像特征, 添加重构损失保持原图像的身份不变; 然后在属性向量获取网络中构建头发特征与头发属性的映射模型, 得到属性向量; 最后将原图像特征与属性向量融合输入到合成网络, 生成最终结果. 在 FFHQ 数据集上进行了多种属性转移实验, 结果表明, 所提方法可以有效地转移头发属性, 生成高分辨率的结果. 大量在 Celeba 数据集上进行的实验结果表明, 与现有的主流属性转移方法相比, 所提方法可以取得更好的视觉效果.

关键词: 头发; 生成对抗网络; 特征融合; 属性转移
中图法分类号: TP391 **DOI:** 10.3724/SP.J.1089.2021.18544

Hair Attribute Transfer via Deep Feature Fusion

Xie Zhifeng[1,2], Su Xu[1], Liu Siwei[1], Zhang Guisong[1], and Ma Lizhuang[2,3]

[1] (*Department of Film and Television Engineering, Shanghai University, Shanghai* 200072)
[2] (*Shanghai Engineering Research Center of Motion Picture Special Effects, Shanghai* 200072)
[3] (*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai* 200240)

Abstract: To tackle the problem that existing attribute transfer methods can't transfer hair attributes effectively, a method of hair attribute transfer based on deep feature fusion is presented. This method includes three subnetworks which are responsible for feature extraction, attribute vector extraction and image synthesis. Firstly, feature extraction network extracts features from original images, and keeps the identity of original images unchanged by adding a reconstruction loss. At the same time, attribute vector extraction network constructs the mapping model of hair features and hair attributes, and generates the attribute vector. Finally, the synthesis network takes the fusion result of image features and the attribute vector as input, and generates final results. Various attribute transfer experiments on FFHQ show that the proposed method can effectively transfer hair attributes and generate high-resolution results. Experiments on Celeba show that the proposed method can achieve better visual quality than existing popular attribute transfer methods.

Key words: hair; generative adversarial networks; feature fusion; attribute transfer

---

收稿日期: 2020-08-05; 修回日期: 2020-09-22. **基金项目**: 国家自然科学基金(61303093); 上海市自然科学基金(19ZR1419100).
谢志峰(1982—), 男, 博士, 副教授, 博士生导师, 主要研究方向为图像处理、深度学习; **苏旭**(1998—), 男, 硕士研究生, 主要研究方向为深度学习、图像处理; **刘思维**(1997—), 女, 硕士研究生, 主要研究方向为深度学习、图像处理; **张桂菘**(1997—), 男, 硕士研究生, 主要研究方向为图像处理、深度学习; **马利庄**(1963—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为计算机图形学、图像处理.

三分靠长相，七分靠打扮，人的外貌受众多因素的影响．作为其中重要的组成部分，发型可以充分地展现一个人的精神面貌，体现一个人的人生态度．以发成型，美化妆容，合适的发型可以迅速提升一个人的整体气质．生成对抗网络的不断发展，改变了传统的图像处理方式，使得改变人们外貌变得十分轻松．Chang 等[1]提出的用于化妆和卸妆的非对称式转移网络可以快速地将指定参考图像的妆容转移到素颜图像上．Portenier 等[2]提出的基于草图的人脸图像编辑网络和 Jo 等[3]提出的基于素描和颜色的人脸编辑网络允许用户按照自己的意愿编辑人脸图像，大量属性转移方法[4-6]可以快速转移人脸属性．上述方法取得了不错的效果，但这些方法主要集中在面部区域．相比于面部区域，头发包含了更加丰富多样的变化：头发的卷曲程度、头发的长短以及是否有刘海等．另外，现有的相关数据集都专注于人脸属性标注，与头发相关的属性标注并不准确，数据集中图像的分辨率也比较低．上述因素导致头发属性转移的研究仍然是具有挑战的难题．

现有与头发相关的研究大部分都集中在头发区域分割[7-9]、三维头发重建[10-11]和头发仿真[12]等方面．市场上已经出现了一些可以变化发型的应用①，然而，这些应用需要借助预先设置好的参数，不能充分地满足用户需求．本文受基于风格的生成对抗网络[13]与风格编码器②的启发，提出了一种基于深度特征融合的头发属性转移方法．该方法通过构建头发特征向量与头发属性的映射模型，可以准确地转移图像中的头发属性，达到变换发型的效果．与多种主流属性转移方法的比较证明了该方法可以取得更好的结果．

