



(12) 发明专利申请

(10) 申请公布号 CN 111915693 A
(43) 申请公布日 2020. 11. 10

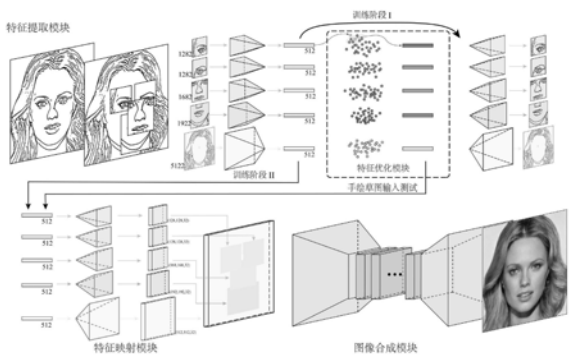
(21) 申请号 202010439641.2
(22) 申请日 2020.05.22
(71) 申请人 中国科学院计算技术研究所
地址 100080 北京市海淀区中关村科学院南路6号
申请人 香港城市大学
(72) 发明人 高林 傅红波 苏万超
(74) 专利代理机构 北京律诚同业知识产权代理有限公司 11006
代理人 祁建国

(51) Int. Cl .
G06T 11/00 (2006.01)
G06N 3/04 (2006.01)
G06N 3/08 (2006.01)
G06T 7/11 (2017.01)

权利要求书3页 说明书11页 附图6页

(54) 发明名称
基于草图的人脸图像生成方法及系统

(57) 摘要
本发明提出一种基于草图的人脸图像生成方法及系统,包括:通过特征提取网络提取手绘草图中人脸的多个部分,得到多个特征向量,将每个该特征向量在流形空间中的合理表达作为优化向量;利用特征映射网络,将优化向量解码并映射为特征张量,拼合所有特征张量得到完整的人脸特征张量,并利用图像合成网络,将该人脸特征张量,合成为高真实感的人脸图像。本发明的目的是解决上述现有技术的输入图像质量要求高,生成结果真实感较差的问题,提出了一种基于草图交互的高质量人脸图像生成方法及系统,该方法可以从用户手绘草图图像对应完整高真实感的人脸图像。



1. 一种基于草图的人脸图像生成方法,其特征在于,包括:

步骤1、通过特征提取网络提取手绘草图中人脸的多个部分,得到多个特征向量,将每个该特征向量在流形空间中的合理表达作为优化向量;

步骤2、利用特征映射网络,将优化向量解码并映射为特征张量,拼合所有特征张量得到完整的人脸特征张量,并利用图像合成网络将该人脸特征张量合成为人脸特征图像。

2. 如权利要求1所述的基于草图的人脸图像生成方法,其特征在于,该步骤1包括:

步骤11、对该手绘草图中人脸图像采用重叠窗口设置左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域,并通过反选左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域得到其他区域,对每个区域 c 单独提取特征 f_s^c ;

步骤12、以每个区域的特征作为训练数据,训练自动编码网络,得到编码器模型 E^c ,该自动编码网络包括多层解码器和多层编码器,并在编码器和解码器中间设有全连接层,且在编码器和解码器的每个卷积或反卷积操作之后添加一个残差块来构建隐含层描述子。

3. 如权利要求2所述的基于草图的人脸图像生成方法,其特征在于,该步骤1包括:

步骤13、构造用于训练的草稿图像数据集 $s = \{s_i\}$ 到隐含层的训练集合,对该训练集合中每张图片进行特征提取,对于某个区域 c ,利用该编码器模型 E^c 构建该训练组成的隐含层特征集合 $F^c = \{f_i^c = E^c(s_i^c)\}$,该隐含层特征集合 F^c 所有的特征点分布在一个低维的流形空间 M^c 内;

步骤14、当输入手绘草图 \tilde{s} 后,使用该编码器模型 E^c 提取到对应区域 c 的特征向量 $f_s^c = E_c(\tilde{s}^c)$,采用检索和插值的方法将第 c 个部分的特征向量 $E_c(\tilde{s}^c)$ 投影至流形空间 M^c 上,表示为 $f_{\tilde{s}}^c$ 到 M^c ;

其中检索和插值的方法具体包括:

步骤141、对区域 c 的特征向量 f_s^c ,通过计算特征向量之间的欧式距离,来检索到在训练集到特征集合 F^c 中最相近的 K 个样例 $\mathcal{K}^c = \{s_k^c\}, (s_k \subset S)$ 表示 K 个最近样例的集合,来表示 \tilde{s}^c 在流形空间 M^c 的相邻特征;

步骤142、通过下式以最小化问题求解插值权重:

$$\min \left\| f_{\tilde{s}}^c - \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c \right\|_2^2, \quad \text{s.t.} \quad \sum_{k \in \mathcal{K}^c} w_k^c = 1$$

其中 w_k^c 是样本 s_k^c 的未知权重,通过求解单个分量的约束最小二乘,可以求出各分量的权重;

步骤143、给定求解的权重 $\{w_k^c\}$, \tilde{s}^c 在流形空间 M^c 上的投影点可以表示为:

$$f_{\text{proj}}^c = \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c$$

其中 f_{proj}^c 是 \tilde{s}^c 的该优化向量,用于图像合成。

4. 如权利要求3所述的基于草图的人脸图像生成方法,其特征在于,该步骤2包括:

步骤21、将隐含层空间中的该优化向量映射到多通道特征中,生成三维特征张量,将每个区域的三维特征张量根据手绘草图 \tilde{s} 中人脸组件的确切位置进行拼合,得到完整的人脸特征张量。

5. 如权利要求4所述的基于草图的人脸图像生成方法,其特征在于,还包括:

步骤3、当给定当前的绘画草图,将该草稿图像数据集中与该绘画草图 k 个最相似的草图图像加权叠加在一起后以30%透明度的阴影图形式显示在画板上,当用户进行绘画时,实时更新该阴影图。

6. 一种基于草图的人脸图像生成系统,其特征在于,包括:

模块1、通过特征提取网络提取手绘草图中人脸的多个部分,得到多个特征向量,将每个该特征向量在流形空间中的合理表达作为优化向量;

模块2、利用特征映射网络,将优化向量解码并映射为特征张量,拼合所有特征张量得到完整的人脸特征张量,并利用图像合成网络将该人脸特征张量合成为人脸特征图像。

7. 如权利要求6所述的基于草图的人脸图像生成系统,其特征在于,该模块1包括:

模块11、对该手绘草图中人脸图像采用重叠窗口设置左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域,并通过反选左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域得到其他区域,对每个区域 c 单独提取特征 f_s^c ;

模块12、以每个区域的特征作为训练数据,训练自动编码网络,得到编码器模型 E^c ,该自动编码网络包括多层解码器和多层编码器,并在编码器和解码器中间设有全连接层,且在编码器和解码器的每个卷积或反卷积操作之后添加一个残差块来构建隐含层描述子。

8. 如权利要求7所述的基于草图的人脸图像生成系统,其特征在于,该模块1包括:

模块13、构造用于训练的草稿图像数据集 $S = \{s_i\}$ 到隐含层的训练集合,对该训练集合中每张图片进行特征提取,对于某个区域 c ,利用该编码器模型 E^c 构建该训练组成的隐含层特征集合 $F^c = \{f_i^c = E^c(s_i^c)\}$,该隐含层特征集合 F^c 所有的特征点分布在一个低维的流形空间 M^c 内;

模块14、当输入手绘草图 \tilde{s} 后,使用该编码器模型 E^c 提取到对应区域 c 的特征向量 $f_s^c = E_c(\tilde{s}^c)$,采用检索和插值的方法将第 c 个部分的特征向量 $E_c(\tilde{s}^c)$ 投影至流形空间 M^c 上,表示为 $f_{\tilde{s}}^c$ 到 M^c ;

其中检索和插值的系统具体包括:

模块141、对区域 c 的特征向量 f_s^c ,通过计算特征向量之间的欧式距离,来检索到在训练集到特征集合 F^c 中最相近的 K 个样例 $\mathcal{K}^c = \{s_k^c\}, (s_k \subset S)$ 表示 K 个最近样例的集合,来表示 \tilde{s}^c 在流形空间 M^c 的相邻特征;

