

文章编号: 2095-2163(2023)09-0184-06

中图分类号: TP391.9

文献标志码: A

# 基于特征迁移网络的大范围地面场景红外图像快速生成研究

牟卿志, 周 荃, 宋西宁, 孙春萍

(潍坊职业学院 机电工程学院, 山东 潍坊 262737)

**摘要:** 目前大范围红外地面场景面临获取难度大、成本高的问题。本文基于小范围红外实测图像与大范围卫星可见光图像之间的信息关联性,通过搭建 VGG19-SANet 特征迁移网络,将红外实测图像的特征信息迁移到大范围卫星可见光图像的内容结构上,从而实现大范围地面场景红外图像的快速生成。实验数据分析表明,通过本文方法得到的大范围地面场景红外图像构建速度快,在图像基本信息相似度、梯度相似度等各方面均接近于实测红外图像,具有较高的置信度,可用于现有的各类仿真实验。

**关键词:** 红外成像仿真; 特征迁移; VGG-SANet

## Research on fast infrared image generation of large scale ground scene based on style transfer network

MU Qingzhi, ZHOU Quan, SONG Xining, SUN Chunping

(School of Mechanical and Electrical Engineering, Weifang Vocational College, Weifang Shandong 262737, China)

**【Abstract】** At present, the large-scale infrared ground scene is faced with the problem of difficult acquisition and high cost. Based on the information correlation between the small range infrared measured images and the large range satellite visible images, this paper builds a VGG19-SANet feature migration network to migrate the feature information of the infrared measured images to the content structure of the large range satellite visible images, thus realizing the rapid generation of infrared images of large range ground scenes. The analysis of experimental data shows that the infrared image of a large range of ground scenes obtained by this method is fast in construction, and is close to the measured infrared image in terms of image basic information similarity, gradient similarity, etc. It has a high degree of confidence and can be used in various existing simulation experiments.

**【Key words】** infrared imaging simulation; style transfer; VGG-SANet

## 0 引言

红外成像制导具有目标识别能力强,全天候打击,识别精度高等特点,已成为现代精确制导武器的重要组成部分<sup>[1-2]</sup>。在对地面红外成像制导武器的研制中,需要提供大量不同条件下的大范围地面场景红外图像作为输入进行测试。然而,由于试验次数限制或非合作目标等原因,外场实测很难得到大范围的地面红外场景<sup>[2]</sup>。目前常采用采集大量不同位置实测图片手动拼接后,再进行人工修正方式,耗费大量时间与人力物力成本。若需要获得不同高度/气候条件下的大范围的地面红外场景,则该过程更加冗长复杂。

红外成像仿真技术可为解决这类问题提供一种极为有效/经济的途径,对现代精确制导武器的开发和研制来说具有重要意义<sup>[3-5]</sup>。传统红外建模仿真方式依据纯理论计算,需要大量时间和精力去完成高精度的建模与计算。当建模精度不足情况下,仿真结果往往与实测数据相差较大。本文不同于传统红外建模仿真方式,在特征迁移工作的启发下,搭建 VGG19<sup>[6]</sup>与 SANet<sup>[7]</sup>网络框架,探索采用基于卫星实拍的大范围可见光图片与小范围红外实测数据相结合的方式,结合可见光图片纹理内容信息与实拍红外数据高置信度特征样式,生成具有与实拍数据相近特征样式的大范围场景红外仿真。实验结果分析表明,该方法生成的大范围红外场景与实测数据具

**作者简介:** 牟卿志(1996-),男,硕士,助教,主要研究方向:计算机视觉、深度学习;周 荃(1981-),女,硕士,副教授,主要研究方向:图像信息处理;宋西宁(1983-),女,硕士,讲师,主要研究方向:图像质量分析与评估;孙春萍(1975-),女,硕士,副教授,主要研究方向:高等数学、数学建模。

**通讯作者:** 牟卿志 Email: 920128084@qq.com

收稿日期: 2022-10-27

有较高的一致性,且图像生成方便快捷,在各类仿真试验中具有一定的实用价值。

## 1 特征迁移神经网络简介

特征迁移<sup>[8-9]</sup>是一种用于通过给定的特征图像,均匀地在内容图像上合成全局和局部特征信息,同时保持其原有图像内容结构来创建新图像的技术。特征迁移工作示意如图 1 所示。

