

基于多模态深度学习的特定虚拟图像 视觉特征自动补偿

何伟^{1†}, 杨大伟¹, 马天福¹, 马崇瑞¹, 苑学贺²

(1. 国网新疆电力有限公司信息通信公司, 新疆 乌鲁木齐 830002; 2. 北京中电普华信息技术有限公司, 北京 100000)

摘 要: 为了提高特定虚拟图像视觉特征缺失补偿效果, 提出了基于多模态深度学习的特定虚拟图像视觉特征自动补偿方法。首先, 通过结合主成分分析(PCA)降维方法和结构张量的非局域全变分方法, 对图像实行冗余信息去除和去噪处理; 其次, 通过稀疏自编码器和卷积池化技术展开深度学习, 获取多模态下图像的彩色图、深度图、灰度图和 3D 曲面法线特征; 最后, 通过卷积神经网络完成特定虚拟图像视觉特征自动补偿。测试结果表明: 这种特征补偿方法能够通过预处理提高图像清晰度, 且图像视觉特征自动补偿质量较高。

关键词: 多模态; 特定虚拟图像; 视觉特征自动补偿; 图像去噪; 卷积神经网络
中图分类号: TP751 **文献标识码:** A

Automatic Compensation of Visual Features of Specific Virtual Images Based on Multimode Depth Learning

HE Wei^{1†}, YANG Dawei¹, MA Tianfu¹, MA Chongrui¹, YUAN Xuehe²

(1. State Grid Xinjiang Information and Telecommunication Company, Urumqi, Xinjiang 830002, China;
2. Beijing China-power Information Technology Co., Ltd., Beijing 100000, China)

Abstract: In order to improve the compensation effect for missing visual features of specific virtual images, a multimodal deep learning based automatic compensation method for visual features of specific virtual images is proposed. Firstly, by combining principal component analysis (PCA) dimensionality reduction method with structural tensor non local total variation method, redundant information is removed and denoised from the image. Secondly, deep learning is carried out through sparse autoencoder and convolutional pooling techniques to obtain color, depth, grayscale, and 3D surface normal features of images under multimodal conditions. Finally, automatic compensation of specific virtual image visual features is achieved through convolutional neural networks. The test results show that this feature compensation method can improve image clarity through preprocessing, and the quality of image visual feature automatic compensation is high.

Key words: multimode; specific virtual images; automatic compensation of visual features; image denoising; convolution neural networks

虚拟图像是一种与传统二维、三维图像不同的全新视觉形式, 用于人机交互界面。随着虚拟现实

技术的发展, 具有多变性、仿真性等审美特性的特定虚拟图像(如 VR、AR 图像等)逐步被广泛应用。

但是,由于特定虚拟图像的特征具有多变性,因此在传输过程中很容易丢失细节视觉特征。为解决这一问题,特定虚拟图像的特征补偿处理成为人们研究的新方向^[1-2]。

于淑侠等^[3]构建图像特征重建网络;提取图像特征信息并实行模型训练;将特征输入到建立的深度图像超分辨率重建网络完成图像视觉特征的自动补偿。但该方法受到图像噪声影响,存在图像补偿后清晰度仍然不高的问题。

王紫薇等^[4]分析了超分辨率图像重建的方法;基于卷积神经网络完成深度图像的超分辨率重建;通过卷积神经网络进一步获取彩色图像的边缘及深度特征信息,进一步优化图像纹理细节,最终完成图像视觉特征的自动补偿。但该方法未对图像去噪,补偿精度较差。

姜红涛等^[5]设计深度学习网络的外部梯度先验框架;结合先验完成图像的超分辨率重建模型建立;通过超分辨率重建模型完成图像视觉特征的自动补偿。但该方法使用深度学习网络提取的图像特征不完整,存在图像补偿精度不佳的问题。