模块142、通过下式以最小化问题求解插值权重:

$$\min \left\| f_s^c - \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c \right\|_2^2, \quad \text{s.t.} \quad \sum_{k \in \mathcal{K}^c} w_k^c = 1$$

其中 w_k^c 是样本 s_k^c 的未知权重,通过求解单个分量的约束最小二乘,可以求出各分量的权重;

模块143、给定求解的权重 $\{w_k^c\}$, \tilde{s}^c 在流形空间 M^c 上的投影点可以表示为:

$$f_{\text{proj}}^c = \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c$$

其中 f_{proj}^c 是 \tilde{s}^c 的该优化向量,用于图像合成。

9.如权利要求8所述的基于草图的人脸图像生成系统,其特征在于,该模块2包括:

模块21、将隐含层空间中的该优化向量映射到多通道特征中,生成三维特征张量,将每个区域的三维特征张量根据手绘草图 \tilde{s} 中人脸组件的确切位置进行拼合,得到完整的人脸特征张量。

10.如权利要求9所述的基于草图的人脸图像生成系统,其特征在于,还包括:

模块3、当给定当前的绘画草图,将该草稿图像数据集中与该绘画草图 K 个最相似的草图图像加权叠加在一起后以30%透明度的阴影图形式显示在画板上,当用户进行绘画时,实时更新该阴影图。

基于草图的人脸图像生成方法及系统

技术领域

[0001] 本发明涉及计算机图形学和计算机视觉,尤其涉及对人脸面部图像合成和草图编辑的方法。

背景技术

[0002] 随着深度学习的发展,视觉领域中图像编辑和翻译的工作越来越多,从手绘草图图像到真实人脸的迁移工作逐渐出现。然而,现有的解决方法大多基于专业的草稿图像或者从真实图像中提取到的轮廓图作为输入,在普通用户手绘出的草稿图像上表现不佳,目前尚无一面部图像合成的交互式系统出现。从草图图像生成真实人脸图像的技术在实际生活中有着广泛的应用,可用于犯罪调查,人物设计等各个方面。在上述应用中,现有研究往往采用图像到图像的迁移方法,然而这种方法约束要尽可能的刻画人物形象的细节,当刻画的细节不够细致时往往很难产生真实的人脸图像,因此限制了使用用户的范围。

[0003] 鉴于草图简单和易使用性,草图图像常被用来描绘物体或面孔图像,现有技术可以实现草图图像对人脸图像编辑的操作,但当只输入草图图像时却很难生成高质量人脸图像。经典的网络模型可以实现草图图像到真实图像的翻译,然而当输入为粗糙的或不完整的手绘草图时,往往难以生成真实结果。现有技术可以对草图补全并生成真实图像,对简单的物体如菠萝,草莓,等简单的物体表现良好,但是迁移至人脸图像生成上表现欠佳。

发明内容

[0004] 为了解决手绘草图图像到真实人脸图像的生成和交互工作,本发明设计一基于草图的交互式系统。利用深度神经网络、数据先验,以及局部-全局处理的优化思路,将人脸图像进行剖分,对子区域进行优化处理后整合生成高质量的人脸图像。本发明的目的是解决上述现有技术的输入图像质量要求高,生成结果真实感较差的问题,提出了一种基于草图交互的高质量人脸图像生成方法及系统,该方法可以从用户手绘草图图像对应完整高真实感的人脸图像。并为该方法设计一使用软件,以阴影图的形式引导/帮助用户进行绘画创作,进而生成对应的真实图像。

[0005] 针对现有技术的不足,本发明提出一种基于草图的人脸图像生成方法,其中包括:

[0006] 步骤1、通过特征提取网络提取手绘草图中人脸的多个部分,得到多个特征向量,将每个该特征向量在流形空间中的合理表达作为优化向量;

[0007] 步骤2、利用特征映射网络,将优化向量解码并映射为特征张量,拼合所有特征张量得到完整的人脸特征张量,并利用图像合成网络将该人脸特征张量合成为人脸特征图像。

[0008] 所述的基于草图的人脸图像生成方法,其中该步骤1包括:

[0009] 步骤11、对该手绘草图中人脸图像采用重叠窗口设置左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域,并通过反选左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域得到其他区域,对每个区域 c 单独提取特征 f_s^c ;

[0010] 步骤12、以每个区域的特征作为训练数据,训练自动编码网络,得到编码器模型 E^c ,该自动编码网络包括多层解码器和多层编码器,并在编码器和解码器中间设有全连接层,且在编码器和解码器的每个卷积或反卷积操作之后添加一个残差块来构建隐含层描述子。

[0011] 所述的基于草图的人脸图像生成方法,其中该步骤1包括:

[0012] 步骤13、构造用于训练的草稿图像数据集 $S = \{s_i\}$ 到隐含层的训练集合,对该训练集合中每张图片进行特征提取,对于某个区域 c ,利用该编码器模型 E^c 构建该训练组成的隐含层特征集合 $F^c = \{f_i^c = E^c(s_i^c)\}$,该隐含层特征集合 F^c 所有的特征点分布在一个低维的流形空间 M^c 内;

[0013] 步骤14、当输入手绘草图 \tilde{s} 后,使用该编码器模型 E^c 提取到对应区域 c 的特征向量 $f_s^c = E_c(\tilde{s}^c)$,采用检索和插值的方法将第 c 个部分的特征向量 $E_c(\tilde{s}^c)$ 投影至流形空间 M^c 上,表示为 f_s^c 到 M^c ;

[0014] 其中检索和插值的方法具体包括:

[0015] 步骤141、对区域 c 的特征向量 f_s^c ,通过计算特征向量之间的欧式距离,来检索到在训练集到特征集合 F^c 中最相近的 K 个样例 $\mathcal{K}^c = \{s_k^c\}, (s_k \subset S)$ 表示 K 个最近样例的集合,来表示 \tilde{s}^c 在流形空间 M^c 的相邻特征;

[0016] 步骤142、通过下式以最小化问题求解插值权重:

$$[0017] \quad \min \left\| f_s^c - \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c \right\|_2^2, \quad \text{s.t.} \quad \sum_{k \in \mathcal{K}^c} w_k^c = 1$$

[0018] 其中 w_k^c 是样本 s_k^c 的未知权重,通过求解单个分量的约束最小二乘,可以求出各分量的权重;

[0019] 步骤143、给定求解的权重 $\{w_k^c\}$, \tilde{s}^c 在流形空间 M^c 上的投影点可以表示为:

$$[0020] \quad f_{\text{proj}}^c = \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c$$

[0021] 其中 f_{proj}^c 是 \tilde{s}^c 的该优化向量,用于图像合成。

[0022] 所述的基于草图的人脸图像生成方法,其中该步骤2包括:

[0023] 步骤21、将隐含层空间中的该优化向量映射到多通道特征中,生成三维特征张量,将每个区域的三维特征张量根据手绘草图 \tilde{s} 中人脸组件的确切位置进行拼合,得到完整的人脸特征张量。

[0024] 所述的基于草图的人脸图像生成方法,其中还包括:

[0025] 步骤3、当给定当前的绘画草图,将该草稿图像数据集中与该绘画草图 K 个最相似的草图图像加权叠加在一起后以30%透明度的阴影图形式显示在画板上,当用户进行绘画时,实时更新该阴影图。

[0026] 本发明还提出了一种基于草图的人脸图像生成系统,其中包括:

[0027] 模块1、通过特征提取网络提取手绘草图中人脸的多个部分,得到多个特征向量,将每个该特征向量在流形空间中的合理表达作为优化向量;

[0028] 模块2、利用特征映射网络,将优化向量解码并映射为特征张量,拼合所有特征张量得到完整的人脸特征张量,并利用图像合成网络将该人脸特征张量合成为人脸特征图像。

[0029] 所述的基于草图的人脸图像生成系统,其中该模块1包括:

[0030] 模块11、对该手绘草图中人脸图像采用重叠窗口设置左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域,并通过反选左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域得到其他区域,对每个区域c单独提取特征 f_s^c ;

[0031] 模块12、以每个区域的特征作为训练数据,训练自动编码网络,得到编码器模型 E^c ,该自动编码网络包括多层解码器和多层编码器,并在编码器和解码器中间设有全连接层,且在编码器和解码器的每个卷积或反卷积操作之后添加一个残差块来构建隐含层描述子。

