

(19)中华人民共和国国家知识产权局



(12)发明专利申请

(10)申请公布号 CN 111179167 A

(43)申请公布日 2020.05.19

(21)申请号 201911271576.0

(22)申请日 2019.12.12

(71)申请人 天津大学

地址 300072 天津市南开区卫津路92号

(72)发明人 陈瑞 张衡 赵毅强 刘继鑫

(74)专利代理机构 天津市三利专利商标代理有限公司 12107

代理人 韩新城

(51)Int.Cl.

G06T 3/40(2006.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

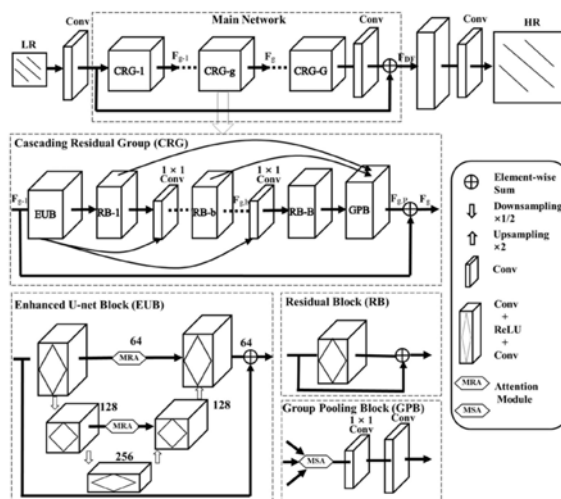
权利要求书2页 说明书6页 附图1页

(54)发明名称

一种基于多阶段注意力增强网络的图像超分辨率方法

(57)摘要

本发明公开一种基于多阶段注意力增强网络的图像超分辨率方法,通过多阶段注意力增强网络实现,所述多阶段注意力增强网络包括:浅层特征提取模块,用于从LR输入中提取浅层特征 F_0 ,基于增强U-net模块的级联残差组件,包括多级级联残差组件模块,用于将所述浅层特征 F_0 输入到多级CRG模块中,以提取深层特征 F_{DF} ;上采样模块,用于对提取到的深层特征 F_{DF} 进行上采样操作,生成高分辨率特征 F_{UP} ,图像重构组件,用于对生成的高分辨率特征 F_{UP} 通过一个卷积层映射成SR图像输出。本发明不仅允许大量的低频信息通过多个层间跳接进行流动,而且还使主干网络能够专注于学习大感受野范围内的抽象特征表示。



CN 111179167 A

1. 基于多阶段注意力增强网络的图像超分辨率方法, 其特征在于, 通过多阶段注意力增强网络实现, 所述多阶段注意力增强网络包括:

浅层特征提取模块, 用于从LR输入中提取浅层特征 F_0 ,

基于增强U-net模块的级联残差组件, 包括多级级联残差组件模块, 用于将所述浅层特征 F_0 输入到多级CRG模块中, 以提取深层特征 F_{DF} ;

上采样模块, 用于对提取到的深层特征 F_{DF} 进行上采样操作, 生成高分辨率特征 F_{UP} ,

图像重构组件, 用于对生成的高分辨率特征 F_{UP} 通过一个卷积层映射成SR图像输出。

2. 根据权利要求1所述基于多阶段注意力增强网络的图像超分辨率方法, 其特征在于, 所述级联残差组件模块包括一个增强U-net块, 多个残差块, 一个群池化块: 增强U-net块利用了不同分辨率的特征图中互补的多尺度特征和自相似性, 在大视场内找到图像中代表全局的结构, 从而有效地融合局部和全局特征, 以获得更好的上下文表示; 残差块进一步改善特征信息流, 提取丰富的局部特征; 在将残差块生成的特征图拼接后, 进一步在一个级联残差组件模块内进行全局特征融合, 从而自适应地保存内容层级的特征信息。

3. 根据权利要求2所述基于多阶段注意力增强网络的图像超分辨率方法, 其特征在于, 第 g 个级联残差组件模块表示为:

$$F_g = F_{g-1} + F_{g,p} = F_{g-1} + W_{GPB} H_g(F_{g-1})$$

其中, W_{GPB} 表示设置给群池化块的卷积层的权重, $H(\cdot)$ 表示第 g 个级联残差组件模块的函数, F_{g-1} , F_g 表示第 g 个级联残差组件模块的输入和输出, $F_{g,p}$ 表示群池化块的输出。