1　相关工作

图像转移方法的本质是学习不同图像域之间的映射关系，头发属性转移可以看做图像转移的一个应用．近年来，图像转移方法已经取得了不错的进展．Isola 等[14]提出了一种基于条件生成对抗网络(conditional generative adversarial networks, cGAN)[15]的框架，它可以从成对的训练数据中学习图像域之间的映射关系．该方法需要大量的成对数据，这通常是难以获取的．基于无监督学习的属性转移方法解决了成对数据难以获取的问题．Liu 等[16]提出的无监督图像到图像转移网络(unsupervised image-to-image translation networks, UNIT)将变分自编码器与耦合生成对抗网络[17]结合在一起，它借助 2 个生成器共享参数学习图像域之间的联合分布．Zhu 等[18]提出的基于循环一致对抗损失的网络(cycle-consistent generative adversarial networks, CycleGAN)添加循环一致损失保持图像域中的关键属性，从而实现了图像在不同域之间转移的效果．

上述方法一次只能学习到 2 个不同图像域之间的映射关系，在处理头发属性转移这种多图像域任务时存在局限性．Perarnau 等[19]提出的可逆条件生成对抗网络(invertible cGAN, IcGAN)采用编码器生成图像的特征表示，使用 cGAN[15]解码基于目标属性的特征表示，该方法可用于多属性转移任务．但由于采用分开训练网络的策略，生成的图像质量比较差．随后，Lample 等[20]提出 FaderNetwork (fader network)，通过添加对抗性约束强制特征对属性保持不变，并以端到端的方式训练网络．上述 2 种方法都依赖于特征向量对属性的一致性．属性编辑网络(attribute editing generative adversarial networks, AttGAN)[4]和星形生成对抗网络(star generative adversarial networks, StarGAN)[5]把目标属性向量作为模型输入的一部分，通过对生成的图像应用属性分类约束转移特定的属性．这 2 种方法在人脸属性转移任务中取得了不错的结果，但把整个目标属性向量作为生成器的输入是没有必要的，并且可能会影响特定的属性转移效果．Liu 等[6]摒弃了将整个目标属性向量作为输入的思想，只把目标属性向量与原属性向量的差值作为输入，从而增加了属性转移的准确性．

2　头发属性转移

如图 1 所示，本文方法由 3 部分组成：图像合成网络、特征提取网络和属性向量获取网络．头发属性转移流程如下：首先，从特征提取网络提取原图像特征向量；然后，构建头发特征向量与头发属性的映射模型，得到可以控制头发属性转移的属性向量；最后，将原图像特征向量与属性向量一起送入合成网络，得到高质量的结果．

① https://hairstyles.knowage.info/

② https://github.com/Puzer/stylegan-encoder

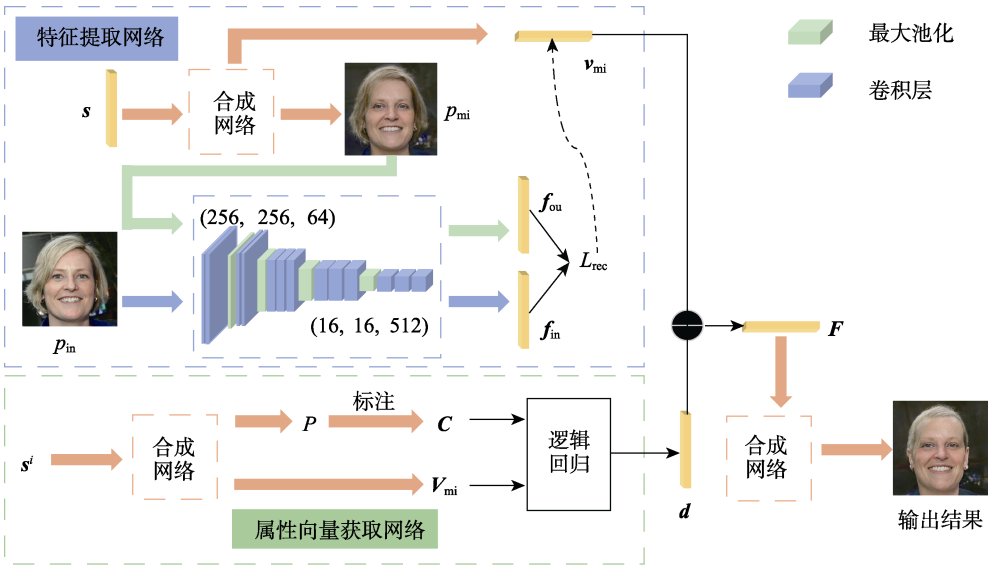

图 1 头发属性转移网络结构

2.1 图像合成网络

近年来，生成对抗网络(generative adversarial networks, GAN)[21]在图像合成中展现了其强大的性能[22-23]. 传统的 GAN 模型通过随机输入向量控制最终生成的图像，随着生成器层数的不断增加，输入向量的影响会逐渐减小. 基于风格的生成对抗网络(style-based GAN, StyleGAN)[13]采用常量作为生成器的原始输入，并将随机输入向量通过映射网络与仿射变换模块转换为“style”，然后将得到的 style 作用到生成器的每一个卷积层后面. 这里的 style 可以理解为一种属性. 多种属性共同作用并影响最终的生成图像.