[0032] 所述的基于草图的人脸图像生成系统,其中该模块1包括:

[0033] 模块13、构造用于训练的草稿图像数据集 $S = \{s_i\}$ 到隐含层的训练集合,对该训练集合中每张图片进行特征提取,对于某个区域c,利用该编码器模型 E^c 构建该训练组成的隐含层特征集合 $F^c = \{f_i^c = E^c(s_i^c)\}$,该隐含层特征集合 F^c 所有的特征点分布在一个低维的流形空间 M^c 内;

[0034] 模块14、当输入手绘草图 \tilde{s} 后,使用该编码器模型 E^c 提取到对应区域c的特征向量 $f_s^c = E_c(\tilde{s}^c)$,采用检索和插值的方法将第c个部分的特征向量 $E_c(\tilde{s}^c)$ 投影至流形空间 M^c 上,表示为 $f_{\tilde{s}}^c$ 到 M^c ;

[0035] 其中检索和插值的系统具体包括:

[0036] 模块141、对区域c的特征向量 f_s^c ,通过计算特征向量之间的欧式距离,来检索到在训练集到特征集合 F^c 中最相近的K个样例 $\mathcal{K}^c = \{s_k^c\}, (s_k \subset S)$ 表示K个最近样例的集合,来表示 \tilde{s}^c 在流形空间 M^c 的相邻特征;

[0037] 模块142、通过下式以最小化问题求解插值权重:

$$[0038] \quad \min \left\| f_{\tilde{s}}^c - \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c \right\|_2^2, \quad \text{s.t.} \quad \sum_{k \in \mathcal{K}^c} w_k^c = 1$$

[0039] 其中 w_k^c 是样本 s_k^c 的未知权重,通过求解单个分量的约束最小二乘,可以求出各分量的权重;

[0040] 模块143、给定求解的权重 $\{w_k^c\}$, \tilde{s}^c 在流形空间 M^c 上的投影点可以表示为:

$$[0041] \quad f_{\text{proj}}^c = \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c$$

[0042] 其中 f_{proj}^c 是 \mathcal{S}^c 的该优化向量,用于图像合成。

[0043] 所述的基于草图的人脸图像生成系统,其中该模块2包括:

[0044] 模块21、将隐含层空间中的该优化向量映射到多通道特征中,生成三维特征张量,将每个区域的三维特征张量根据手绘草图 \mathcal{S} 中人脸组件的确切位置进行拼合,得到完整的人脸特征张量。

[0045] 所述的基于草图的人脸图像生成系统,其中还包括:

[0046] 模块3、当给定当前的绘画草图,将该草稿图像数据集中与该绘画草图 K 个最相似的草图图像加权叠加在一起后以30%透明度的阴影图形式显示在画板上,当用户进行绘画时,实时更新该阴影图。

[0047] 由以上方案可知,本发明的优点在于:

[0048] 本发明设计的系统能够实现任意手绘草图做输入,自动实现阴影图引导、高真实感人脸图像生成操作。

[0049] 系统流程图如图1所示,该系统涉及图像特征提取,特征空间的流形投影,特征空间特征插值,高质量图像生成和交互式界面设计等多项技术。

[0050] 如图2所示,展示了本发明在特征空间中查看相邻部件草图,在特征空间进行特征插值后的图像生成效果。图中第一和第五列图表示在特征空间中相邻的两张草图图像,中间三列图像为均匀插值的特征向量输入解码器中,解码生成的草图图像。从图示结果可以看出,在特征空间中进行特征插值可以得到平滑的图像过渡效果,由此可证明特征空间分布的连续性。

[0051] 如图3所示,展示了本发明在调节其他区域混合权重后所表现出的生成效果。左上为用户画出的草图图像,其余五张图像为设置其他区域中的不同混合权重参数得到的生成效果。从图中可以看出,随着混合权重系数的不断增大,生成图像与输入的对应性越来越高。

[0052] 如图4所示,展示了本发明调节多个区域混合权重后所表现的生成效果对比。第一幅图展示的是输入的草图图像,第二幅图像展示了输入草图直接生成的效果。第三幅图展示的是使用特征优化后的生成效果,以及第四张图展示的是对不同区域调整不同混合参数后对生成结果。从图中可以看出用户通过调整混合参数可以在优化结果和用户生成结果之间做一个平滑调整。

[0053] 如图5所示,展示了本发明在用户绘画时添加线条和细节编辑的效果。左侧部分的图像展示的是随着用户笔触的增多,线条约束增多,所生成的结果更加具象,同时生成的图像全局的光影也发生变化。右侧图像展示的是用户对细节调整的效果,可以看出我们的方法在细节控制上更加自由,局部细节线条的变化并不会影响其他区域的形状表达。

[0054] 如图6所示,展示了本发明的交互式界面设计,界面中有两个图像显示区域,绘图板“SketchingCanvas”和展示板“SynthesizedFace”,分别显示用户绘画的草图内容以及生成的图像效果。同时在界面右上有一个调节条,用户可以通过调节对应区域权重的大小来实现用户手绘草图对生成结果的影响强弱。同时界面上方有不同的控件,用户可以选择画

笔、橡皮擦并调整对应的大小。还可以通过选择“Realtime”和“Shadow”来决定是否打开实时生成功能和阴影引导功能。

[0055] 如图7所示,展示了本发明在用户实际使用的过程中,绘画板和展示板上不同的图像展示。其中在绘画板上,会有阴影引导,阴影图会根据用户的画笔实时更新。同时,生成的结果会在展示板实时展示,用户可以看到当前的生成效果。

[0056] 如图8所示,展示了本发明在实际使用时生成的结果,第一,三行为用户绘画的草图图像,第二、四行为对应的人脸图像生成结果。

[0057] 如图9所示,展示了本发明在人脸图像上进行插值应用时所生成的效果。左右为使用的两张图像,中间的图像为插值得到的人脸图像,可以看出本发明在处理人脸图像插值时可以得到平滑的变形序列。

[0058] 如图10所示,展示了本发明在人脸混合上生成的效果,将不同人脸的不同部位进行拼合,输入不同部位的草图图像,生成一张新的完整的人脸图像。

附图说明

[0059] 图1为本发明系统流程示意图;

[0060] 图2为不同部分在隐含层上的插值生成效果图;

[0061] 图3为在特征优化后,在其他区域设置不同的调节混合参数的效果展示图;

[0062] 图4为在特征优化后,对不同区域调节混合参数的效果展示图;

[0063] 图5为添加细节生成效果(左)和局部细节编辑生成效果(右)图;

[0064] 图6为交互式界面效果展示图;

[0065] 图7为操作界面上绘画板显示图和生成效果展示图;

[0066] 图8为生成效果展示图;

[0067] 图9为图像插值生成效果图;

[0068] 图10为人脸图像拼图合成效果展示图。

具体实施方式

[0069] 现有技术中由于没有考虑用户输入的图像质量参差不齐,未经训练的普通用户难以高效的绘制出人脸草图图像,输入草图信息表达不够充分导致生成的人脸图像真实感较差或无法生成人脸图像的情况产生;同时,普通用户在绘画时往往难以掌握面部的五官位置比例,在绘画创作中存在一定的困难。

[0070] 发明人对人脸的五官结构进行分析,考虑人脸五官分布的相似性,将人脸图像分为五个部分(左眼,右眼,鼻子,嘴巴和其他区域),每个区域单独处理可以更好的刻画局部细节。其目的是为了对每个区域单独编辑,防止局部编辑影响全局的效果。将面部区域分块后单独操作可以更好的表达局部细节。局部编辑的效果可以参考图5。

[0071] 当用户绘画出草图后,本发明会对每个部分单独提取特征,采用局部线性嵌入的方法(即LLE,LocallyLinearEmbedding)对特征进行优化,并利用优化后对特征来生成真实的人脸图像。同时,为了解决普通用户无法绘制出合理的人脸草图,本发明设计一个界面,在用户绘画时,当用户进行绘画时会根据用户所画出对草图来进行预测和推荐,并以阴影的形式放置与绘画板下,以此来引导用户进行绘画。考虑到用户的专业水平不同,本申请设

置有控制条,用户可以通过调整每个部分的权重来控制绘画草图对生成效果控制强弱。

[0072] 本发明提供一种草图交互的高质量人脸图像生成方法和系统,包括以下关键点:

[0073] 关键点1,人脸图像特征提取模块

[0074] 采取人脸图像分区域处理的方法,鉴于人脸的结构清晰,将人脸图像分为5个部分(左眼,右眼,鼻子,嘴巴和其他部分),每个区域图像单独提取特征;为了更好的控制每个部分的细节,学习到每个面部部分的局部特征。采用自动编码网络(即AE,auto-encoder)对每个区域数据单独训练,通过编码器来获取草图图像在隐含层中对应的特征向量。

[0075] 关键点2,特征优化模块

[0076] 考虑到草图表达的多样性,为了将提取到的特征优化至合理到流形空间中,利用特征投影模块将各部分投影到对应的流形分布空间上,使得优化后的特征向量符合数据的合理表达,从而实现将任意草图图像提取到的隐含层特征优化在合理的表达空间中。

[0077] 关键点3,特征映射模块

[0078] 鉴于每个子区域对特征提取操作是独立的,在获取到优化后的特征向量后,需要对每个子区域对隐含层特征单独解码来获取各自的特征张量。并采用一定的放回操作,实现子区域特征张量到完整的人脸特征图的合并操作。

[0079] 关键点4,图像合成模块

[0080] 在获取到融合后的特征张量后,图像合成模块将特征张量转换成真实的人脸图像。该模块采用条件生成对抗网络结构,以人脸特征图作为输入,通过对抗学习与图像约束来实现高质量人脸图像的生成。

[0081] 关键点5,交互界面的设计与显示

[0082] 交互式界面设计支持不同绘画水平的用户进行创作,可以实时生成绘画结果。在绘画过程中使用阴影图的形式对用户进行引导,以此来提供给用户更多的布局和细节引导。用户可以通过工具栏中的控件选取对应的功能,并通过调节栏调节控制区域约束的强弱,来调整生成效果。

[0083] 为让本发明的上述特征和效果能阐述的更明确易懂,下文特举实施例,并配合说明书附图作详细说明如下。

[0084] 如图1所示,本发明方法包括:

[0085] S1:输入用户手绘草图 $\tilde{\mathbf{s}}$;

[0086] S2:利用特征提取网络,对用户手绘草图图 $\tilde{\mathbf{s}}$ 提取得到每个部分,共c对特征向量 $\mathbf{f}_{\tilde{\mathbf{s}}}^c = E_c(\tilde{\mathbf{s}}^c)$;

[0087] S3:在特征流形空间内对特征向量 $\mathbf{f}_{\tilde{\mathbf{s}}}^c$ 进行投影优化,得到特征向量 $\mathbf{f}_{\tilde{\mathbf{s}}}^c$ 在流形空间 M^c 中的合理表达 $\mathbf{f}_{\text{proj}}^c$,合理指的是在真实数据分布的流形空间内找到一个最接近真实数据分布的表达;

[0088] S4:利用特征映射网络,将优化后的特征向量 $\mathbf{f}_{\text{proj}}^c$ 解码并映射为特征张量,并将每一部分的特征张量进行拼合得到完整的人脸特征张量;

[0089] S5:利用图像合成网络,输入人脸特征张量,生成高真实感的人脸图像,作为该草图的人脸图像生成结果。

[0090] S6:交互式界面设计与实现

[0091] 其中,S2中所述方法包括:

[0092] S21:本发明在获取用户手绘草图后,需要对图像提取特征。整张图像进行处理难以表达面部无关对多样性,为此采取人脸图像分区域处理的方法;鉴于人脸的结构清晰,本发明将人脸图像分为5个部分(左眼,右眼,鼻子,嘴巴和其他部分),对每个区域图像单独提取特征,记为 S^c , $c=\{1,2,3,4,5\}$;为了更好的处理部分之间的细节,本发明采用重叠窗口来设置前四个部分的分割区域(区域大小的选择根据数据库中人物对应部分的分布范围决定),对于其他部分的选取我们选择移除前四个部分的人脸图像。

[0093] S22:为了更好的控制每个部分的细节,学习到每个面部部分的局部特征,需要设计特征提取网络并进行训练,用于实时操作优化。其中,包括特征网络和训练数据处理两个部分。

[0094] S221:本发明采用自动编码网络(即AE,auto-encoder)对每个区域数据单独训练,在完成草稿图像到草稿图像的AE网络训练后,我们可以通过编码器 E_c 来获取草图图像在隐含层中对应的特征向量,通过解码器 D_c 来实现特征向量到草图图像的生成;对于自动编码网络的设计,本发明采用五层解码器和五层编码器的形式,并在编码器和解码器中间采用一个全连接层的形式,来确保每个部分生成的隐含层特征都是512维。对于隐含层特征的维度我们试验了不同维度的隐含层特征(128,256,512),经过实验我们发现对于草图细节重建和表达来说,512维度表达效果较好,低维的特征往往导致重建结果较为模糊。同时在反复实验后,我们在编码/解码层的每个卷积/反卷积操作之后添加一个残差块来构建隐含层描述子。网络层中添加残差块可以减少网络的信息损失,提升网络模型表达效果。

[0095] S222:对特征提取网络进行训练时,首先需要构建训练数据。鉴于目前无包含草图图像的真实人脸图像数据集,因此本发明在彩色图像上进行边界轮廓提取来模拟草图图像。对于草图图像数据集的构建,本发明采用PhotoShop进行线条提取操作后,采用草图简化的方法对线条进行精简操作。在对网络参数进行训练时,训练以自监督方式进行,输入的草图图像和重建图像之间的均方误差(MSE)损失,并使用Adam优化器对网络参数进行优化求解。

[0096] 其中S3中所述方法包括:

[0097] S31:考虑到用户输入的草图图像可能与真实的人脸图像存在出入,因此我们需要将其映射到真实的分布空间(流形空间)中。对于输入草图图像 s ,使用特征提取模块获取到每个部分 c 对应到特征向量 f_s^c ,鉴于草图表达具有随机性和多样性,图像所提取的特征可能并不符合数据集的特征分布。为了将提取到的特征优化至合理到流形空间中,实现即使绘画草图图像不完整也可以获取到一个合理的特征表达,本发明设计特征优化模块,可以将对任意草图图像提取隐含层特征并优化在合理的表达空间中。

[0098] S32:首先需要构造用于训练的草稿图像数据集 $S=\{s_i\}$ 到隐含层特征集合。将训练集合中的每张图片都进行特征提取,对于某个部分 c ,我们可以利用在特征提取模块中训练好的编码器模型 E^c 来构建训练集组成的隐含层特征集合 $F^c=\{f_i^c=E^c(s_i^c)\}$,图2展示的结果表明在特征空间上相似的草图图像在特征空间对分布上是相近的,因此我们认为所有的特征点 F^c 分布在一个低维的流形空间内,定义为 M^c 。

[0099] S33:当输入草图 $\tilde{\mathbf{s}}$ 后,使用预训练好的编码器 E^c 提取到对应部分 c 的特征向量,定义为 $\mathbf{f}_{\tilde{\mathbf{s}}}^c = E_c(\tilde{\mathbf{s}}^c)$ 。为了增加其表达一个人脸的真实性,在局部线性假设下,我们遵循LLE的思想,采用检索和插值的方法将第 c 个部分的特征向量 $E_c(\tilde{\mathbf{s}}^c)$ 投影至流行空间 M^c 上,表示为 $\mathbf{f}_{\tilde{\mathbf{s}}}^c$ 到 M^c 。其中包括最近点查找和特征插值两个部分:

[0100] S331:对部分 c 的特征向量 $\mathbf{f}_{\tilde{\mathbf{s}}}^c$,本发明首先通过计算特征向量之间的欧式距离,来查找到在训练集到特征集合 F^c 中最相近的 K 个样例。通过实验和对比发现当 $K=10$ 时,合成结果既可以保证表面真实性,同时保持适当的变化。定义 $\mathcal{K}^c = \{\mathbf{s}_k^c\}, (\mathbf{s}_k \subset S)$ 来表示 K 个最近样例的集合, $\{\mathbf{f}_k^c\}, k \in \mathcal{K}^c$ 来表示 $\tilde{\mathbf{s}}^c$ 在流形空间 M^c 的相邻特征。

