



(12) 发明专利

(10) 授权公告号 CN 112785502 B

(45) 授权公告日 2024. 04. 16

(21) 申请号 202110094795.7

(22) 申请日 2021.01.25

(65) 同一申请的已公布的文献号

申请公布号 CN 112785502 A

(43) 申请公布日 2021.05.11

(73) 专利权人 江南大学

地址 214122 江苏省无锡市滨湖区蠡湖大道1800号

(72) 发明人 晏涛 井花花 李明悦

(74) 专利代理机构 无锡承果知识产权代理有限公司 32373

专利代理师 张亮

(51) Int. Cl.

G06T 3/4053 (2024.01)

G06T 3/4046 (2024.01)

G06T 7/90 (2017.01)

G06V 10/74 (2022.01)

G06V 10/774 (2022.01)

G06V 10/82 (2022.01)

G06N 3/0455 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/0475 (2023.01)

G06N 3/048 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

CN 105023275 A, 2015.11.04

CN 109447919 A, 2019.03.08

CN 111062872 A, 2020.04.24

审查员 祝子豪

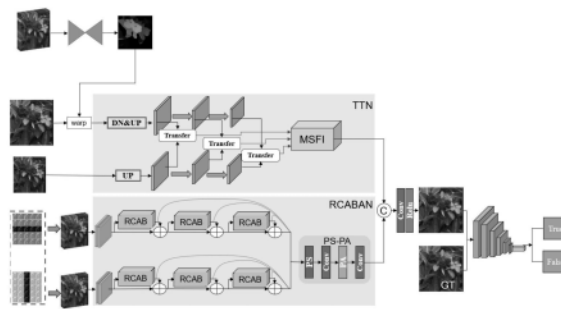
权利要求书1页 说明书5页 附图5页

(54) 发明名称

一种基于纹理迁移的混合相机的光场图像超分辨率方法

(57) 摘要

本发明公开一种基于纹理迁移的混合相机的光场图像超分辨率方法,步骤包括光场图像颜色校正、视差图计算和超分辨率网络构建三个部分。其中,光场图像颜色校正利用WCT风格转换模型使得光场图像接近单反图像的颜色;视差图计算是利用EPI全卷积神经网络模型计算视差图;超分辨率网络构建包含三个模块:第一个模块将光场的一行和一列子视点,经过RCABA提取相应子视点的特征;第二模块是PS-PA将RCABA网络得到的低频信息融合后重建为高分辨率图像;第三模块使用纹理迁移网络提取将单反图高分辨率的纹理信息并迁移到光场中心视点。本发明既利用了光场其他视点的补充信息又利用了单反图的高频信息,可以得到更高质量的超分辨率图。



1. 一种基于纹理迁移的混合相机的光场图像超分辨率方法,其特征在于,步骤包括光场图像颜色矫正、视差图计算和超分辨率网络构建三个部分;

其中,所述光场图像颜色矫正是利用WCT风格转换模型,首先训练多个解码器,将图像输入预训练好的VGG网络,提取不同的层结果作为编码输出,保持编码器的权重不变,设计解码网络,利用损失函数设计解码器;然后通过WCT层的白化操作和上色操作得到编码-WCT-解码的网络结构,从而实现光场图像颜色的矫正;

所述视差图计算是利用EPI全卷积神经网络模型来计算光场中心视点的视差图;

所述超分辨率网络构建包括三个模块:第一个模块是将光场的一行和一系列子视点,经过3D的残差通道注意力信息融合机制RCABA提取相应子视点的特征;第二个模块是重建模块PS-PA将RCABA的网络得到的低频信息融合后重建为高分辨率图像,其包括PixelShuffle层、卷积层和像素注意力模块,PixelShuffle层是将低分辨的特征图,通过卷积和多通道间的重组得到倍高分辨率的特征图;第三个模块是使用纹理迁移网络,输入是单反图和光场中心视点,输出是从单反图中迁移的纹理特征,提取单反图高分辨率的纹理信息并将纹理信息迁移到光场中心视点上。