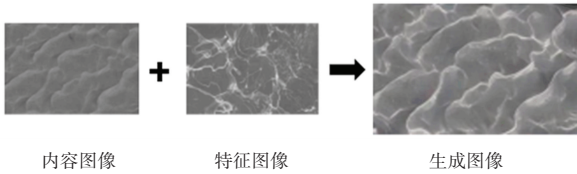


图 1 特征迁移工作示意

Fig. 1 Example of style transfer

在特征迁移网络发展初期,其一般用于艺术及生活方面的应用,如画作风格迁移、笔触模仿等<sup>[10-11]</sup>。随着模型研究、数据集组建及测试工作的深入,在合理调节参数及设置训练数据集的情况下,特征迁移网络在保持内容结构而不丢失特征样式丰富性的需求方面依旧呈现出优异的表现,从而逐步开始应用于样本扩增、数据增强等高价值任务领域,为空间探测、高空红外成像等稀少样本目标场景的数据扩增提供了新颖有效的方法<sup>[12-13]</sup>。

近期,一种新的特征关注网络(SANet)和损失函数被提出。对于指定特征样式迁移,由 SANet 和解码器组成前馈网络,学习内容特征信息的相关性。SANet 使用可学习的相似性内核,将内容特征图进行特定表示,并训练模型以恢复相同的结果。推理过程中,将输入图像之一替换为样式图像,并根据特征样式尽可能还原内容图像。本文基于文献<sup>[9]</sup>的思路,在大范围地面场景红外图像快速生成方面也取得了较好结果。

## 2 基于 VGG 与 SANet 的特征迁移网络

### 2.1 整体架构

整体网络架构由 VGG19 网络模块与 SANet 网络组成。VGG19 网络用作编码器与解码器模块,编码器模块与 SANet 网络结合构成前馈网络,来对图像间的相关性进行学习,解码器模块用于对推理后的图像进行恢复与生成,最后通过调试定义的损失函数,对生成图像进行评价训练。整体网络工作流程如图 2 所示。

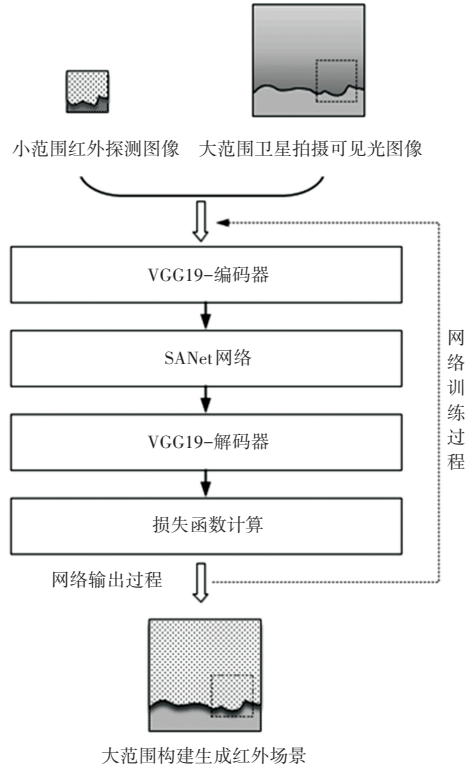


图 2 VGG-SANet 网络工作流程

Fig. 2 VGG-SANet network workflow

### 2.2 VGG 网络

VGG 系列网络由牛津大学 Visual Geometry Group 团队研发搭建,主要贡献是证明增加网络深度可以提升网络精度与训练效果。VGG 网络结构相对传统神经网络的改进是采用连续的若干个 3x3 的卷积核代替较大卷积核(11x11,7x7,5x5)。对于给定的感受野(与输出有关的输入图片的局部大小),采用堆积的小卷积核,优于采用大的卷积核。因为多层非线性层可以增加网络深度来保证学习更复杂的模式,而且代价更小。在 VGG 中,使用了 3 个 3x3 卷积核来代替 7x7 卷积核、2 个 3x3 卷积核来代替 5x5 卷积核,其主要目的是在保证具有相同感知野的条件下,提升网络的深度,在一定程度上提升神经网络的效果。

常用 VGG 架构有 16 层(VGG16)与 19 层(VGG19)两种。在 VGG16 架构中,包含有 16 个隐藏层(13 个卷积层和 3 个全连接层),在 VGG19 架构中,包含有 19 个隐藏层(16 个卷积层和 3 个全连接层),在部分层间采用池化层进行连接。VGG 架构的具体实现如图 3 所示。