[0101] S332:利用相邻特征 $\{\mathbf{f}_k^c\}, k \in \mathcal{K}^c$ 的线性组合来对特征向量 $\mathbf{f}_{\tilde{\mathbf{s}}}^c$ 进行重建表示,并通过最小化重建误差来进行权重参数求解。这相当于通过以下最小化问题求解插值权重:

$$[0102] \quad \min \left\| \mathbf{f}_{\tilde{\mathbf{s}}}^c - \sum_{k \in \mathcal{K}^c} w_k^c \cdot \mathbf{f}_k^c \right\|_2^2, \quad \text{s.t.} \quad \sum_{k \in \mathcal{K}^c} w_k^c = 1$$

[0103] 其中 w_k^c 是样本 \mathbf{s}_k^c 的未知权重。通过求解单个分量的约束最小二乘问题,可以求出各分量的权重。给定求解的权重 $\{w_k^c\}$, $\tilde{\mathbf{s}}^c$ 在流形空间 M^c 上的投影点可以表示为:

$$[0104] \quad \mathbf{f}_{\text{proj}}^c = \sum_{k \in \mathcal{K}^c} w_k^c \cdot \mathbf{f}_k^c$$

[0105] $\mathbf{f}_{\text{proj}}^c$ 是 $\tilde{\mathbf{s}}^c$ 的优化后的特征向量,可以传递到特征映射模块和图像合成模块进行图像合成。

[0106] 其中,S4中所述方法包括:

[0107] S41:给定输入草图图像 $\tilde{\mathbf{s}}$ 后,我们可以使用特征提取模块提取特征并利用特征投影模块将各部分投影到对应的空间分布流形上,使得优化后的特征向量在一个合理的表达空间中。当获取到每个部分的特征向量后,我们需要利用该特征来合成完整的人脸图像,如果直接将特征解码成草图图像进行拼合时,在区域衔接处会出现图片导致生成效果不佳。鉴于草稿图像只有一个通道,所以相邻组件在重叠区域的不一致性很难通过草图生成网络自动解决。这促使我们将隐含层空间中的特征向量映射到多通道特征中(即生成三维特征张量)。实验中发现这极大地改善了信息流,并且有效的融合了特征图而不是简单的草稿图像上的拼接,有助于解决各部分之间的不一致性。由于不同部分的特征具有不同的语义含义,因此我们对每个部分都设计了一个解码器模型,解决不同的部位提取到的高维特征空间分布存在差异的问题,从而实现从一个512维度的特征向量到一个三维特征张量的转换。每个解码器都有一个全连接层和五个解码层组成。解码得到的每个特征张量都有32个通道,并且与草图图像中对应部分的区域大小相同的尺寸。

[0108] S42:在获取到每个部分的张量表示后,我们需要将左眼、右眼、鼻子、嘴巴的结果

特征张量,根据输入人脸草图图像中人脸组件的确切位置,放回其他部分的特征张量中,以保留人脸组件之间的原始空间关系。在合并的时候,我们使用一个固定的合并顺序(即,左/右眼睛,鼻子,嘴巴,其他部分)来合并特征图。

[0109] 其中,S5中所述方法包括:

[0110] 在获取到S4中融合后的特征张量后,图像合成模块将特征张量转换成真实的人脸图像。该模块采用条件生成网络的结构,以特征张量作为条件输入,并在训练的时候使用判别器模型进行对抗训练。我们的生成网络包含一个编码器单元,残差块和解码器单元。在该模块中,我们使用多尺度判别器网络对不同尺度对生成图像进行判别。在训练时将输入图像和生成图像下采样到不同的图像大小,并利用对应图像大小的判别器来对图像进行判别。我们使用此来设置来隐式学习部件之间的高级关联。在网络训练的时候,采用Adam优化器进行参数求解优化,特征提取模块与图像合成模块网络共同训练,从而实现从特征向量到真实图片的生成。对于特征提取模块和图像合成模块,除了生成网络误差外,我们还加入真实图像与重建图像之间的L1误差以进一步引导生成器,从而确保生成图像的像素级质量。我们在鉴别器训练时,使用感知误差比较真实图像和生成图像之间的特征差异。

[0111] 其中,S6中所述方法包括:

[0112] S61:考虑到用户绘画水平到差异,普通用户可能难以准确定位到用户的五官位置。为了协助用户绘画,尤其是对绘画水平较低对用户,我们提供一种阴影引导的绘画界面,可以在用户绘画的过程中实时更新阴影图,以此来提供给用户更多的布局和细节引导。当给定当前的绘画图像 \tilde{S} ,使用S33中找到的训练集S中K个最相似的草图图像。找到最相似的K个图片后,将草图图像加权叠加在一起后,每张图片的权重为对S3中权重 $\{w_k^c\}$ 单位化后的数值,并将重叠后的图片以30%透明度的形式显示在画板上。当用户进行绘画时,阴影图会实时进行更新。合成的图片会显示在窗口右边。

[0113] S62:考虑到,绘画水平高的用户可以自行很好的绘制出人脸图像,用户可以通过控制上方shadow的控件来选择是否需要显示阴影图。当用户完成绘画或者想实时查看生成结果是,可以通过点击“Convert”控件或“Realtime”复选框来选择准换或者实时生成的功能。

[0114] S62:考虑到用户水平到差异可能导致对细节的表达差异,本发明设计一个调节条调节混合参数,通过调整原图像特征和优化特征之间的插值权重来实现用户控制每个部分绘画细节对生成结果影响的效果,增加绘图细节的可控性。定义 wb^c 表示部分c的混合参数,混合后的特征向量可以计算为:

$$[0115] \quad f_{\text{blend}}^c = wb^c \times f_s^c + (1 - wb^c) \times f_{\text{proj}}^c$$

[0116] 将 f_{blend}^c 输入S4和S5中,得到一个新的合成图像。

[0117] 以下为与上述方法实施例对应的系统实施例,本实施方式可与上述实施方式互相配合实施。上述实施方式中提到的相关技术细节在本实施方式中依然有效,为了减少重复,这里不再赘述。相应地,本实施方式中提到的相关技术细节也可应用在上述实施方式中。

[0118] 本发明还提出了一种基于草图的人脸图像生成系统,其中包括:

[0119] 模块1、通过特征提取网络提取手绘草图中人脸的多个部分,得到多个特征向量,将每个该特征向量在流形空间中的合理表达作为优化向量;

[0120] 模块2、利用特征映射网络,将优化向量解码并映射为特征张量,拼合所有特征张量得到完整的人脸特征张量,并利用图像合成网络将该人脸特征张量合成为人脸特征图像。

[0121] 所述的基于草图的人脸图像生成系统,其中该模块1包括:

[0122] 模块11、对该手绘草图中人脸图像采用重叠窗口设置左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域,并通过反选左眼区域、右眼区域、鼻子区域和嘴巴区域的分割区域得到其他区域,对每个区域c单独提取特征 f_s^c ;

[0123] 模块12、以每个区域的特征作为训练数据,训练自动编码网络,得到编码器模型 E^c ,该自动编码网络包括多层解码器和多层编码器,并在编码器和解码器中间设有全连接层,且在编码器和解码器的每个卷积或反卷积操作之后添加一个残差块来构建隐含层描述子。

[0124] 所述的基于草图的人脸图像生成系统,其中该模块1包括:

[0125] 模块13、构造用于训练的草稿图像数据集 $S = \{s_i\}$ 到隐含层的训练集合,对该训练集合中每张图片进行特征提取,对于某个区域c,利用该编码器模型 E^c 构建该训练组成的隐含层特征集合 $F^c = \{f_i^c = E^c(s_i^c)\}$,该隐含层特征集合 F^c 所有的特征点分布在一个低维的流形空间 M^c 内;

[0126] 模块14、当输入手绘草图 \tilde{s} 后,使用该编码器模型 E^c 提取到对应区域c的特征向量 $f_s^c = E_c(\tilde{s}^c)$,采用检索和插值的方法将第c个部分的特征向量 $E_c(\tilde{s}^c)$ 投影至流形空间 M^c 上,表示为 f_s^c 到 M^c ;

[0127] 其中检索和插值的系统具体包括:

[0128] 模块141、对区域c的特征向量 f_s^c ,通过计算特征向量之间的欧式距离,来检索到在训练集到特征集合 F^c 中最相近的K个样例 $\mathcal{K}^c = \{s_k^c\}, (s_k \subset S)$ 表示K个最近样例的集合,来表示 \tilde{s}^c 在流形空间 M^c 的相邻特征;

[0129] 模块142、通过下式以最小化问题求解插值权重:

$$[0130] \quad \min \left\| f_s^c - \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c \right\|_2^2, \quad \text{s.t.} \quad \sum_{k \in \mathcal{K}^c} w_k^c = 1$$

[0131] 其中 w_k^c 是样本 s_k^c 的未知权重,通过求解单个分量的约束最小二乘,可以求出各分量的权重;

[0132] 模块143、给定求解的权重 $\{w_k^c\}$, \tilde{s}^c 在流形空间 M^c 上的投影点可以表示为:

$$[0133] \quad f_{\text{proj}}^c = \sum_{k \in \mathcal{K}^c} w_k^c \cdot f_k^c$$

[0134] 其中 f_{proj}^c 是 \tilde{s}^c 的该优化向量,用于图像合成。

[0135] 所述的基于草图的人脸图像生成系统,其中该模块2包括:

[0136] 模块21、将隐含层空间中的该优化向量映射到多通道特征中,生成三维特征张量,将每个区域的三维特征张量根据手绘草图 \mathcal{S} 中人脸组件的确切位置进行拼合,得到完整的人脸特征张量。

[0137] 所述的基于草图的人脸图像生成系统,其中还包括:

[0138] 模块3、当给定当前的绘画草图,将该草稿图像数据集中与该绘画草图K个最相似的草图图像加权叠加在一起后以30%透明度的阴影图形式显示在画板上,当用户进行绘画时,实时更新该阴影图。