为了解决上述方法中存在的问题,进一步提高虚拟图像补偿效果,提出基于多模态深度学习的特定虚拟图像视觉特征自动补偿方法,自动补偿特定虚拟视觉图像。

1 特定虚拟图像预处理

冗余信息和噪声干扰容易使虚拟图像视觉特征的非线性映射产生偏差,为了提高图像视觉特征补偿精度,需要对虚拟图像进行预处理。首先使用主成分分析法去除图像中的冗余图像信息,然后通过融合结构张量的非局域全变分法完成图像的去噪处理^[6-8]。

1.1 图像冗余信息预处理

受灯光、设备参数设置等因素的影响,采集到的特定虚拟图像多变性特征较多,因此存在无用信息较多的问题,通过主成分降维分析法去除特定虚拟图像中的无用图像数据信息^[9]。图像降维通过互相关联聚类变量的离散投影实现。

用 $\mathbf{F}_N$ 表示主成分聚类的类间离散矩阵,计算全局离散投影矩阵 $\mathbf{D}_Y$, 其公式表达如下:

$$\mathbf{D}_Y = \mathbf{E}_1 \mathbf{F}_N + \mathbf{E}_2 \mathbf{F}_E \quad (1)$$

其中, $\mathbf{E}_1$ 、 $\mathbf{E}_2$ 表示加权因子, $\mathbf{F}_E$ 表示类内离散矩阵。

利用空间中线性判别降维方法和主成分分析降维方法对 $\mathbf{F}_E$ 实行降维转换处理,其公式表达如下:

$$\mathbf{F}_E = \sum_{o=1}^v \sum_{k=1}^{m_o} (\vec{c}_k^{(o)} - \vec{q}^{(o)}) (\vec{c}_k^{(o)} - \vec{q}^{(o)})^T \quad (2)$$

其中, v 表示聚类数, o 表示聚类维度, k 表示样本数量, m_o 表示单一聚类投影维度所对应的样本总数, $\vec{c}_k^{(o)}$ 聚类中存在的样本, $\vec{q}^{(o)}$ 表示聚类 o 的聚类中心。

通过加权处理更新 $\mathbf{F}_E$, 去除离散的冗余信息,获得图像主成分全局最佳聚类投射矩阵 $\mathbf{D}_Y'$, 其公式表达如下:

$$\mathbf{D}_Y' = (\mathbf{E}_1 \mu + \mathbf{E}_2) \mathbf{F}_E \quad (3)$$

其中, μ 表示常数。

由此通过主成分聚类降维映射获得主成分最优映射投影,完成虚拟图像的去冗余预处理。

1.2 图像去噪

非局部正则化算法具有纹理特征描述佳的优点,可以在去噪的同时保留重要纹理特征。将结构张量融入该算法进行全变分约束,实现图像预处理。

算法使用的非局部算子 $\nabla_{NL} i(c, u)$ 公式表达如下:

$$\nabla_{NL} i(c, u) = [i(u) - i(c)] \sqrt{e_{NL}(c, u)}, \quad (4)$$

其中, $i(u)$ 、 $i(c)$ 表示空间 Γ 中的图像两个像素点, $\sqrt{e_{NL}(c, u)} \in \Gamma \times \Gamma \rightarrow T$ 表示符合对称性的像素点间的边缘权重, ∇ 表示梯度算子。由公式 (4) 可得像素的非局部 (NL) 梯度 $|\nabla_{NL} i(c)|$ 公式表达如下:

$$|\nabla_{NL} i(c)| \mathbf{D}_Y' \sqrt{\int_{\Gamma} [i(u) - i(c)]^2 e_{NL}(c, u) u} \quad (5)$$

图像的结构张量 $K_o(\nabla O_{\zeta})$ 公式表达如下:

$$K_o(\nabla O_{\zeta}) = H_{\sigma} * (\nabla O_{\zeta} \otimes \nabla O_{\zeta})$$

$$|\nabla_{NL} i(c)| \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \quad (6)$$