VGG-16网络	VGG-19网络	信息流方向
图像信息输入		↓
conv3-64(ReLU1_1) conv3-64(ReLU1_2)	conv3-64(ReLU1_1) conv3-64(ReLU1_2)	
池化层		
conv3-128(ReLU2_1) conv3-128(ReLU2_2)	conv3-128(ReLU2_1) conv3-128(ReLU2_2)	
池化层		
conv3-256(ReLU3_1) conv3-256(ReLU3_2) conv3-256(ReLU3_3)	conv3-256(ReLU3_1) conv3-256(ReLU3_2) conv3-256(ReLU3_3) conv3-256(ReLU3_4)	
池化层		
conv3-512(ReLU4_1) conv3-512(ReLU4_2) conv3-512(ReLU4_3)	conv3-512(ReLU4_1)→SANet conv3-512(ReLU4_2) conv3-512(ReLU4_3) conv3-512(ReLU4_4)	
池化层		
conv3-512(ReLU5_1) conv3-512(ReLU5_2) conv3-512(ReLU5_3)	conv3-512(ReLU5_1)→SANet conv3-512(ReLU5_2) conv3-512(ReLU5_3) conv3-512(ReLU5_4)	
池化层		
全连接层(FC-4096)		
全连接层(FC-4096)		
全连接层(FC-1000)		
Softmax层		
图像信息输出		

图3 VGG网络架构(包含SANet)

Fig. 3 VGG network architecture (including SANet)

### 2.3 VGG 编码与解码

通过 VGG 网络进行卷积与池化等操作,其实是对输入图像不断进行信息化编码的过程。该过程将

原始图像像素映射到特征图谱集合,被多种网络架构用于学习有效的表示。通常认为,在更高卷积层数生成的特征图谱,反应的信息越高维,概括性的描述能力越强,反之则对细节的描述能力更强。在编码过程中,这一特性被用来选择特征图谱,并输入到其他网络架构(如 SANet)网络中进行信息处理。在解码过程中,网络架构的输出又通过 VGG 网络的逆过程进行还原,完成最终的编码与解码过程。

### 2.4 VGG-SANet 网络结构

在前向推理过程中,输入一幅图像后,SANet 网络选用 VGG 编码器中的 ReLU4\_1 与 ReLU5\_1 的特征图谱作为输入,记图像  $I_c$  与  $I_s$  得到的特征图谱为  $F_c^{r-4,1}$ 、 $F_c^{r-5,1}$ 、 $F_s^{r-5,1}$ , 其作为两组输入分别输入到两个 SANet 网络中,计算生成输出  $F_{cs}^{r-4,1}$  与  $F_{cs}^{r-5,1}$ 。该两组信息再次结合  $F_c^{r-4,1}$  与  $F_c^{r-5,1}$  进行计算处理,得到最终的生成特征图谱  $F_{cs}^m$ ,并通过 VGG 解码器还原得到输出的合成图像  $I_{cs}$ 。

在反馈训练过程中,输出的合成图像  $I_{cs}$  通过 VGG 编码器再次进行编码,与  $I_s$  图像及先前提取的  $\overline{F_c^{r-4,1}}$  及  $\overline{F_c^{r-5,1}}$  进行计算,得到损失函数用以指导网络的训练。VGG-SANet 网络架构如图 4 所示。