4. 根据权利要求2所述基于多阶段注意力增强网络的图像超分辨率方法, 其特征在于, 所述增强U-net块是将多分辨率注意力机制整合到具有残差结构的基本U-net中形成, 以全局方式充分利用层级特征, 可将浅层特征和深层特征合并在一起; 基于该增强U-net块的结构, 第 m 个自下而上块生成的深层特征—高分辨率特征, 用如下等式来表示:

$$F_D^m = H_D^m(F_D^{m-1}, T_{MRA}(F_E^m))$$

其中, $H_D^m(\cdot)$ 是第 m 个自下而上块的函数, $T_{MRA}(\cdot)$ 表示多分辨率注意力的重新缩放函数, F_E^m 表示第 m 个自上而下块的输出, 为低层次特征, F_D^{m-1} 表示前面的第 $m-1$ 个自下而上块的输出, 为低分辨率特征图。

5. 根据权利要求4所述基于多阶段注意力增强网络的图像超分辨率方法, 其特征在于, 所述多分辨率注意力机制, 在全局范围内确认存在互相补充的特征图的并进行重新校准, 根据U-Net结构的编码器和解码器之间特征图的变化统计量计算通道权重, 然后使用得到的权重因子来有选择性增强更多内容信息的特征图并抑制无用特征; 将编码器的大小为 $H \times W$ 的 C 个特征图 $F_E = [f_1^E, \dots, f_C^E]$ 作为输入, 给定与 F_E 相同的尺寸, 解码器输出特征图 $F_D = [f_1^E, \dots, f_C^E]$ 作为计算参考模板, 池化矢量 $z^E \in R^C$ 的第 c 个尺度计算如下:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (f_c^E(i, j) - \bar{f}_c^D(i, j))^2$$

其中, \bar{f}_c^D 是 F_D 的平均特征图的第 c 个通道输出, R 表示实数集合, (i, j) 表示图像中像素

点的坐标；

使用Sigmoid函数作为权重门函数，得到权重集合为：

$$w = f(W_U \delta(W_D z^E))$$

W_U 和 W_D 表示卷积层的权重集，分别将输出特征的通道维度变为 C 和 C/r ， $f(\cdot)$ 和 $\delta(\cdot)$ 分别表示Sigmoid门和ReLU， r 表示减少的特征通道数，最后获得通道注意力权重图 w 以重新调整输入：

$$\tilde{f}_c^E = w_c \cdot f_c^E$$

其中， w_c 和 f_c^E 分别表示第 c 通道的权重因子和特征图。

6. 根据权利要求5所述基于多阶段注意力增强网络的图像超分辨率方法，其特征在于，所述群池化块与多支流视觉注意力机制结合，通过操作缩放多个网络流的特征图，以全局方式利用特征间的相互依赖性关系，找到对高分辨率图像重构有重要影响的特征图； B 个特征流被送入群池化块中，群池化投票因子 $S \in R^{BC}$ 计算为：

$$s_c = \sum_{i=1}^H \sum_{j=1}^W y_c(i, j) / \sum_{c=1}^C \sum_{i=1}^H \sum_{j=1}^W Y(i, j, c)$$

其中， s_c 表示第 c 个特征图的投票因子， $y_c(i, j)$ 表示第 c 个特征图，

$Y(\cdot)$ 表示与群池化块的所有通道连接中的特征图， R^{BC} 表示实向量空间。

7. 根据权利要求6所述基于多阶段注意力增强网络的图像超分辨率方法，其特征在于，与内容重要性排序的注意力机制结合，通过测量不同特征图之间的内容相似度，使更多信息以粗糙至精细的方式传输到网络重建层，以提高重建图像的视觉质量；其对特征图的相似度计算如下：

$$q_c = \frac{(t_c^M)^T \cdot \bar{t}^M}{\|t_c^M\| \|\bar{t}^M\|}$$

其中， t_i^M 和 \bar{t}^M 表示每个特征图的最大池化过程，通过 $1/C \cdot \sum t_i$ 计算平均特征图 \bar{t} ，将 t_c^M 和 \bar{t}^M 向量化以计算特征图得分 q_c ， $T = [t_1, \dots, t_C]$ 表示大小为 $H \times W$ 的 C 个特征图，上标 T 表示转置。

一种基于多阶段注意力增强网络的图像超分辨率方法

技术领域

[0001] 本发明涉及图像超分辨率技术领域,特别是涉及一种基于多阶段注意力增强网络的图像超分辨率方法。

背景技术

[0002] 单帧图像超分辨率(SISR)旨在从其对应的低分辨率(LR)输入图像中重建高分辨率(HR)图像。SISR技术有望最大程度地减少重建误差并产生良好的主观视觉效果。HR图像可以提供更多所需信息。因此,图像超分辨率(SR)在许多计算机视觉任务中具有重要应用,例如视频监控、卫星成像和消费者摄影。对于同一LR图像,由于HR图像的估计具有多种解决方案,因此SISR是一个典型的不适定反问题。另外,包括噪声和模糊在内的成像质量下降问题将使HR图像的恢复过程更加困难。因此,有必要探索可靠地自然图像先验知识,以及HR-LR图像对的正确映射,以找到具有高视觉质量的最佳解决方案。