其中, H_{σ} 表示方差为 ζ 的高斯滤波器, O_{ζ} 表示经高斯滤波后的图像, $\otimes$ 表示张量积运算, $*$ 表示卷积运算。将图像中非局域信息的相似性权重值与结构张量的相似性权重值结合,以保留图像空间域内灰度相似性及特殊纹理特性,从而避免去噪权重调节时消除纹理特征。

结合上述公式建立结构张量与非局域全变分

去噪模型 $R(i)$ 公式表达如下:

$$R(i) = \min_{i \in T^m} \int_{\Gamma} \frac{1}{2} \|i - \iota\|^2 + \mu_{1,2} |\nabla_{STNL} i| e_{NL}(c, u) K_{\sigma}(\nabla O_{\zeta}) di \quad (7)$$

其中, $\nabla_{STNL} i$ 表示非局域全变分, ι 表示含有噪声的图像。基于模型,完成特定虚拟图像的除冗及去噪预处理,为后续特征提取奠定良好数据基础。

2 多模态下特定虚拟图像视觉特征提取

完成特定虚拟图像的预处理后,通过稀疏自编码器和卷积池化技术展开深度学习,基于预处理后的图像进行多模态下特定虚拟图像的特征提取^[10-11]。

传统的卷积神经网络在对图像实行特征提取时,往往只能提取到图像的单一 RGB 图像特征值,导致最终的图像视觉特征自动补偿精度欠佳,故融合稀疏自编码器和卷积池化技术,利用 L-BFGS 算法迭代深度学习训练获取多模态下图像的彩色图、深度图、灰度图和 3D 曲面法线特征,对传统卷积神经网络实行优化处理^[12]。

利用卷积神经网络的隐藏层对图像的特征实行提取。用 c 简化代指预处理结果去噪模型 $R(i)$ 表示输入图像,使用 Sigmoid 激活函数 f_i 作为映射,得到隐藏层图像输出特征 u 公式表达如下:

$$u = f_i(e c + n) \quad (8)$$

其中, $\vartheta = \{e, n\}$ 表示网络参数, e 表示权值矩阵, n 表示偏置向量。

在隐藏层实行二次映射,得到图像输出特征 x 公式表达如下:

$$x = \tilde{f}_i(\tilde{e} u + \tilde{n}) \quad (9)$$

其中, $\tilde{f}_i$ 表示二次映射函数, $\tilde{e} = e^T$ 表示矩阵权值的转置, $\tilde{f}_i$ 表示偏置向量的转置, $\tilde{\vartheta} = \{\tilde{e}, \tilde{n}\}$ 参数可通过训练完成调整,影响需要训练参数个数。

通过迭代学习优化输入的图像,通过权值矩阵和偏置向量的调整,得到隐藏层最小总体代价函数。

但是直接使用稀疏自编码器完成图像特征训练时存在丢失细节问题,故再通过卷积和池化处理对其实行优化,以学习到的图像中局部小块图像特征作为探测器,对原图像实行卷积操作,最终获得

不同图像不同区域的不同特征激活值。使用最大值池化和均值池化方法对卷积后的图像的彩色图、深度图、灰度图和 3D 曲面法线性特征实行池化采样。

基于此,完成多模态下特定虚拟图像特征提取,将提取的特征输入训练网络中进行自动补偿。

3 特定虚拟图像视觉特征自动补偿

将上节提取的特征输入至卷积神经网络,通过非线性映射将彩色图、深度图等隐藏特征映射至图像特征结构中,重建视觉特征完成图像视觉特征自动补偿^[13-14],增强图像细节信息。