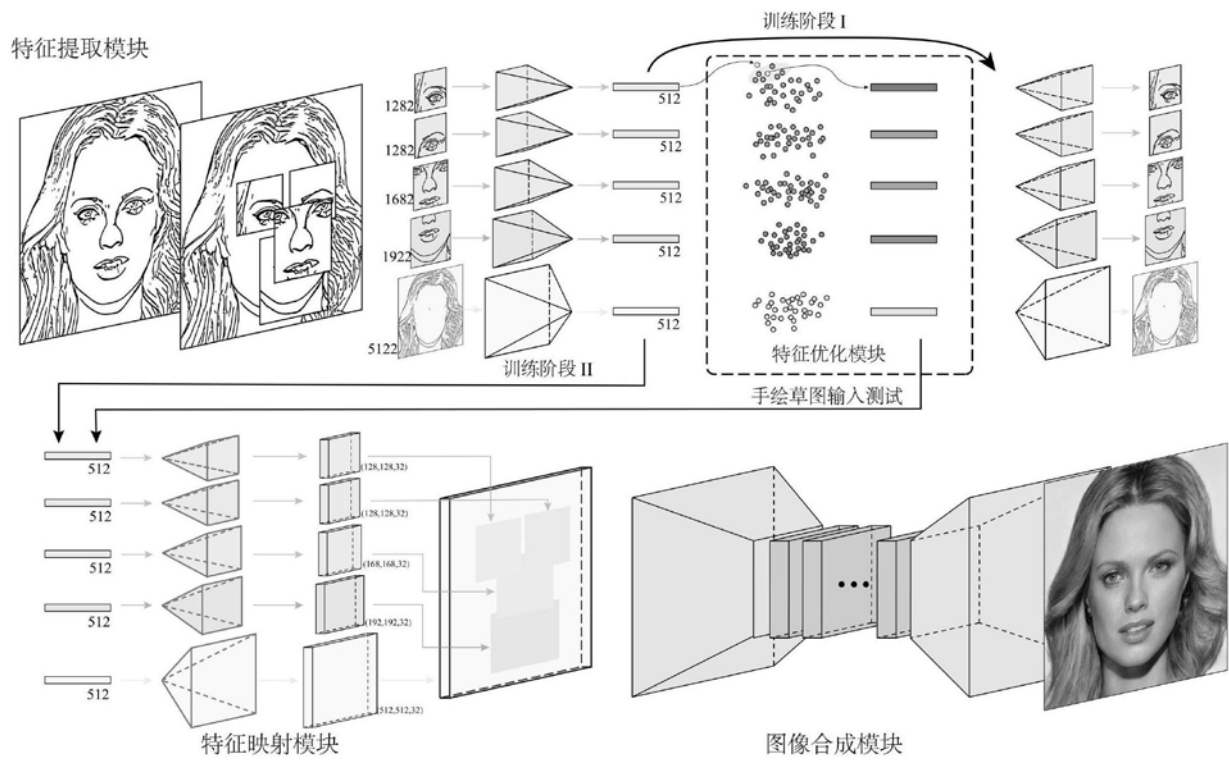


图1

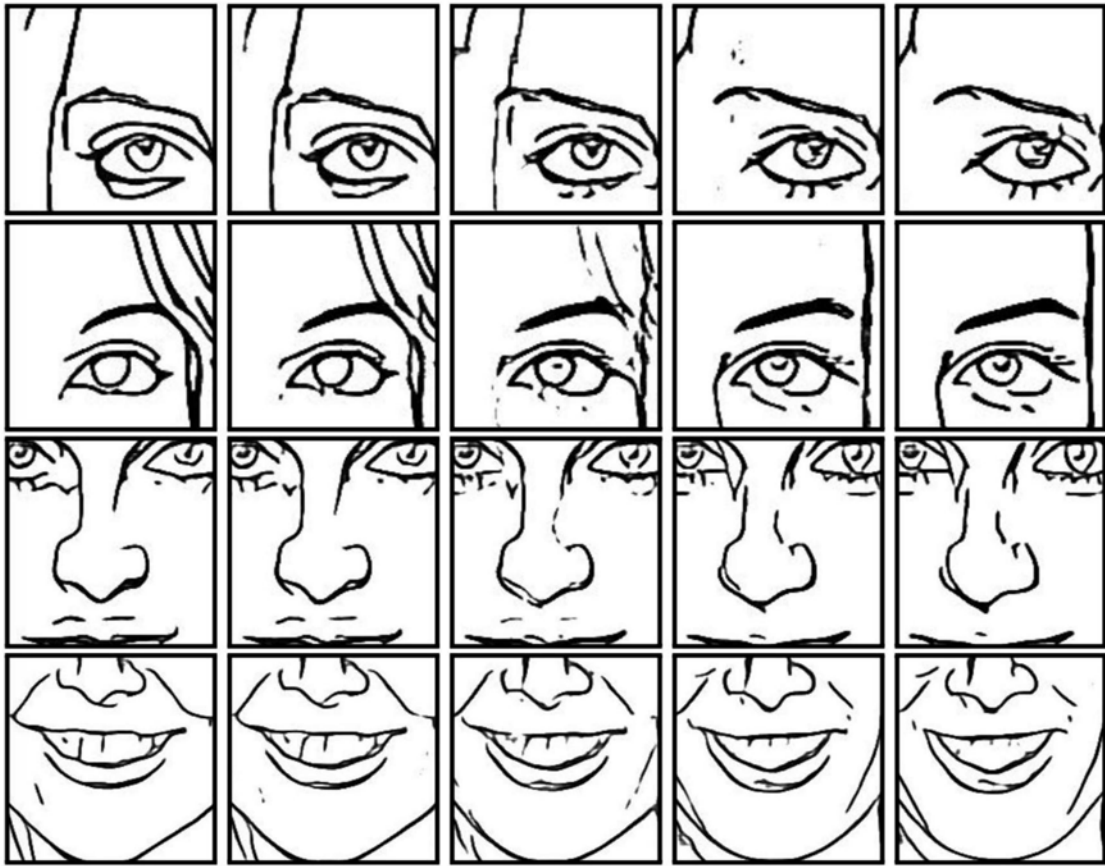


图2

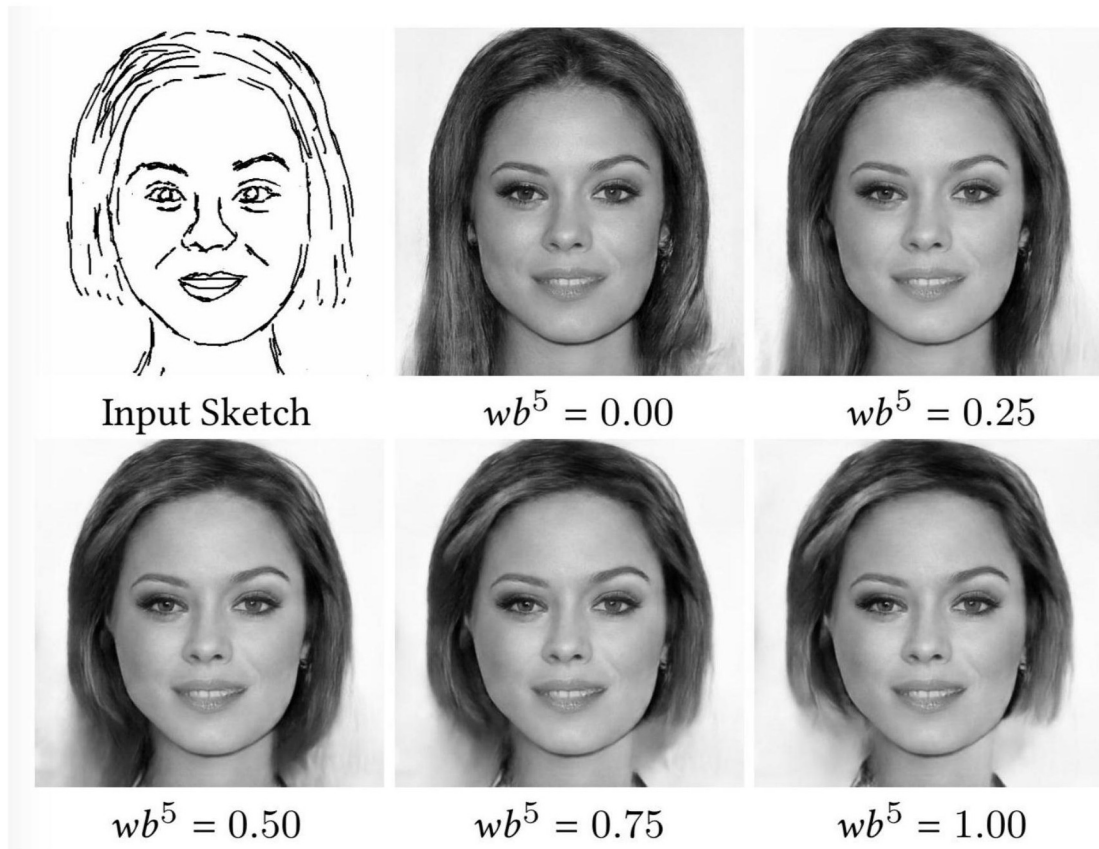


图3

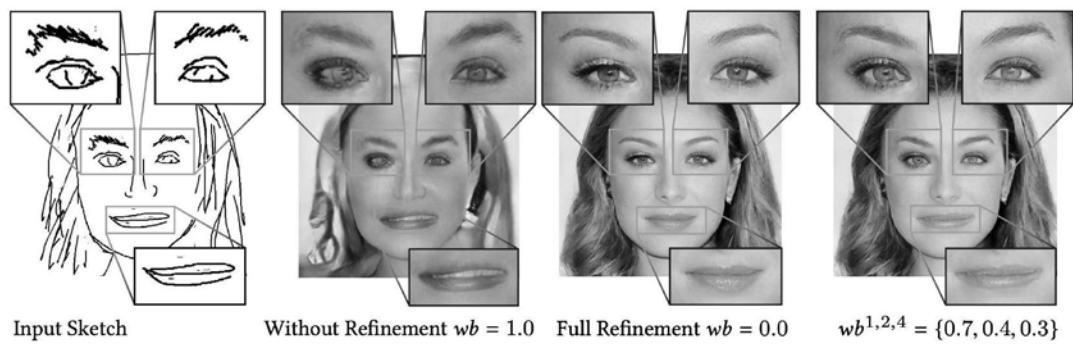


图4

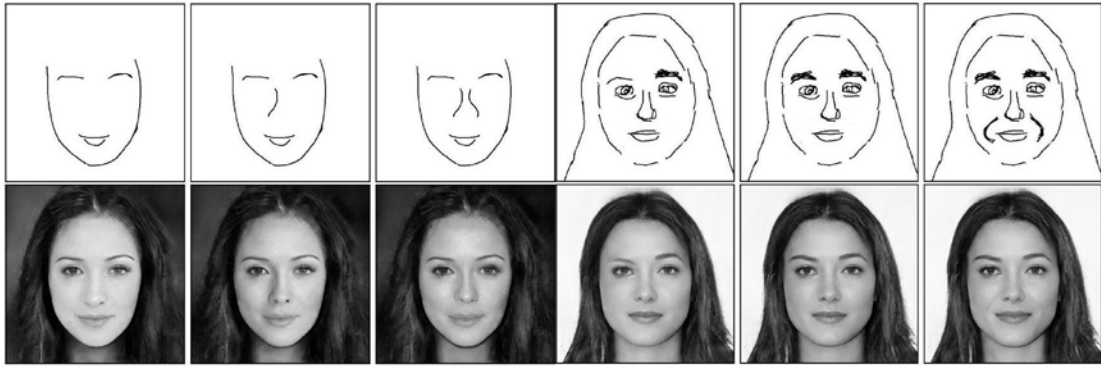


图5