如图 2 所示，本文的合成网络借鉴了这种思想. 给定一个随机初始向量 $\boldsymbol{s}$，映射网络 M 首先生成中间向量 $\boldsymbol{v}_{mi}$. 将 $\boldsymbol{v}_{mi}$ 输入到仿射变换模块 A 可以得到一个 style. M 由 8 个全连接层组成；多个 style 可以共同影响输出图像的样式，改变中间向量可以变换输出图像的样式. 因此，构建中间向量与输出图像头发属性的映射模型，可以达到转移头发属性的目的.

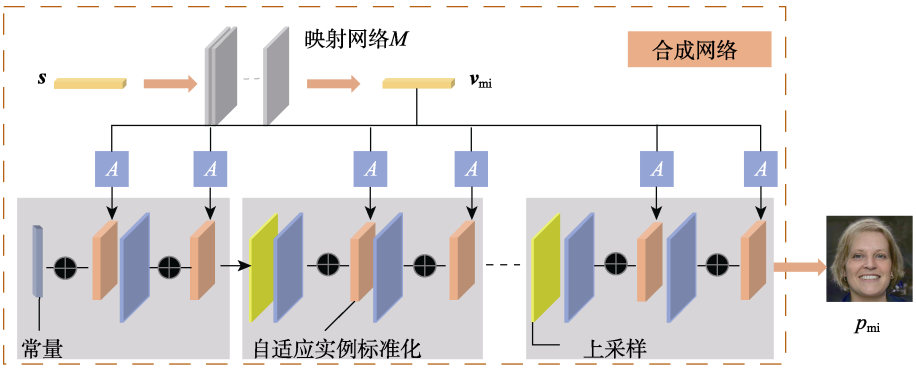

图 2 合成网络结构

本文共使用 18 个仿射变化模块，可以使输出图像的分辨率从 4×4 逐步提升到 1 024×1 024. $\boldsymbol{s}$ 的形状为 (1, 512)，$\boldsymbol{v}_{mi}$ 的形状为 (18, 512). 因此，合成网络也可以看作形状为 (1, 512) 的向量到分辨率为 (1 024, 1 024, 3) 的图像的映射网络.

FFHQ 数据集包含分辨率为 1 024×1 024 的头发图像 70 000 幅，并涵盖了基本的头发属性. 因此，本文使用该数据集训练图像合成网络. 由于使用预训练的模型，因此图 2 并没有展示鉴别器网络.

2.2 特征提取网络

传统的特征提取方法，如尺度不变特征变换、方向梯度直方图和局部二值模式等均依赖于烦琐的预处理操作，而且通常无法提取深度特征. 近年来，卷积神经网络凭借其强大的学习能力，在图像

特征提取方面取得了不错的结果．本文将卷积神经网络 VGG19[24]作为特征提取的骨干网络．

由于 VGG19 模型提取的特征向量不符合合成网络的生成域分布，并且不能直接解码为输入图像，因此不能将其直接作用于图像合成网络．如图 1 所示，本文融合 VGG19 模型与合成网络构建新的特征提取网络，并通过重建损失保留原图像的身份特征．

给定一幅输入图像 p_{in}，VGG19 模型首先提取其特征，作为参考向量 $\boldsymbol{f}_{\text{in}}$．同时，随机初始化 10 000 个形状为 $(1, 512)$ 的向量，将其平均值赋值给初始向量 $\boldsymbol{s}$．将 $\boldsymbol{s}$ 输入到合成网络，得到中间向量 $\boldsymbol{v}_{\text{mi}}$ 和中间图像 p_{mi}；$\boldsymbol{v}_{\text{mi}}$ 的形状为 $(18, 512)$，p_{mi} 形状为 $(1\,024, 1\,024, 3)$．然后 VGG19 模型提取 p_{mi} 的特征作为待优化向量 $\boldsymbol{f}_{\text{ou}}$．本文的思想是：如果待优化向量 $\boldsymbol{f}_{\text{ou}}$ 与参考向量 $\boldsymbol{f}_{\text{in}}$ 相似，那么此时的中间图像 p_{mi} 和输入图像 p_{in} 一致．为了约束 $\boldsymbol{f}_{\text{in}}$ 与 $\boldsymbol{f}_{\text{ou}}$ 保持一致，本文提出重构损失

$$L_{\text{rec}} = \left(F\left(S(\boldsymbol{s})\right) - F\left(p_{\text{in}}\right)\right)^2 .$$