视觉特征补偿卷积神经网络结构如图 1 所示。

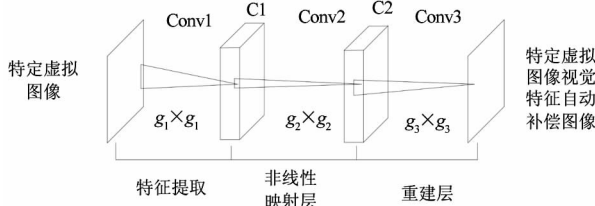


图 1 视觉特征补偿结构图

如图 1 所示,将上节提取到的图像隐藏层特征 x 作为非线性映射层的输入^[15]。

非线性映射即通过非线性方法映射一个高维度矢量至另一个高维度矢量,其公式表达如下:

$$G_2 = \max(0, E_2 \times x + N_2) \quad (10)$$

其中, E_2 表示含有 n_2 个空域大小为 $g_1 \times g_1$ 的滤波器, N_2 表示 n_2 向量。

最后一层重建层通过卷积运算对上层得到的图像实行视觉图像特征自动补偿,其公式表达如下:

$$G(x) = E_3 \times G_2(x) + N_3 \quad (11)$$

其中, E_3 表示含有 n_3 个空域大小为 $g_1 \times g_1$ 的滤波器, N_3 表示 n_3 向量。

基于此,通过卷积神经网络模型完成特定虚拟图像视觉特征自动补偿。

4 测试与分析

为了验证基于多模态深度学习的特定虚拟图像视觉特征自动补偿方法的有效性,对其完成如下测试。

本文数据采用 Middlebury 立体图像公开数据集集中的 500 张特定虚拟人脸图像作为训练集(包括 Art、Book、Dolls、Laundry、Moebius、Reindeer 标准数据集),六种标准数据集共选用 300 张图像作为测试集。通过下采样操作获得低分辨率图像,具体操作如下:将训练集中的图像分别进行缩放,即截取图像原尺寸的 60%再放大到原尺寸,将原始图像转变为细节丢失的低分辨率图像。补偿重建倍数为 4。将训练集输入卷积神经网络中进行训练,训练的软硬件环境为:Matlab 配套神经网络工具箱、Windows 10 操作系统、英特尔 i5-9400F 处理器、NVIDIA GeForce GTX1070 显卡、Python 编程语言、Tensor Flow 深度学习框架。网络训练参数如表 1 所示:

表 1 网络训练参数表

参数类型	参数
激活函数	PreLu
基本学习率	0.001
输出步幅	18
训练迭代次数	40000
神经网络单元数	185

完成训练后,从 Art 测试集中选用 1 张特定虚拟人脸图像作为测试结果展示案例,如图 2 所示:



图 2 实验对象

采用基于多模态深度学习的特定虚拟图像视觉特征自动补偿方法(所提方法)、文献[3]方法和文献[4]方法完成对比测试。

4.1 预处理效果

由图 3 对比可知,三种方法都使图像像素点变得更加清晰,但所提方法的预处理效果最好,去除

了冗余噪点,纹理特征更加清晰。

4.2 视觉特征提取

由图 4 对比可知,所提方法提取的视觉特征关键点最多,能够通过特征点构造图像人脸的五官轮廓,而其他两种方法虽然能够提取人脸脸型轮廓,但五官轮廓的关键点提取不足,证明本文方法的视觉特征提取效果较好。



(a)文献[3]方法的图像预处理效果



(b)文献[4]方法的图像预处理效果



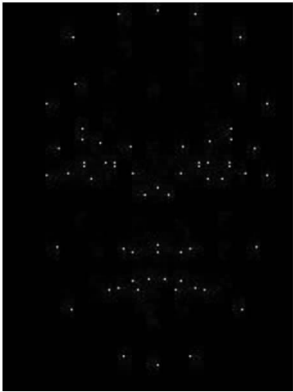
(c)所提方法的预处理效果

图 3 三种方法的图像预处理效果

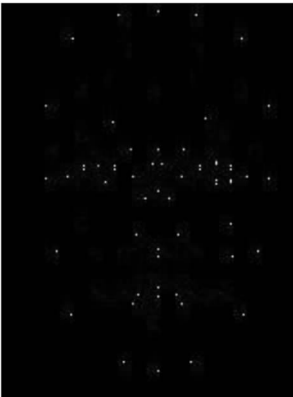
4.3 图像特征补偿效果

采用所提方法、文献[3]方法、文献[4]方法和

文献[5]方法完成特定虚拟图像特征补偿处理,其结果如图 5 所示:



(a)文献[3]方法的图像视觉特征提取效果



(b)文献[4]方法的图像视觉特征提取效果



(c)所提方法的视觉特征提取效果

图 4 视觉特征提取效果

分析图 5 可知,所提方法的图像细节更加清晰,特征补偿效果明显优于文献[3]方法和文献[4]方法,表明所提方法的特定虚拟图像的预处理效果更好。

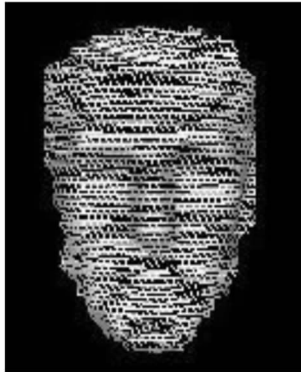
4.4 补偿质量定量分析

采用均方根误差 (RMSE) 和结构相似性 (SSIM)作为客观评价指标。RMSE 用于衡量重建

图像与原始图像之间的差异程度,重建图像与原始图像之间的差异越小,即补偿重建质量越好,较小的 RMSE 值意味着重建图像与原始图像之间的像素差异较小,重建图像更接近原始图像,图像质量更高;比较原始图像与经过压缩、降噪或其他处理后的图像之间的相似程度时,通常情况下,当图像经过压缩或其他处理后,SSIM 指数会下降,意味着图像的结构相似性较原始图像有所损失。采用所提方法、文献[3]方法和文献[4]方法完成特定虚拟图像视觉特征自动补偿,不同方法的图像补偿结果 RMSE 值和 SSIM 值对比结果如图 6 和图 7 所示:



(a)文献[3]方法的图像特征补偿效果



(b)文献[4]方法的图像特征补偿效果



(c)所提方法的图像特征补偿效果

图 5 三种方法的图像特征补偿效果

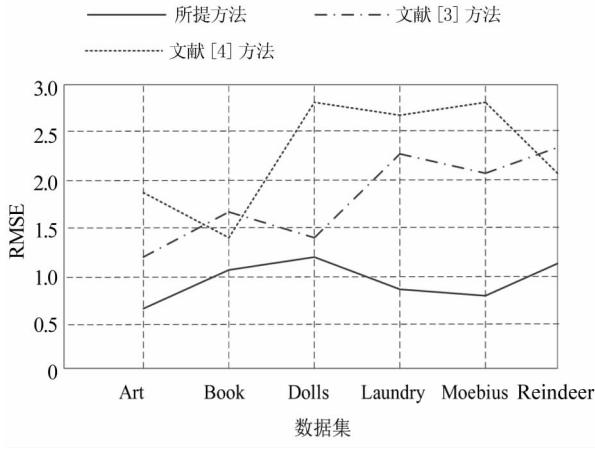


图 6 不同方法的图像补偿结果 RMSE 值

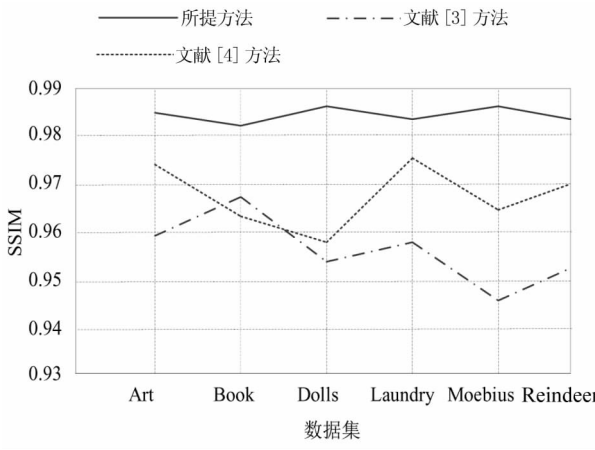


图 7 不同方法的图像补偿结果 SSIM 值

分析图 6 和图 7 可知,与文献[3]方法和文献[4]方法相比,所提方法对特定虚拟图像完成补偿后的 RMSE 值较低、SSIM 较高,表明所提方法的图像视觉特征补偿质量更好。

通过上述实验可知,所提方法提高了图像自动补偿清晰度;其次,通过多模态深度学习方法的图像特征自动补偿,提高了图像自动补偿的质量,获得了更好的特定虚拟图像视觉特征自动补偿效果。

5 结 论

随着虚拟技术的发展,虚拟图像逐步被广泛应用,特定虚拟图像的视觉特征自动补偿技术成为人们研究的重点,因此,提出基于多模态深度学习的特定虚拟图像视觉特征自动补偿。首先,对图像实行预处理;其次,完成多模态下图像的特征提取;最后,完成特定虚拟图像视觉特征的自动补偿。测试

结果表明本文方法能够有效提高图像补偿质量,具有重要现实应用意义。

参考文献

[1] 张芳,赵东旭,肖志涛,等.单幅图像超分辨率重建技术研究进展[J].自动化学报,2022,48(11):2634—2654.

[2] 杨夏宁,王帮海,李佐龙.分层特征融合网络航拍图像超分辨率重建[J].计算机工程与应用,2022,58(19):224—232.

[3] 于淑侠,胡良梅,张旭东,等.彩色图像多尺度引导的深度图像超分辨率重建[J].光电工程,2020,47(4):42—51.

[4] 王紫薇,邓慧萍,向森,等.基于 CNN 的彩色图像引导的深度图像超分辨率重建[J].图学学报,2020,41(2):262—269.

[5] 姜红涛,孙京,谢成.内外梯度联合约束的图像超分辨率重建方法[J].计算机工程,2022,48(1):220—227+235.

[6] 金羽锋,肖裔新,符采灵.应用于 VA-LCD Mura 补偿的图像预处理算法[J].液晶与显示,2021,36(7):999—1005.

[7] 李孟航,高珊珊,韩慧健,等.结合 L₀ 优化与拉普拉斯算子的图像平滑方法[J].计算机辅助设计与图形学学报,2021,33(7):1000—1014.

[8] 陆县委,刘辉,尚振宏.基于自监督深度学习的 NVST 图像去噪[J].激光与光电子学进展,2021,58(6):252—263.

[9] 韩银萍,刘丽,王冰洁,等.基于鲁棒主成分分析的混沌穿墙成像雷达杂波抑制方法[J].电子器件,2020,43(1):142—146.

[10] 何俊,张彩庆,李小珍,等.面向深度学习的多模态融合技术研究综述[J].计算机工程,2020,46(5):1—11.

[11] 沈克成,施佳,王晗.可见光—远红外图像的多模态能见度深度学习模型[J].计算机辅助设计与图形学学报,2021,33(6):939—946.

[12] 聂茜茜,肖斌,毕秀丽,等.基于超像素级卷积神经网络的多聚焦图像融合算法[J].电子与信息学报,2021,43(4):965—973.

[13] 董安勇,杜庆治,苏斌,等.基于卷积神经网络的红外与可见光图像融合[J].红外技术,2020,42(7):660—669.

[14] 胡晓辉,张建国.基于改进卷积神经网络的图像超分辨率算法研究[J].计算机应用研究,2020,37(3):947—950+956.

[15] 徐少平,林珍玉,张贵珍,等.采用深度学习与图像融合混合实现策略的低照度图像增强算法[J].电子学报,2021,49(1):72—76.