2. 根据权利要求1所述的一种基于纹理迁移的混合相机的光场图像超分辨率方法,其特征在于,所述白化操作为根据原图的协方差矩阵的对角矩阵和正交矩阵计算原图的内容提取结果;所述上色操作为白化操作的逆操作,求逆时由计算风格图经过Encoder后的协方差矩阵的对角矩阵和正交矩阵参与计算进行还原。

3. 根据权利要求1所述的一种基于纹理迁移的混合相机的光场图像超分辨率方法,其特征在于,所述损失函数包括重建损失、感知损失、判别器损失,计算公式分别如下式(1-1)、(1-2)、(1-3)所示:

$$L_{rec} = -SSIM(I^{HR} - I^{SR}) \quad (1-1)$$

$$L_{per} = MSE(\varphi(I^{HR}) - \varphi(I^{SR})) \quad (1-2)$$

$$L_D = -\log(D(I^{HR})) - \log(1 - (G(I^{LR}))) \quad (1-3)$$

式(1-1)中, L_{rec} 是重建损失, I^{HR} 是真值, I^{SR} 是超分辨率后的结果,SSIM是结构相似度;式(1-2)中, L_{per} 是感知损失, φ 表示从VGG中提取特征图,MSE是均方误差;式(1-3)中, L_D 是判别器损失, I^{LR} 为低分辨率的输入,D是判别器,G是生成器。

4. 根据权利要求1所述的一种基于纹理迁移的混合相机的光场图像超分辨率方法,其特征在于,所述残差通道注意力信息融合机制RCABA通过残差通道注意力模块RCAB实现对不同通道的特征重新进行加权分配。

5. 根据权利要求4所述的一种基于纹理迁移的混合相机的光场图像超分辨率方法,其特征在于,所述残差通道注意力模块RCAB是将通道注意力模块CA和残差结合在一起。

一种基于纹理迁移的混合相机的光场图像超分辨率方法

技术领域

[0001] 本发明涉及一种基于纹理迁移的混合相机的光场图像超分辨率方法,属于计算机图像处理技术领域。

背景技术

[0002] 超分辨率一般包含两种:其一是参考当前低分辨率图像,不依赖其他相关图像的超分辨率技术,称之为单幅图像的超分辨率。其二是参考多幅图像或多个视频帧的超分辨率技术,称之为多帧视频/多图超分辨率(multi-frame super resolution)。单张图像获取的信息太少,视频需要借助光流等不精确的方法来对齐帧造成超分辨率的误差,光场图像一般仅借助光场内部视点有局限性。

[0003] 因此,本发明提供了一种基于纹理迁移的混合相机的光场图像超分辨率方法,以优化上述问题。

发明内容

[0004] 本发明的目的在于提供一种基于纹理迁移的混合相机的光场图像超分辨率方法,通过利用光场图像部分子视点的信息和单反图的高频信息,设计了基于纹理迁移的混合相机的光场图像超分辨率网络结构,从而实现了既利用了光场其他视点的补充信息又利用了单反图的高频信息,可以得到更高质量的超分辨率图。

[0005] 本发明的单反图是指利用单反相机拍摄的图像,像素范围为0~255。

[0006] 为实现上述发明目的,本发明采用如下技术方案:

[0007] 一种基于纹理迁移的混合相机的光场图像超分辨率方法,步骤包括光场图像颜色矫正、视差图计算和超分辨率网络构建三个部分;

[0008] 其中,所述光场图像颜色矫正是利用WCT风格转换模型使得光场图像接近单反图像的颜色;

[0009] 所述视差图计算是利用EPI全卷积神经网络模型来计算光场中心视点的视差图;