[0003] 常用的图像插值方法通过各种插值内核来预测未知像素值来生成HR图像。尽管它们计算简单且快速,但是由于对相邻像素加权的過程易于抑制高频图像分量,因此它们的插值性能受到限制。为了提供更多有用重建信息,基于重建的方法采用各种正则化技术作为先验知识。但是,它们仍然很难恢复精细的图像结构和纹理,因为以启发方式获得的先验知识不足以代表自然图像的多样化模式。为了获得更好的视觉质量,学习类的方法通过从HR-LR示例图像或图像对中学习映射关系来解决SR问题。尽管这些方法获得了良好的视觉质量,但是它们的建模能力相对于大规模训练数据仍然有限,这将限制图像重建质量的提高。当前,用于SR任务的深度学习技术已比以前的SR方法取得了更加优越的性能。由于使用纯数据驱动的方式,使用端到端训练模式的深度卷积神经网络(CNN)具有强大的特征表示能力。目前已有一系列基于CNN的SISR方法,从LR输入及其对应的HR训练样本中学习映射函数。现有的大多数基于CNN的SR模型主要通过设计更深或更广的网络以学习更多具有可区分性的高级特征,而忽略了跨不同尺度和通道特征间的交互,因此缺乏对特征关系的表示能力并且限制了图像SR结果的准确性和稳定性。

发明内容

[0004] 本发明的目的是针对现有技术中存在的技术缺陷,而提供一种基于多阶段注意力增强网络的图像超分辨率方法,以进一步的提高单帧图像超分辨率的性能。

[0005] 为实现本发明的目的所采用的技术方案是:

[0006] 一种基于多阶段注意力增强网络的图像超分辨率方法,通过多阶段注意力增强网络实现,所述多阶段注意力增强网络包括:

[0007] 浅层特征提取模块,用于从LR输入中提取浅层特征 F_0 ,

[0008] 基于增强U-net模块的级联残差组件,包括多级级联残差组件模块,用于将所述浅层特征 F_0 输入到多级CRG模块中,以提取深层特征 F_{DF} ;

[0009] 上采样模块,用于对提取到的深层特征 F_{DF} 进行上采样操作,生成高分辨率特征

F_{UP} ,

[0010] 图像重构组件,用于对生成的高分辨率特征 F_{UP} 通过一个卷积层映射成SR图像输出。

[0011] 其中,所述级联残差组件模块包括一个增强U-net块,多个残差块,一个群池化块:增强U-net块利用了不同分辨率的特征图中互补的多尺度特征和自相似性,在大视场内找到图像中代表全局的结构,从而有效地融合局部和全局特征,以获得更好的上下文表示;残差块进一步改善特征信息流,提取丰富的局部特征;在将残差块生成的特征图拼接后,进一步在一个级联残差组件模块内进行全局特征融合,从而自适应地保存内容层级的特征信息。

[0012] 其中,第 g 个级联残差组件模块表示为:

[0013] $F_g = F_{g-1} + F_{g,P} = F_{g-1} + W_{GPB} H_g(F_{g-1})$

[0014] 其中, W_{GPB} 表示设置给群池化块的卷积层的权重, $H(\cdot)$ 表示第 g 个级联残差组件模块的函数, F_{g-1} , F_g 表示第 g 个级联残差组件模块的输入和输出, $F_{g,P}$ 表示群池化块的输出。

[0015] 其中,所述增强U-net块是将多分辨率注意力机制整合到具有残差结构的基本U-net中形成,以全局方式充分利用层级特征,可将浅层特征和深层特征合并在一起;基于该增强U-net块的结构,第 m 个自下而上块生成的深层特征--高分辨率特征,用如下等式来表示:

[0016] $F_D^m = H_D^m(F_D^{m-1}, T_{MRA}(F_E^m))$

[0017] 其中, $H_D^m(\cdot)$ 是第 m 个自下而上块的函数, $T_{MRA}(\cdot)$ 表示多分辨率注意力的重新缩放函数, F_E^m 表示第 m 个自上而下块的输出,为低层次特征, F_D^{m-1} 表示前面的第 $m-1$ 个自下而上块的输出,为低分辨率特征图。