其中，$\boldsymbol{s}$ 表示随机初始向量，对于任意输入图像，$\boldsymbol{s}$ 被重新赋值并优化；$S(\boldsymbol{s})$ 表示向量经过合成网络后得到的输出图像，即中间图像 p_{mi}；$F(p)$ 表示将图像 p 输入到 VGG19 模型中第 16 层的输出特征向量，即 $F(p_{\text{in}}) \Rightarrow \boldsymbol{f}_{\text{in}}$，$F(p_{\text{mi}}) \Rightarrow \boldsymbol{f}_{\text{ou}}$．当损失达到最优时，中间图像 p_{mi} 和输入图像 p_{in} 达到一致．相较于 $\boldsymbol{f}_{\text{in}}$，$\boldsymbol{v}_{\text{mi}}$ 源于合成网络，可直接作用于合成网络并解码为 p_{in}．因此，本文将中间向量 $\boldsymbol{v}_{\text{mi}}$ 视为 p_{in} 的特征向量．

2.3　属性向量获取网络

如第 2.1 和第 2.2 节所述，本文的合成网络可以视为连接特征向量与对应的生成图像的桥梁，而生成图像具有一些头发属性，如刘海、长短和曲直等．通过探究这些属性的变化在特征向量域变化的关系，可以变换特征向量实现控制头发属性转移的目的．

本文用属性向量表示属性的变化在特征向量域变化的关系，即 $\boldsymbol{V}_{\text{mi}} = \{S(\boldsymbol{s}^i)\} = \{\boldsymbol{v}_{\text{mi}}^i\}, i = 1, \cdots, n$．

首先将一组随机向量 $\boldsymbol{s}^i$ 输入合成网络模型中，得到一组中间向量 $\boldsymbol{v}_{\text{mi}}^i$ 和一组图像 $P = \{p^i\}$．$\boldsymbol{v}_{\text{mi}}^i$ 和 p^i 是一一对应的关系．其中，i 表示输入样本的序号，n 表示随机向量的个数，且 n 的值为 10 007．中间向量的形状为 $(18, 512)$．如第 2.2 节所述，$\boldsymbol{v}_{\text{mi}}^i$ 可以视为图像 p^i 的头发特征向量．

为了拟合头发特征向量与头发属性，人工为生成的每一幅图像 p^i 标注头发属性．得到向量 $\boldsymbol{c}^i = \{c_j^i\}_{j=1}^5$．

图像可以由一个包含 5 种头发属性值的向量表示．j 表示属性序号．c_1^i 表示属性“男性”；c_2^i 表示属性“长发”；c_3^i 表示属性“直发”；c_4^i 表示属性“有刘海”；c_5^i 表示属性“光头”．这些属性涵盖了头发的各种外形变化．当图像 p^i 具有属性 c_j^i 时，图像对应的 c_j^i =1，否则 c_j^i =0．例如，$\boldsymbol{c}^i = [1, 0, 1, 1, 0]$ 可以表示 p^i 是一个留着短直发而且有刘海的男性；$\boldsymbol{c}^i = [0, 1, 0, 0, 0]$ 可以表示 p^i 是一个留着长卷发而且没有刘海的女性．另外，鉴于光头的特殊性，当 $c_5^i = 1$ 时不再考虑其他头发属性的取值．标注工作完成后，可以得到一组属性向量 $\boldsymbol{C} = \{\boldsymbol{c}^i\}, i = 1, \cdots, n$．

本文以长发属性 c_2^i 为例，说明如何获取属性向量．此时，根据 c_2^i 取值的不同，可以将 10 007 个特征向量分为 2 类：c_2^i=1 表示此时的特征向量对应的生成图像为长发；c_2^i=0 表示此时的特征向量对应的生成图像为短发．通过拟合特征向量与长发属性的取值(0/1)，可以得到长发属性变化在特征向量域变化的关系，称之为属性向量．

相比基于神经网络的方法，逻辑回归算法占用内存小且耗时少．本文采用逻辑回归算法拟合头发特征向量与头发属性值．首先将中间向量 $\boldsymbol{v}_{\text{mi}}^i$ 转换为 9 216 维向量 $\boldsymbol{v}_{\text{mi}}^i = [v_1^i, v_2^i, \cdots, v_k^i]$．其中，$k$ 表示中间向量的维度，k 的值为 9 216．对于 $\boldsymbol{v}_{\text{mi}}^i$，$c_2^i$ 的预测值为 $y^i = 1/(1 + \mathrm{e}^{-\boldsymbol{z}^i})$．其中，

$$\boldsymbol{z}^i = \boldsymbol{W}^{\mathrm{T}} * \boldsymbol{v}_{\text{mi}}^i = w_0 + w_1 v_1^i + \cdots + w_k v_k^i .$$