[0010] 所述超分辨率网络构建包括三个模块:第一个模块将光场的一行和一系列子视点,经过3D的残差通道注意力信息融合机制RCABA提取相应子视点的特征;第二个模块是重建模块(PS-PA)将RCABA网络得到的低频信息融合后重建为高分辨率图像(即垂直分辨率大于等于720的图像),主要由像素重组(PixelShuffle)层、卷积层和像素注意力模块组成;第三个模块是使用纹理迁移网络,提取单反图高分辨率的纹理信息并将纹理信息迁移到光场中心视点上。

[0011] 进一步地,所述光场图像颜色矫正的方法为:首先训练多个解码器,将图像输入预训练好的VGG网络,提取不同的层结果作为编码输出,保持编码器的权重不变,设计解码网络,利用损失函数设计解码器;然后通过WCT层的白化和上色两部操作得到编码-WCT-解码的网络结构,从而实现光场图像颜色的矫正。

[0012] 更进一步地,所述白化操作为根据原图的协方差矩阵的对角矩阵和正交矩阵计算

原图的内容提取结果;所述上色操作为白化操作的逆操作,求逆时由计算风格图经过Encoder后的协方差矩阵的对角矩阵和正交矩阵参与计算进行还原。

[0013] 更进一步地,所述损失函数包括重建损失、感知损失、判别器损失,计算公式分别如下式(1-1)、(1-2)、(1-3)所示:

$$[0014] \quad L_{\text{rec}} = -\text{SSIM}(I^{\text{HR}} - I^{\text{SR}}) \quad (1-1)$$

$$[0015] \quad L_{\text{per}} = \text{MSE}(\varphi(I^{\text{HR}}) - \varphi(I^{\text{SR}})) \quad (1-2)$$

$$[0016] \quad L_{\text{D}} = -\log(D(I^{\text{HR}})) - \log(1 - (G(I^{\text{LR}}))) \quad (1-3)$$

[0017] 式(1-1)中, L_{rec} 是重建损失, I^{HR} 是真值, I^{SR} 是超分辨率后的结果,SSIM是结构相似度;式(1-2)中, L_{per} 是感知损失, φ 表示从VGG中提取特征图,MSE是均方误差;式(1-3)中, L_{D} 是判别器损失, I^{LR} 为低分辨率的输入,D是判别器,G是生成器。

[0018] 进一步地,所述残差通道注意力信息融合机制RCABA通过残差通道注意力模块RCAB实现对不同通道的特征重新进行加权分配。

[0019] 更进一步地,所述残差通道注意力模块RCAB是将通道注意力模块CA和残差结合在一起。

[0020] 本发明相较于现有技术优势和有益效果在于:

[0021] 本发明利用了光场图像部分子视点的信息和单反图的高频信息,设计了基于纹理迁移的混合相机的光场图像超分辨率网络结构,相比于同类型工作,本设计既利用了光场其他视点的补充信息又利用了单反图的高频信息,可以得到更高质量的超分辨率图。

附图说明

[0022] 图1为本发明的一种基于纹理迁移的混合相机的光场图像超分辨率方法的网络结构总图;

[0023] 图2为本发明的WCT风格转换模型的网络结构图;

[0024] 图3为本发明的残差通道注意力模块RCAB的主要结构图;

[0025] 图4为本发明的通道注意力模块CA的结构图;

[0026] 图5为本发明的重建模块PS-PA的网络结构图;

[0027] 图6为本发明的像素注意力模块的结构图;

[0028] 图7为本发明的纹理迁移网络的算法流程图;

[0029] 图8为本发明的位于 $u=2$ 的整行视点以及 $v=2$ 的整列视点图像集;

[0030] 图9为本发明的合成场景及其插值的结果图,其中,Result列为合成场景的结果图,Bicubic列为合成场景插值的结果图;

[0031] 图10为本发明的真实场景及其插值的结果图,其中,Result列为真实数场景的结果图,Bicubic列为真实场景插值的结果图。

具体实施方式

[0032] 以下结合附图和优选实施例对本发明作进一步说明,本实施例在以本发明技术方案为前提下进行实施,给出了详细的实施方式和过程,但本发明的保护范围不限于下述的实施例。