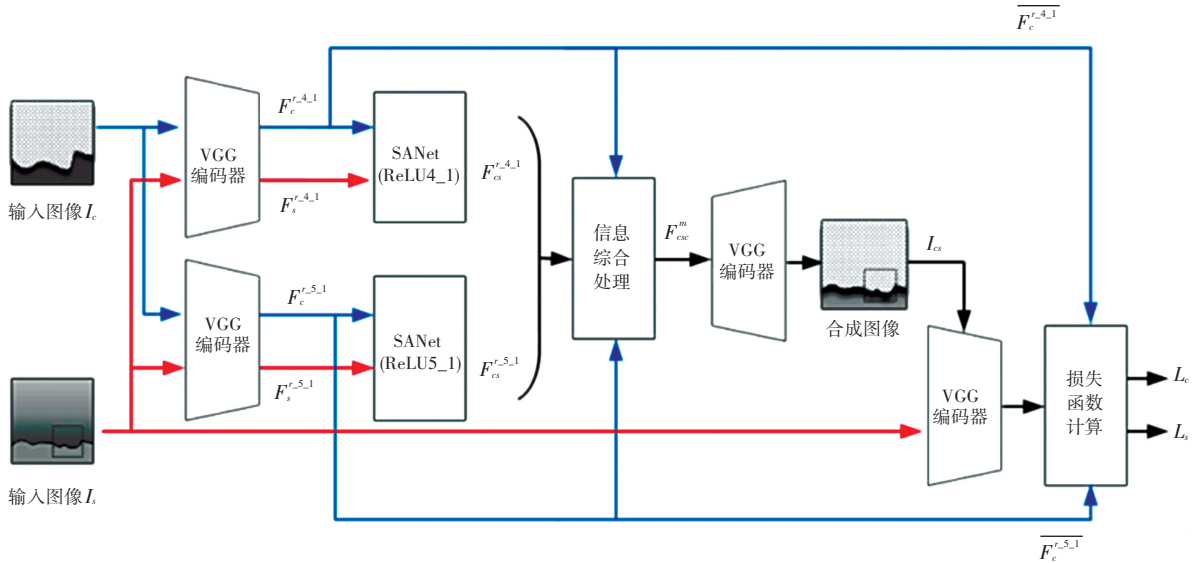


图4 VGG-SANet 网络架构

Fig. 4 VGG-SANet Network Architecture

### 2.5 SANet 网络实现

如图 5 所示,单个 SANet 网络的输入为两幅图像在 VGG 某一层的特征图谱  $F_c$  与  $F_s$ , 将处理后的  $\overline{F_c}$ 、 $\overline{F_s}$  及  $F_s$  分别进行运算后得到  $f(\overline{F_c})$ 、 $g(\overline{F_s})$  和  $h(\overline{F_s})$ 。则有  $f(\overline{F_c}) = W_f \overline{F_c}$ 、 $g(\overline{F_s}) = W_g \overline{F_s}$ 、 $h(\overline{F_s}) = W_h \overline{F_s}$ , 而  $W_f$ 、 $W_g$ 、 $W_h$  是可通过学习更新的权重矩阵。最终得到:

$$F_{cs}^i = \frac{1}{C(F)} \sum_{v_j} \exp(f(\overline{F_c}^i)^T g(\overline{F_s}^j) h(\overline{F_s}^j)) \quad (1)$$

$$C(F) = \sum_{v_j} \exp(f(\overline{F_c}^i)^T g(\overline{F_s}^j)) \quad (2)$$

其中,  $i$  为输出位置的索引,  $j$  为所有可能输出位置的枚举索引。

在通过单个 SANet 网络得到输出的  $F_{cs}$  后,与

VGG 指定层特征图谱  $F_c$  进行运算, 得到  $F_{csc}^{r-4-1}$  与  $F_{csc}^{r-5-1}$ , 再对  $F_{csc}^{r-5-1}$  进行上采样并联合卷积后, 得到预期输出的  $F_{csc}^m$ 。

$$F_{csc} = F_c + W_{cs} F_{cs} \quad (3)$$

$$F_{csc}^m = conv_{3 \times 3}(F_{csc}^{r-4-1} + upsampling(F_{csc}^{r-5-1})) \quad (4)$$