使用梯度下降的方法拟合全部特征向量的预测值与实际值，得到优化后的权重矩阵 $\boldsymbol{W}$．将 $\boldsymbol{W}$ 转换为 $(18, 512)$ 的向量 $\boldsymbol{d}$，即属性向量．该向量可以控制输出图像的长度属性转移．

以此类推，一共可以得到 4 个属性向量，分别控制 4 种头发属性转移．属性向量不依赖于输入图像 p_{in}，只与 $\boldsymbol{V}_{\text{mi}}$ 和 $\boldsymbol{C}$ 有关．因此，对于任意给定的输入图像，都可以使用这 4 个属性向量．此外，对每种头发属性单独处理，增加了头发属性转移的多样性．本文方法既可以独立控制某一种头发属性转移，也可以融合多个属性向量，实现多属性转移任务．

3 实验及结果分析

3.1 参数配置，效率分析与数据集

本文实验基于 TensorFlow1.10 和 Keras2.2 框架．实验环境为 Ubuntu16.04.6, Intel Xeon E5-2609CPU@1.7 GHz, 2 个 TITAN Xp 显卡．

本文的图像合成网络采用预训练的模型，使用自适应矩估计优化器(adaptive moment estimation, Adam)训练，β_1 和 β_2 的取值分别为 0.9 和 0.999，学习率设置为 0.002，使用 ReLU 作为激活函数，取 α=0.2．本文的特征提取模块采用 VGG19 作为骨干网络，迭代次数为 1 000，学习率为 1.

在应用中，给定任意一幅输入图像，本文方法可以分为 2 个阶段: (1) 提取输入图像的特征向量; (2) 将图像的特征向量与第 2.3 节描述的属性向量融合，并输入合成网络．由于合成网络是预训练模型，并且属性向量也是预先得到的，本文方法只需训练特征提取网络，即阶段(1)．对于每幅输入图像，阶段(1)均需要重新初始化并优化 $\boldsymbol{s}$ 得到图像的特征向量 $\boldsymbol{v}_{\mathrm{mi}}$，训练过程如第 2.2 节所述．对于每幅输入图像，阶段(1)平均需要花费 2 min．尽管效率不高，多种实验结果依然可以证明本文方法的有效性．

本文使用 FFHQ 数据集训练合成网络，其包括分辨率为 1 024×1 024 像素的 70 000 幅高清人脸图像．在与现有方法进行比较时，使用 Celeba 数据集．该数据集包含 202 599 幅名人的面部图像，每幅图像都带有 40 个二进制属性，其中包括了曲直、刘海和光头等头发相关属性．在测试现有方法时，图像由初始分辨率 178×218 被裁剪为 178×178，然后被调整为 1 024×1 024．用 Celeba 数据集测试本文方法时，图像大小被调整为 1 024×1 024．

3.2 头发属性转移结果

图 3 展示了本文方法在 FFHQ 数据集上的单头发属性转移结果，第 1 列为原图像；第 2 列为中间图，即重构图像；第 3~7 列分别代表一种头发属性转移结果．其中，当原图像没有刘海时，第 5 列表示添加刘海，如第 1 行和第 3 行的结果所示．当原图像有刘海时，第 5 列表示去掉刘海，如第 2 行和第 4 行的结果所示．由于正确地构建了头发特征向量与每一种头发属性的映射模型，本文方法可以精确地转移每一种头发属性．如第 1 行第 4 列的结果，原图像是短发而且没有刘海，而将属性转换为长发后得到结果仍然是没有刘海的．再比如第 2 行第 7 列的结果，原图像是有刘海的直发，将属性转换为卷发之后得到的结果依然是有刘海的．