[0033] 如图1所示为本发明的一种基于纹理迁移的混合相机的光场图像超分辨率方法的

网络结构总图。本发明的光场图像超分辨率方法,包括光场图像颜色矫正、视差图计算和超分辨率网络结构三个部分。具体说明如下:

[0034] 1、光场图像颜色矫正

[0035] 由于光场相机获取的光场图像的颜色偏暗,颜色与现实场景的视觉颜色差别较大,单反图的颜色比较接近真实。因此,本发明借助已有的风格转换模型将光场图像的颜色转换为单反图的颜色。

[0036] 而传统的风格转换模型通过计算原图和风格图的内容损失(content loss)和风格损失(style loss)来保障迁移的效果,这导致每种风格转换都需要专门训练对应的网络,训练是十分耗时的。为了有效优化传统风格转换模型的缺点,本发明采用WCT(Whitening and coloring transforms)风格转换模型进行转换。如图2所示为本发明的WCT风格转换模型的网络结构图。

[0037] 本发明的光场图像颜色矫正的方法为:首先训练多个解码器,将图像输入预训练好的VGG网络,提取不同的层结果作为编码输出,保持编码器(Encoder)的权重不变,设计解码网络,利用损失函数设计解码器,这样得到的编解码器实际上是对图片进行一个恒等变换。WCT层分为白化和上色两步操作,白化操作会根据原图的协方差矩阵的对角矩阵和正交矩阵,计算原图的内容提取结果;上色操作是白化操作的逆操作,只不过求逆时是计算风格图经过Encoder后的协方差矩阵的对角矩阵和正交矩阵参与计算进行还原,这样相当于在恢复原图时把风格图的风格融合到其中。经过上面的两点可以得到编码-WCT-解码的网络结构,从而实现光场图像颜色的矫正。

[0038] 进一步地,所述损失函数包括重建损失 L_{rec} 、感知损失 L_{per} 、判别器损失 L_D 三部分。其中,重建损失 L_{rec} 的计算公式如下式(1-1)所示:

$$[0039] \quad L_{rec} = -SSIM(I^{HR} - I^{SR}) \quad (1-1)$$

[0040] 式(1-1)中, I^{HR} 是真值, I^{SR} 是超分辨率后的结果,SSIM是结构相似度。感知损失 L_{per} 改善SR结果的视觉质量,计算公式如下式(1-2)所示:

$$[0041] \quad L_{per} = MSE(\phi(I^{HR}) - \phi(I^{SR})) \quad (1-2)$$

[0042] 式(1-2)中, ϕ 表示从VGG中提取特征图,MSE是均方误差。

[0043] 判别器损失 L_D 的计算公式如下式(1-3)所示:

$$[0044] \quad L_D = -\log(D(I^{HR})) - \log(1 - (G(I^{LR}))) \quad (1-3)$$

[0045] 式(1-3)中, I^{LR} 为低分辨率的输入,D是判别器,G是生成器。

[0046] 2、视差图计算

[0047] 利用一个现有的网络模型EPINet来计算光场中心视点的视差图。主要采用了一个多流(multi-stream)的网络结构,从不同的方向来获取EPI的信息。网络的输入是四个不同的角度: 0° 、 45° 、 90° 、 135° ,输出是中心视点的视差图。由于两个相机的光轴是平行的,因此通过视差图将单反图投影(warp)到光场中心视点的位置,实现对齐。根据以下公式(2-1)可以获得单反图的视差图,进而将单反图warp到光场的中心视点。

$$[0048] \quad D = \frac{Bd}{b} \quad (2-1)$$

[0049] 式(2-1)中,D是单反图的视差图,B是两个相机之间的距离,b是光场的基线,d是光场中心视点的视差图。

[0050] 3、超分辨率网络构建

[0051] 本发明的超分辨率网络构建主要分为三个模块：第一个模块主要是利用光场中心视点包含的一行和一列视点作为输入给中心视点提供丰富的低频信息和缺失的细节信息。由于低分辨率的输入包含丰富的低频信息，但却在通道间被平等对待，从而阻碍了网络的表达能力，为了优化这种问题，本发明利用残差通道注意力信息融合机制(RCABA)来提高网络的表达能力，然后通过跳跃连接将残差信息聚合起来直接进行传播，避免信息受损，使主网络专注于学习高频信息。残差通道注意力模块(RCAB)的主要网络结构如图3所示。