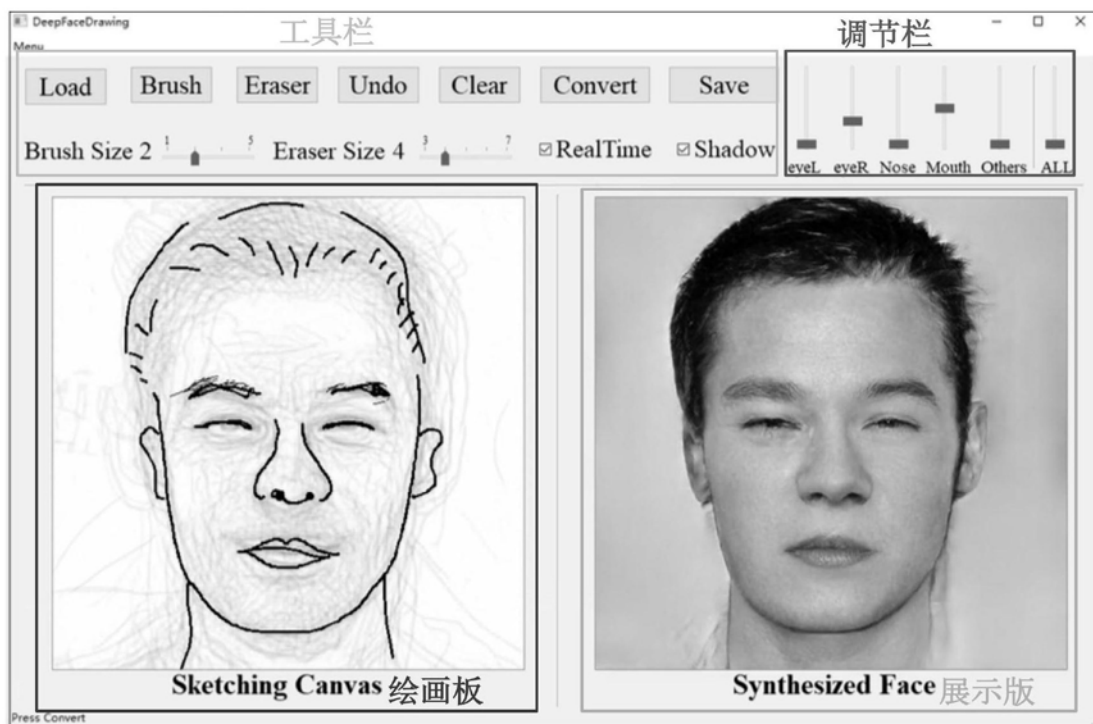


图6

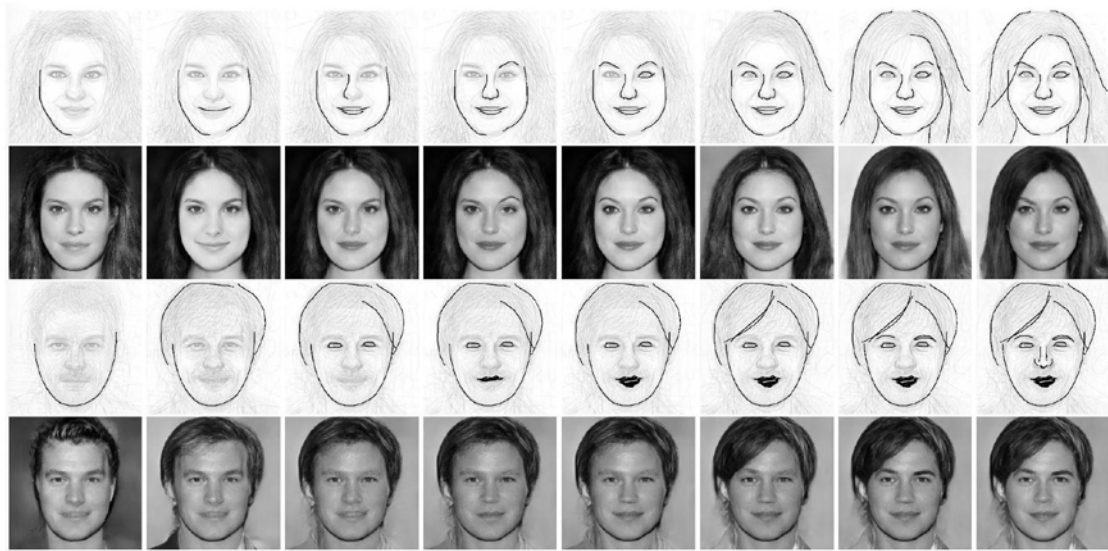


图7

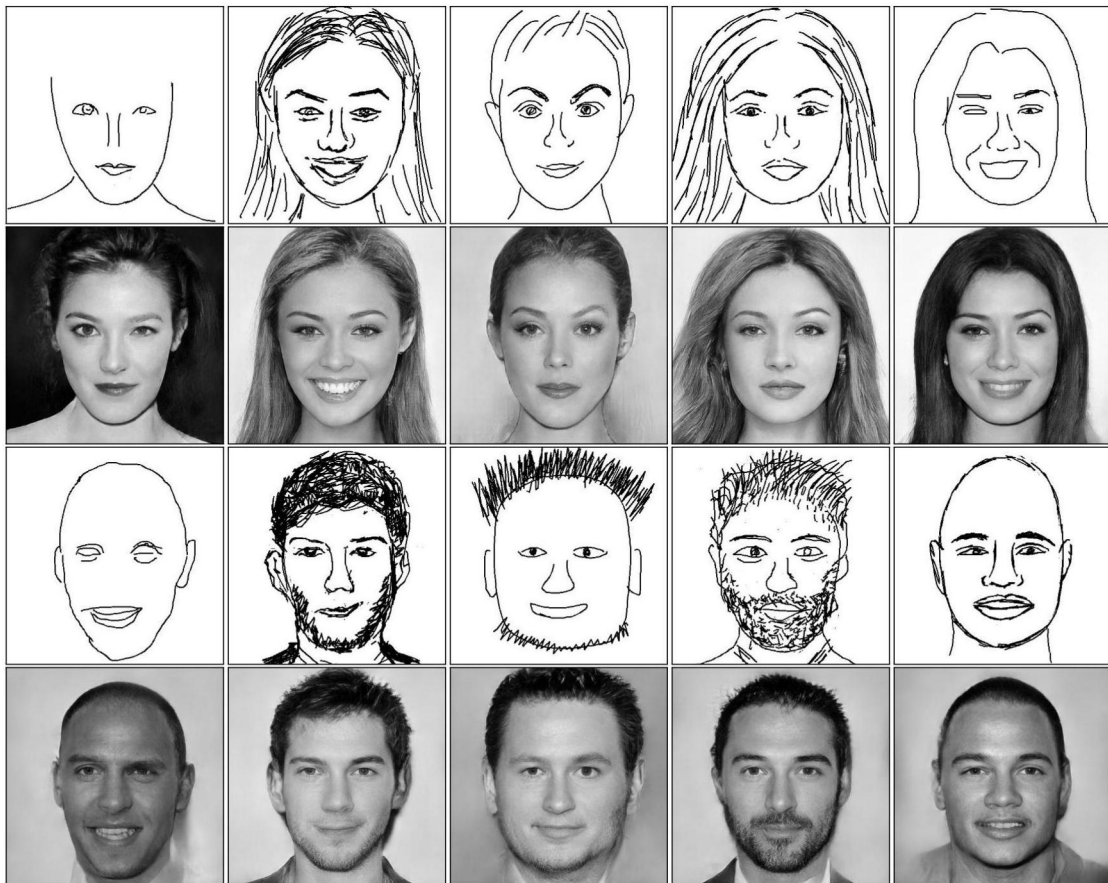


图8

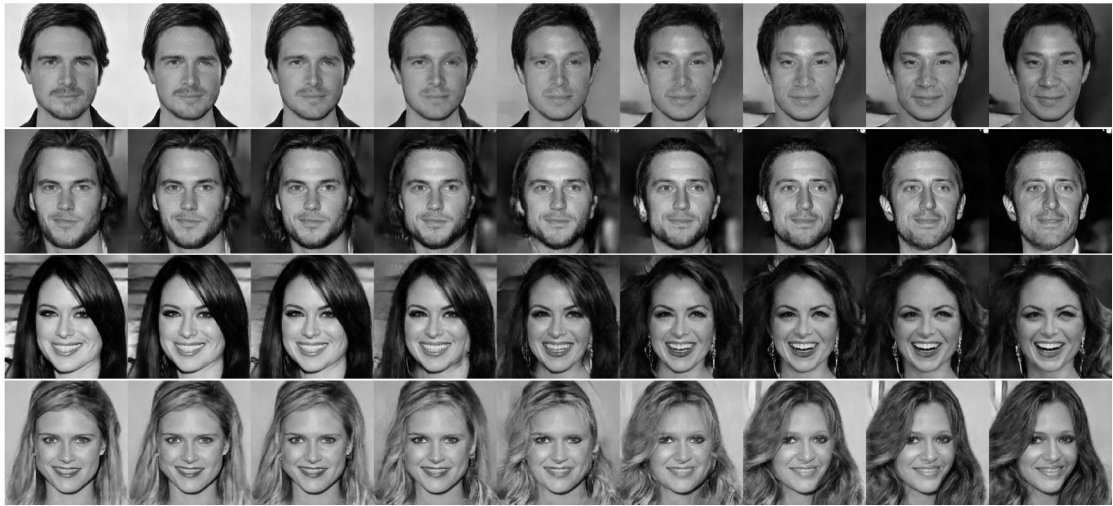


图9

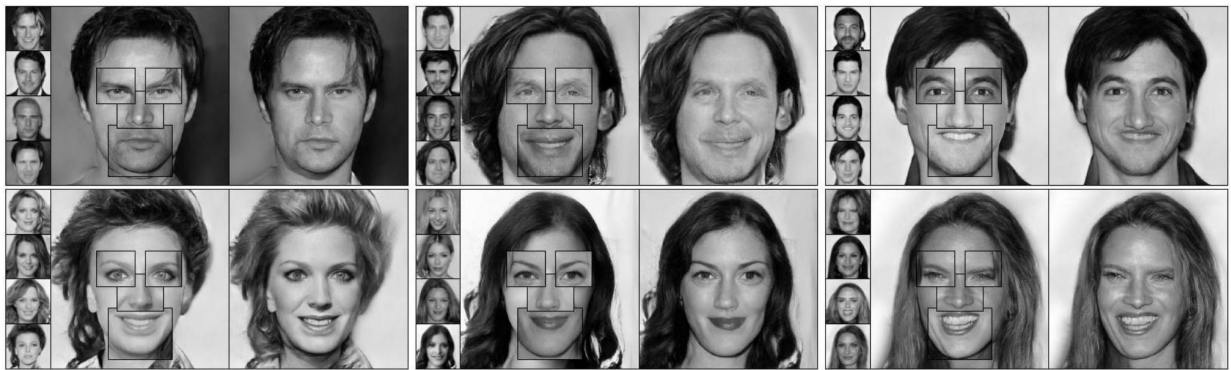


图10