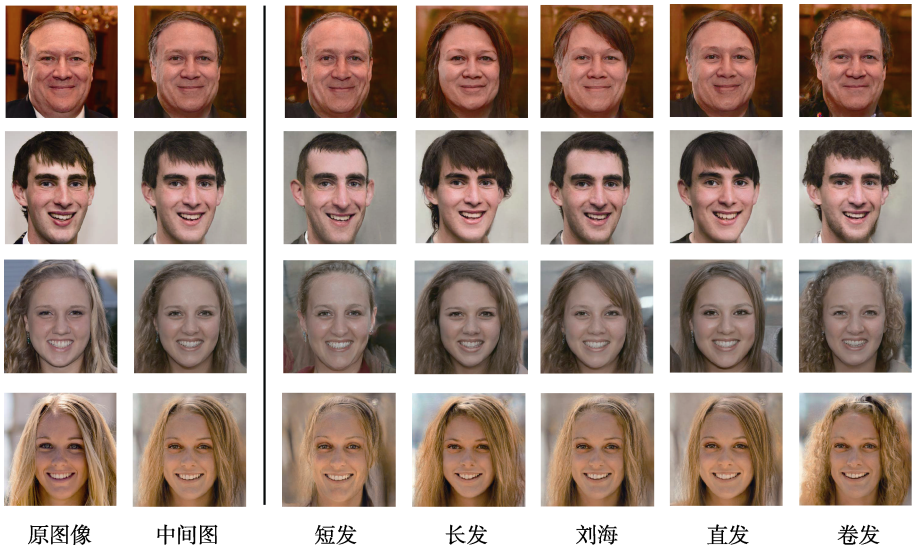

图 3 在 FFHQ 数据集上的单属性转移结果

图 4 展示了本文方法在 FFHQ 数据集上的多头发属性转移结果．第 1 列为原图像；第 2 列和第 3 列为单属性转移结果；第 4 列为多属性转移结果．从第 1 行前 4 列所展示的结果可以观察到，原图像是有刘海，将属性转移为长发后依然是没有刘海的．同时将头发属性转移为长发和有刘海后得到的结果既是长发又有刘海．从第 2 行前 4 列所展示的结果可以观察到，原图像是长直发，在将头发属性转移为短发之后依然是直发，同时将头发属性转移为短发和卷发得到的结果为卷曲的短发．图 4 中的结果充分地说明了本文方法不仅可以精确地转移单一头发属性，并且可以同时转移多种头发属性．

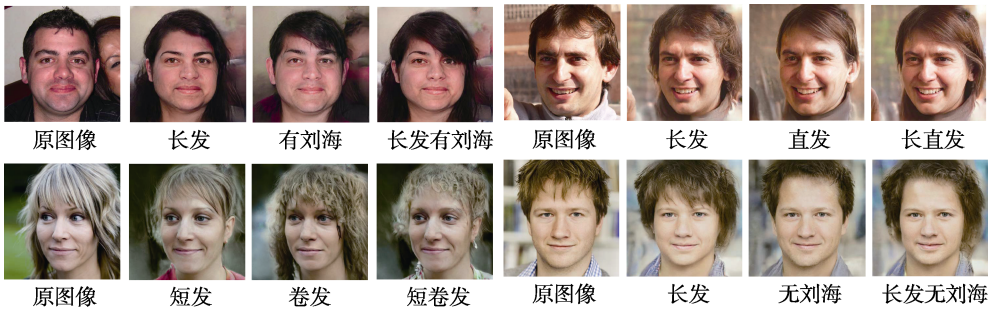

图 4　在 FFHQ 数据集上的多属性转移结果

3.3　定性评估

为了证明本文方法的有效性，分别将其与主流的 3 种属性转移方法 AttGAN[4], StarGAN[5]和选择性属性转移网络(selective transfer GAN, STGAN)[6]进行比较. 图 5 展示了 4 种方法在 Celeba 数据集上的头发属性转移结果. 可以看到，现有的 3 种属性转移方法都无法准确地转移相应的头发属性，无法生成真实的结果. 本文方法可以准确地转移每一种头发属性，生成高质量的结果.

3.4　定量评估

为了定量评估本文方法，本文进行了一项用户调查研究，一共考虑 4 种头发属性转移任务. 在每种属性转移任务中，向 100 余名参与者按随机顺序展示由 4 种方法生成的结果，每一个参与者依据 3 个标准选出效果最好的方法：生成结果是否能够保持原图像的身份特征、是否可以精确地转移对应属性、是否可以生成真实自然的结果.

表 1 展示了 4 种方法在每种属性转移任务的得票百分比，每一列的结果相加为 100%. 从表 1 中可以看到本文方法在 4 种头发属性转移任务上都获得了更多选票，进一步证明了本文方法的有效性.