[0052] RCAB就是将通道注意力模块CA和残差结合在一起。如图4所示为通道注意力模块CA的结构图，通道注意力模块CA将一个 $H \times W \times C$ 的特征，先进行一个空间的全局平均池化得到一个 $1 \times 1 \times C$ 的通道描述。接着，再经过一个下采样层和一个上采样层得到每一个通道的权重系数，将权重系数和原来的特征相乘即可得到缩放后的新特征，整个过程实际上就是对不同通道的特征重新进行了加权分配。其中，激活函数为Sigmoid。

[0053] 第二个模块是重建模块PS-PA(如图5所示)，是将RCABA网络融合的低频信息重建为高分辨率图像。主要由像素重组(PixelShuffle)层、卷积层以及像素注意力模块组成，这样有利于卷积层的表达能力，更好的重建图像。PixelShuffle的主要功能是将低分辨率的特征图，通过卷积和多通道间的重组得到倍高分辨率的特征图。具体来说，就是将原来一个低分辨率的像素划分为个更小的格子，利用个特征图对应位置的值按照一定的规则来填充这些小格子。按照同样的规则将每个低分辨率像素划分出的小格子填满就完成了重组过程。

[0054] 像素注意力模块能够生成 $3D(C \times H \times W)$ 矩阵作为注意力特征。换句话说，像素注意力模块会为特征图的所有像素生成权重系数，它可以提高卷积的表达能力，其结构如图6所示。

[0055] 第三个模块是纹理迁移网络，输入是单反图和光场中心视点，输出是从单反图中迁移的纹理特征。因为参考图不能很好的和光场中心视点对齐，所以利用纹理迁移可以减少误差，防止不必要的错误信息误导网络。

[0056] 如图7所示为纹理迁移网络的算法流程图。具体流程为：首先将光场中心视点插值放大，为了使参考图和光场中心视点保持一个频域，同时将参考图降采样再上采样，然后用VGG提取光场中心视点、处理后的参考图、未经处理的参考图的特征在特征空间上匹配，计算他们之间的相似度，把相似度最高的特征迁移到光场中心视点(注意：匹配使利用处理后的参考图，迁移特征是利用未经处理的参考图)。利用内积来计算特征之间的相似度，公式如下式(3-1)所示：

$$[0057] \quad s_{i,j} = \left\langle P_i(\phi(I^{LR\uparrow})), \frac{P_j(\phi(I^{Ref\uparrow\downarrow}))}{\|P_j(\phi(I^{Ref\uparrow\downarrow}))\|} \right\rangle \quad (3-1)$$

[0058] 式(3-1)， $P_i(\cdot)$ 表示从神经特征图的第 i 块采样， $s_{i,j}$ 是第 i 个LR块和第 j 个参考图块的相似度， $I^{LR\uparrow}$ 是上采样后的光场中心视点图像， $I^{Ref\uparrow\downarrow}$ 是下采样后又上采样的单反图， $\|\cdot\|$ 是归一化。

[0059] 参考图块的特征被归一化，然后将相似度最高的特征提取出来，提取公式如下式(3-2)：

$$[0060] \quad P_{\omega(x,y)}(M) = P_{j^*}(\phi(I^{Ref})), j^* = \arg \max_j S_j(x, y) \quad (3-2)$$

[0061] 式(3-2)中, M 是交换的特征图, x 、 y 是图像的位置坐标, j^* 是相似度最高的位置下标, I^{Ref} 是单反图, $S_j(x, y)$ 在以位置 (x, y) 为中心的光场中心视点块与第 j 个单反图块之间的相似度。