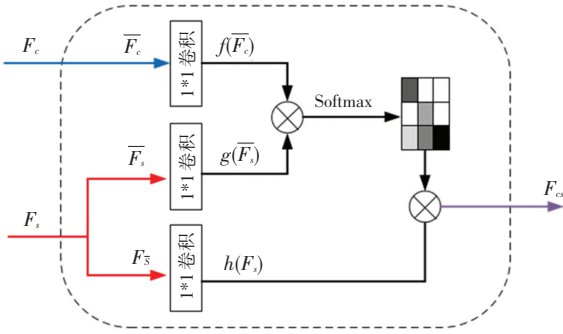


图5 SANet 网络实现

Fig. 5 SANet Network

### 2.6 损失评价函数

网络构建后, 需要通过损失评价函数, 对每次网络输出的结果进行评价, 并以此更新网络中各矩阵的权值直到收敛, 以实现网络最佳性能。反馈训练网络示意图如图6所示。

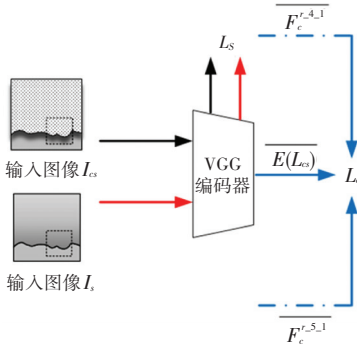


图6 反馈训练网络

Fig. 6 Feedback training network

在反馈训练网络中, 函数  $L_s$  通过输入图像  $I_s$  与  $I_{cs}$  进行 VGG 编码与特征图谱计算实现, 函数  $L_c$  通过计算 VGG 编码得到的  $E(I_{cs})$  及推理网络中的  $F_c^{r-4-1}$ 、 $F_c^{r-5-1}$  来实现。最后通过  $L_s$  与  $L_c$ , 设定附加项  $L_{id}$ , 共同得到总的损失评价函数  $L$ 。其计算公式为:

$$L_c = \| \overline{E(I_{cs})}^{r-4-1} - \overline{F_c}^{r-4-1} \|_2 + \| \overline{E(I_{cs})}^{r-5-1} - \overline{F_c}^{r-5-1} \|_2 \quad (5)$$

$$L_s = \sum_{i=1}^L \| \mu(\phi_i(I_{cs})) - \mu(\phi_i(I_s)) \|_2 + \| \sigma(\phi_i(I_{cs})) - \sigma(\phi_i(I_s)) \|_2 \quad (6)$$

$$L_{id} = \lambda_{id1} (\| I_{cc} - I_c \|_2 + \| I_{ss} - I_s \|_2) +$$

$$\lambda_{id2} \sum_{i=1}^L (\| \phi_i(I_{cc}) - \phi_i(I_c) \|_2 + \| \phi_i(I_{ss}) - \phi_i(I_s) \|_2) \quad (7)$$

最后可得:

$$L = \lambda_c L_c + \lambda_s L_s + L_{id} \quad (8)$$

其中,  $I_{cc}$  (或  $I_{ss}$ ) 表示从两个相同内容 (或样式) 图像合成的输出图像, 每个  $\phi_i$  表示编码器中的各层编码函数;  $\lambda_{id1}$  和  $\lambda_{id2}$  是损失评价函数  $L_{id}$  的调节权重;  $\lambda_c$  与  $\lambda_s$  为总损失评价函数  $L$  的调节权重。

在实际调整训练参数的过程中, 通过对数据集的测试发现, 保持  $L_c$ 、 $L_s$ 、 $L_{id}$  三者的数量级在同一水平, 且令  $L_c$ 、 $L_s$  起主导作用, 训练收敛精度较高, 测试集会取得较好的结果表现。在本文实验中, 权重参数设置为  $\lambda_c = 1.5$ ,  $\lambda_s = 3$ ,  $\lambda_{id1} = 1$ ,  $\lambda_{id2} = 40$ 。

通过式(5)~式(8)并引入数据集训练, 最终实现网络的生成。

## 3 实验结果与分析

### 3.1 实验结果分析思路

在本文实验中, 为保证可见光-红外图集的特征能够准确对应, 采用已有数据集 OCTBVS<sup>[14]</sup> 作为可见光-红外配对图集, 提供训练内容。在测试实验中, 选取实拍红外图像区域部分及可见光整体部分作为网络输入, 从而生成特征迁移后大范围红外仿真图像。此外, 还将大范围可见光图像进行灰度转化调节作为对比组 (此时手动调整至灰度与真实红外图像一致, 对其他指标进行观察), 进而验证方法的有效性。

### 3.2 实验结果

实验结果如图7~图10所示:

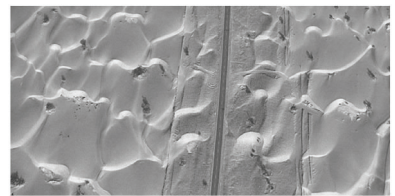


图7 实拍可见光图像

Fig. 7 Real visual image

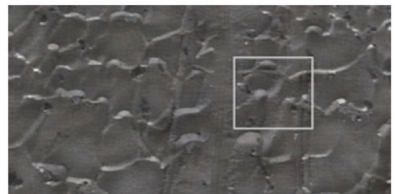


图8 实拍红外图像

Fig. 8 Real infrared image

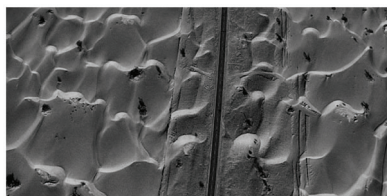


图9 可见光转灰度模拟红外图像

Fig. 9 Simulated image generated by visible light to gray scale

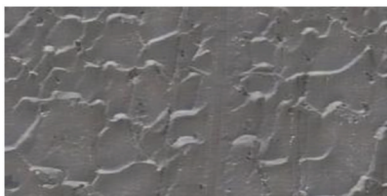


图10 特征迁移生成模拟红外图像

Fig. 10 Simulated image generated by style transfer

由此可见,实拍可见光图像与实拍红外图像(图7、图8)整体结构虽大体一致,但在纹理、边缘亮暗分布等细节方面明显表达有不同特征;可见光转灰度图像(图9)除灰度与实拍红外图像(图8)人工调节一致外,其在结构、亮暗边缘等方面上依然存在较大差异;而特征迁移生成图像(图10)通过结合真实小范围红外图像所提供的信息,在细节层面与实拍红外图像高度一致,在整体灰度层面也有较好的表现。

### 3.3 指标分析

在图像指标对比分析中,选取平均灰度 $\bar{F}$ 、平均梯度 $\bar{G}$ 、信息熵 $H(U)$ 、梯度相似度GSIM作为指标,用作图像相似度的对比指标参考。其中, $\bar{F}$ 、 $\bar{G}$ 、 $H(U)$ 三者作为单幅图像的通用数据指标,已有非常成熟的定义:

(1)平均灰度:

$$\bar{F} = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i, j)}{MN} \quad (9)$$

其中,图像中单个像素点灰度记为 $f(i, j)$ 。

(2)平均梯度:

$$\bar{G} = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} G(i, j)}{MN} \quad (10)$$

其中,梯度采用Sobel算子计算,记为 $G(i, j)$ 。

(3)信息熵:

$$H(U) = E[-\log p_i] = -\sum_{i=1}^n p_i \log p_i \quad (11)$$

其中,信息熵记为 $H(U)$ , $n$ 阶灰度中各阶取值的概率分别为 $p_1, p_2, \dots, p_i, \dots, p_n$ 。

对于两幅图像对比用的梯度相似度GSIM指标,则定义采用像素4邻域绝对值之和,表征图像梯度幅度值。设原始图像为 $x$ ,像素坐标为 $(i, j)$ ,则在坐标 $(i, j)$ 上的梯度表达式记为 $G_x(i, j)$ ,则有:

$$G_x(i, j) = |x(i, j) - x(i-1, j)| + |x(i, j) - x(i, j-1)| + |x(i, j) - x(i+1, j)| + |x(i, j) - x(i, j+1)| \quad (12)$$

$$GSIM = \frac{2 \sum_{i=1}^M \sum_{j=1}^N G_x(i, j) G_y(i, j)}{\sum_{i=1}^M \sum_{j=1}^N [G_x(i, j)]^2 + \sum_{i=1}^M \sum_{j=1}^N [G_y(i, j)]^2} \quad (13)$$

式中: $G_x(i, j)$ 、 $G_y(i, j)$ 分别为图像 $x$ 和图像 $y$ 在 $(i, j)$ 处的梯度幅度值,梯度相似度的值越大,说明两幅图像的相似程度越高。

### 3.4 指标分析结果

表1为对红外真值图像(图8)、可见光转灰度图像(图9)、特征迁移生成图像(图10)按3.3节所规定指标进行的量化分析。通过对比,验证本方法对于地面场景红外图像进行仿真生成的准确性与可行性。分析表1指标数据可知,除因对比需要进行人工调节灰度指标外,特征迁移生成图像的各项图像数据指标与实拍红外图像均有较高的一致性(差异<15%),证明存在小范围真实红外图像前提下,本文方法可在大范围可见光地面场景中,规律性的生成相关纹理、细节,快速构建与实际效果相近的大范围红外仿真地面场景,从而在相关仿真实验尤其是半实物仿真实验过程中,可有效提高效率、节约成本。

表1 指标分析结果(归一化)

Tab. 1 Index analysis results (By normalized)

	$\bar{F}$	$\bar{G}$	$H(U)$	GSIM
红外真值图像	1	1	1	1
可见光转灰度图像	0.991	1.434	1.688	1.422
特征迁移生成图像	0.952	0.970	1.021	0.937

## 4 结束语

快速准确的构建生成大范围地面场景的红外图像是目前红外成像领域的研究热点。本文通过构建基于VGG-SANet的深度神经网络,在小范围红外实测图像的特征信息与大范围卫星可见光图像的内

(下转第193页)