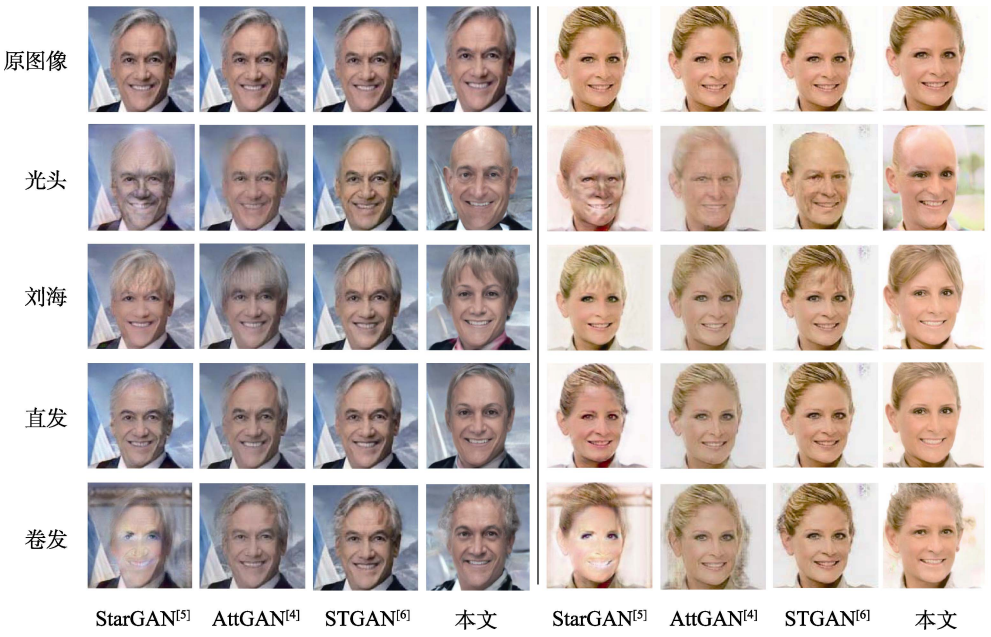

图 5　4 种方法在 Celeba 数据集的头发属性转移结果

表 1　不同方法在各属性转移任务的用户调查得票　%

方法	属性			
	光头	刘海	直发	卷发
AttGAN[4]	11.11	27.78	16.67	22.22
StarGAN[5]	11.11	11.11	27.78	5.56
STGAN[6]	16.67	5.56	22.22	22.22
本文	61.11	55.55	33.33	50.00

3.5　属性转移强度控制

得到原图像特征 $\boldsymbol{v}_{\mathrm{mi}}$ 以及属性向量 $\boldsymbol{d}$ 后，将两者相加的结果 $\boldsymbol{F}$ 送入合成网络，生成最终结果. $\boldsymbol{F}$ 的计算式为 $\boldsymbol{F}=\boldsymbol{v}_{\mathrm{mi}}+\lambda\times\boldsymbol{d}$.

本文通过为属性向量添加权重 λ 控制属性转移的强度. 图 6 展示了本文方法在 FFHQ 数据集上的属性转移强度控制结果，其中的数值为 λ 的取

值. 当 λ=0 时，生成的图像相当于原图像的重构图像. 从图 6 可以观察到，随着 λ 的取值发生变化，属性转移的强度也随之变化. 从第 2 行的结果可以观察到，当 $\lambda>0$ 时，随着 $|\lambda|$ 值的增加，头发越来越直；当 $\lambda<0$ 时，随着 $|\lambda|$ 值的增加，头发越来越卷曲. 由此可见，$\lambda>0$ 与 $\lambda<0$ 分别代表属性转移的 2 种不同方向，并且属性转移的程度与 $|\lambda|$ 的取值呈正相关关系.

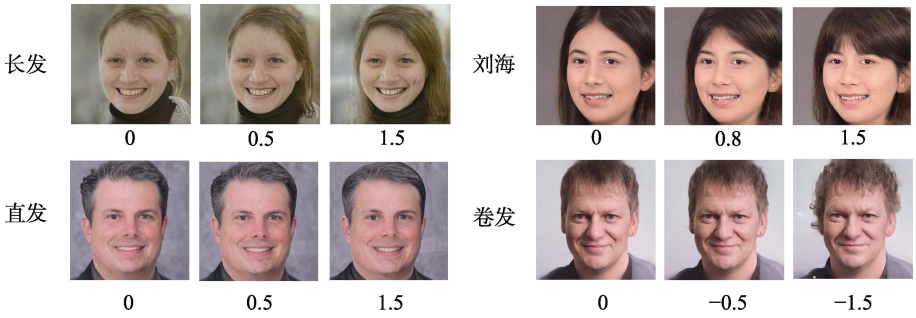

图 6 在 FFHQ 数据集上的属性转移强度控制结果

4 结 语

本文提出了一种基于深度特征融合的头发属性转移方法，该方法可以有效地转移头发属性，生成具有真实感且高质量的结果. 与多种主流属性转移方法的比较结果证明了本文方法的优越性. 但本文方法仍然存在可以优化的地方：原图像的身份特征保持得并不是很好，线性方法拟合头发特征向量与头发属性存在一定的局限性. 在未来的工作中将尝试添加更多约束保持原图像的身份特征，同时通过使用非线性的方法进一步改进生成结果.