[0062] 然后将相似度高的特征图迁移到光场中心视点上, 再进入多尺度残差模块中进行融合。最后将两个网络的输出经过几层卷积层融合在一起得到最终的超分辨率结果。

[0063] 上述的超分辨率网络是作为生成器的结构, 判别器共有六层, 前五层的每一层都由卷积层, BN层以及MaxPooling层组成, 卷积核都是 3×3 , 激励函数设置为relu; 最后一层使用一个全连接层将网络的输出与真值对比, 评估生成器网络的优劣, 从而对网络参数进行相应调整, 最后一层的激励函数为sigmoid。

[0064] 网络训练

[0065] 为了充分说明本发明的光场图像超分辨率方法, 本发明通过下述网络训练进行展示说明:

[0066] 首先使用blender构建虚拟光场相机和单反相机渲染出虚拟场景数据集, 将数据分割成 256×256 大小得到最终虚拟场景光场数据集和 1024×1024 的单反数据集, 再使用光场相机和单反相机采集真实场景光场数据集。

[0067] 考虑到计算资源的合理运用, 网络在训练过程中仅使用光场处于中间行和列的3D EPI。训练时, 首先将光场数据分解为 9×9 的子视点图像阵列, 图像分辨率为 $H \times W$ 提取位于 $u = 4$ 的整行视点以及 $v = 4$ 的整列视点图像集, 如图8所示 (以 $u = 2, v = 2$ 为例), 组成两组大小为 $9 \times H \times W$ 的3D EPI, 作为残差通道注意力融合网络的输入。纹理迁移网络的输入是光场的中心视点以及对应位置的反单图。网络生成器的损失函数为上述公式(1-1)和公式(1-2), 判别器损失函数为上述公式(1-3)。

[0068] 网络使用60组场景进行训练, 以训练800轮的结果为准, 在i7-8700 CPU, GTX 1070 GPU的设备上训练时间为10小时。

[0069] 图9和图10为本发明使用的数据集上测试的结果, 其中图9的Result列为合成场景的结果图, Bicubic列为合成场景插值的结果图; 图10的Result列为真实数场景的结果图, Bicubic列为真实场景插值的结果图。通过在合成数据集和真实场景测试的结果与插值后的结果对比, 可以看出本发明的算法恢复的纹理细节比较好并且清晰。

[0070] 以上所述仅为本发明的较佳实施例, 但本发明的保护范围并不局限于此。凡依本发明申请专利范围未违背本发明涉及原则所做的均等变化、简化与修饰, 皆应属本发明的涵盖范围。