参考文献(References):

[1] Chang H W, Lu J W, Yu F, *et al*. PairedCycleGAN: asymmetric style transfer for applying and removing makeup[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 40-48

[2] Portenier T, Hu Q Y, Szabó A, *et al*. FaceShop: deep sketch-based face image editing[J]. ACM Transactions on Graphics, 2018, 37(4): Article No.99

[3] Jo Y J, Park J Y. SC-FEGAN: face editing generative adversarial network with user's sketch and color[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2019: 1745-1753

[4] He Z L, Zuo W M, Kan M N, *et al*. AttGAN: facial attribute editing by only changing what you want[J]. IEEE Transactions on Image Processing, 2019, 28(11): 5464-5478

[5] Choi Y J, Choi M J, Kim M Y, *et al*. StarGAN: unified generative adversarial networks for multi-domain image-to-image translation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 18-23

[6] Liu M, Ding Y K, Xia M, *et al*. STGAN: a unified selective transfer network for arbitrary image attribute editing[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 15-22

[7] Shih H C. An unsupervised hair segmentation and counting system in microscopy images[J]. IEEE Sensors Journal, 2015, 15(6): 3565-3572

[8] Muhammad U R, Svanera M, Leonardi R, *et al*. Hair detection, segmentation, and hairstyle classification in the wild[J]. Image and Vision Computing, 2018, 71: 25-37

[9] Wang Dan, Shan Shiguang, Zhang Hongming, *et al*. Coarse-to-fine hair segmentation method[J]. Journal of Software, 2013, 24(10): 2391-2404(in Chinese)
(王丹，山世光，张洪明，等. 一种由粗到细的头发分割方法[J]. 软件学报, 2013, 24(10): 2391-2404)

[10] Hu L W, Ma C Y, Luo L J, *et al*. Single-view hair modeling using a hairstyle database[J]. ACM Transactions on Graphics, 2015, 34(4): Article No.125

[11] Zhou Y, Hu L W, Xing J, *et al*. HairNet: single-view hair reconstruction using convolutional neural networks[C] //Proceedings of European Conference on Computer Vision. Heidelberg: Springer, 2018: 249-265

[12] Zhang Pian, Hou Jin. Real-time hair simulation and rendering based on fast Gaussian kernel interpolation[J]. Journal of Computer-Aided Design & Computer Graphics, 2017, 29(2): 320-327(in Chinese)
(张骅，侯进. 高斯核函数快速插值的头发实时仿真与渲染[J]. 计算机辅助设计与图形学学报, 2017, 29(2): 320-327)

[13] Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 4396-4405

[14] Isola P, Zhu J Y, Zhou T H, *et al*. Image-to-image translation with conditional adversarial networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Los Alamitos: IEEE Computer Society Press, 2017: 5967-5976

[15] Mirza M, Osindero S. Conditional generative adversarial nets[OL]. [2020-08-05]. https://arxiv.org/abs/1411.1784

[16] Liu M Y, Breuel T, Kautz J. Unsupervised image-to-image translation networks[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 700-708

[17] Liu M Y, Tuzel O. Coupled generative adversarial networks[C] //Proceedings of the 30th International Conference on Neural Information Processing Systems. New York: ACM Press, 2016: 469-477

[18] Zhu J Y, Park T, Isola P, *et al*. Unpaired image-to-image translation using cycle-consistent adversarial networks[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2242-2251

[19] Perarnau G, Weijer J V D, Raducanu B, *et al*. Invertible conditional GANs for image editing[OL]. [2020-08-05]. https://arxiv.org/abs/1611.06355

[20] Lample G, Zeghidour N, Usunier N, *et al*. Fader networks: manipulating images by sliding attributes[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 5969-5978

[21] Goodfellow I, Pouget-Abadie J, Mirza M, *et al*. Generative adversarial nets[C] //Proceedings of the 27th International Conference on Neural Information Processing Systems. New York: ACM Press, 2014: 2672-2680

[22] Huang Yuan, Bai Cong, Li Hongkai, *et al*. Image captioning based on conditional generative adversarial nets[J]. Journal of Computer-Aided Design & Computer Graphics, 2020, 32(6): 911-918(in Chinese)
(黄远, 白琮, 李宏凯, 等. 基于条件生成对抗网络的图像描述生成方法[J]. 计算机辅助设计与图形学学报, 2020, 32(6): 911-918)

[23] He Jijun, Shen Yuan, Guo Yutang, *et al*. Head area extraction and portrait synthesis method using GAN[J]. Journal of Computer-Aided Design & Computer Graphics, 2020, 32(4): 599-605(in Chinese)
(何冀军, 申远, 郭玉堂, 等. 用于人像提取及半身像合成的生成对抗网络算法[J]. 计算机辅助设计与图形学学报, 2020, 32(4): 599-605)

[24] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[OL]. [2020-08-05]. https://arxiv.org/abs/1409.1556