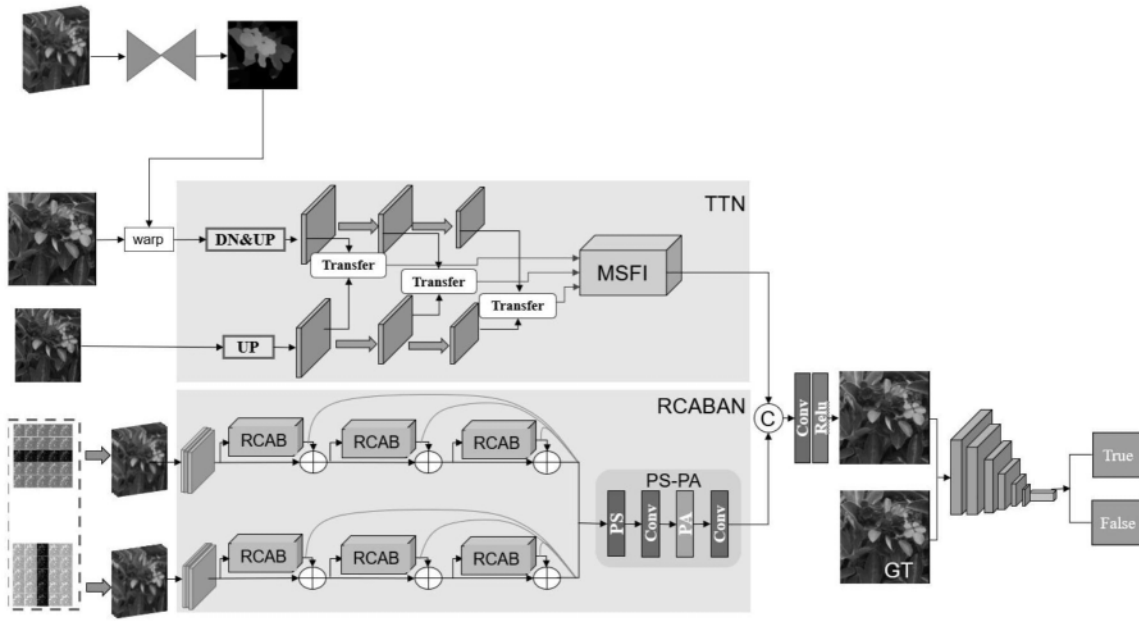


图1

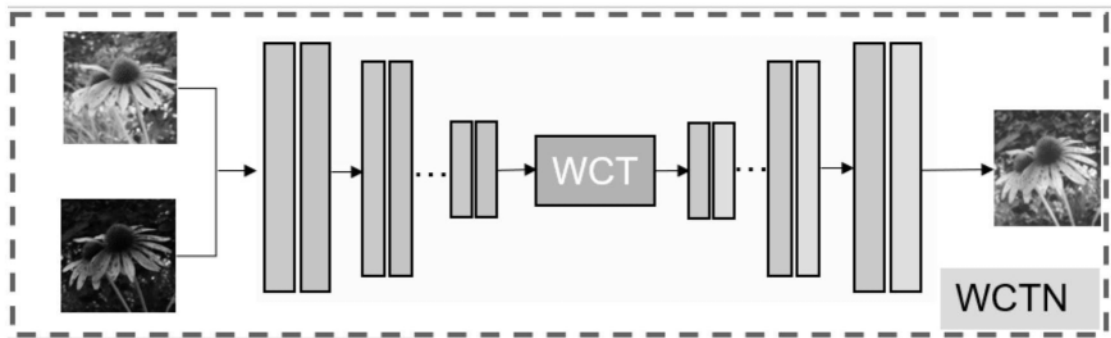


图2

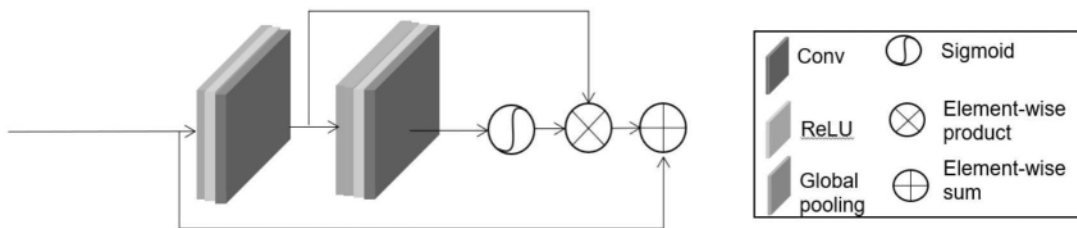


图3

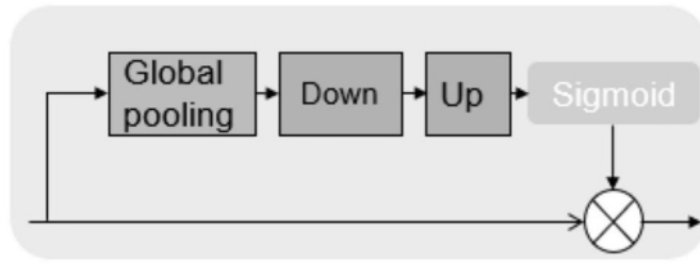


图4

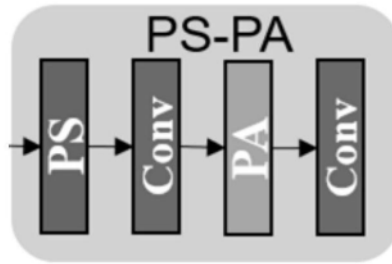


图5

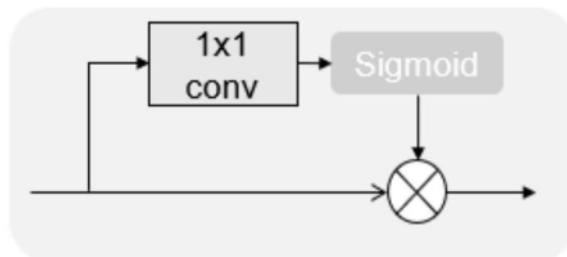


图6

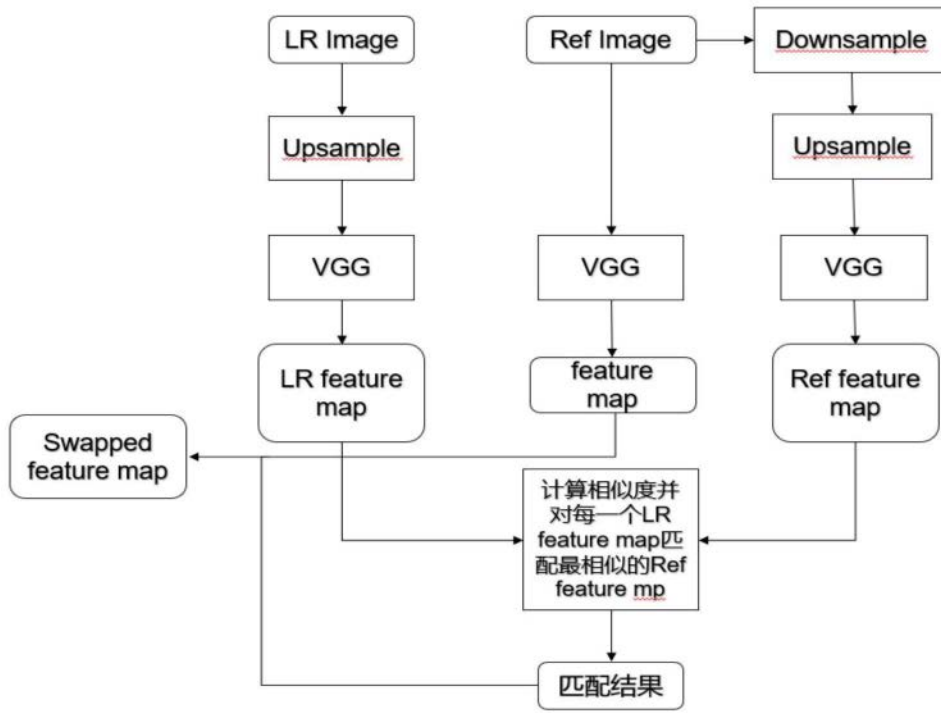


图7

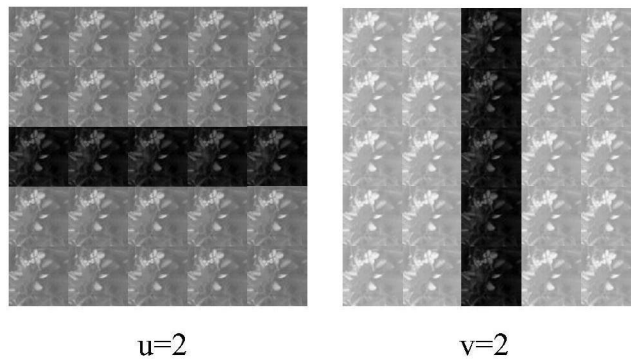


图8



图9



图10