[0018] 其中,所述多分辨率注意力机制,在全局范围内确认存在互相补充的特征图的并进行重新校准,根据U-Net结构的编码器和解码器之间特征图的变化统计量计算通道权重,然后使用得到的权重因子来有选择性增强更多内容信息的特征图并抑制无用特征;将编码器的大小为 $H \times W$ 的 C 个特征图 $F_E = [f_1^E, \dots, f_C^E]$ 作为输入,给定与 F_E 相同的尺寸,解码器输出特征图 $F_D = [f_1^E, \dots, f_C^E]$ 作为计算参考模板,池化矢量 $z^E \in R^c$ 的第 c 个尺度计算如下:

[0019]
$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (f_c^E(i, j) - \bar{f}_c^D(i, j))^2$$

[0020] 其中, \bar{f}_c^D 是 F_D 的平均特征图的第 c 个通道输出, R 表示实数集合, (i, j) 表示图像中像素点的坐标;

[0021] 使用Sigmoid函数作为权重门函数,得到权重集合为:

[0022] $w = f(W_U \delta(W_D z^E))$

[0023] W_U 和 W_D 表示卷积层的权重集,分别将输出特征的通道维度变为 C 和 C/r , $f(\cdot)$ 和 $\delta(\cdot)$ 分别表示Sigmoid门和ReLU, r 表示减少的特征通道数,最后获得通道注意力权重图 w 以重新调整输入:

$$[0024] \quad \tilde{f}_c^E = w_c \cdot f_c^E$$

[0025] 其中, w_c 和 f_c^E 分别表示第 c 通道的权重因子和特征图。

[0026] 其中, 所述群池化块与多支流视觉注意力机制结合, 通过操作缩放多个网络流的特征图, 以全局方式利用特征间的相互依赖性关系, 找到对高分辨率图像重构有重要影响的特征图; B 个特征流被送入群池化块中, 群池化投票因子 $S \in R^{BC}$ 计算为:

$$[0027] \quad s_c = \sum_{i=1}^H \sum_{j=1}^W y_c(i, j) / \sum_{c=1}^C \sum_{i=1}^H \sum_{j=1}^W Y(i, j, c)$$

[0028] 其中, s_c 表示第 c 个特征图的投票因子, $y_c(i, j)$ 表示第 c 个特征图, $Y(\cdot)$ 表示与群池化块的所有通道连接中的特征图, R^{BC} 表示实向量空间。

[0029] 其中, 与对内容重要性排序的注意力机制结合, 通过测量不同特征图之间的内容相似度, 使更多信息以粗糙至精细的方式传输到网络重建层, 以提高重建图像的视觉质量; 其对特征图的相似度计算如下:

$$[0030] \quad q_c = \frac{(t_c^M)^T \cdot \bar{t}^M}{\|t_c^M\| \|\bar{t}^M\|}$$

[0031] 其中, t_i^M 和 \bar{t}^M 表示每个特征图的最大池化过程, 通过 $1/C \cdot \sum t_i$ 计算平均特征图 \bar{t} , 将 t_c^M 和 \bar{t}^M 向量化以计算特征图得分 q_c , $T = [t_1, \dots, t_C]$ 表示大小为 $H \times W$ 的 C 个特征图, 上标 T 表示转置。

[0032] 本发明提出的多阶段注意力增强网络 MAAN, 用于学习视觉特征的共同表示和交互; 通过将增强 U-net 组件堆叠成深度残差结构中, 将高分辨率特征和低分辨率内容信息结合起来, 以提供有用的特征。另外, 利用多分辨率、多流和内容排序注意力机制来增强特征学习能力, 使得本发明可以充分利用多尺度和多层次特征之间的互补信息。

附图说明

[0033] 图1是本发明的多阶段注意力增强网络的整体结构网络图;

[0034] 图2为本发明的三种注意力模块融合到网络多阶段处理模块的融合示意图。

具体实施方式

[0035] 以下结合附图和具体实施例对本发明作进一步详细说明。应当理解, 此处所描述的具体实施例仅仅用以解释本发明, 并不用于限定本发明。

[0036] 本发明提出了一个用于单帧图像超分辨的多阶段注意力增强网络, 通过学习跨尺度和跨通道的全局内容感知特征组来重建高分辨率图像。所提出的网络采用了多个增强 U-net 组件并使用级联残差连接, 不仅允许大量的低频信息通过多个层间跳接进行流动, 而且还使主干网络能够专注于学习大感受野范围内的抽象特征表示。

[0037] 此外, 进一步的, 本发明还提出多分辨率、多支流和内容重要性排序的注意力模块, 并将其融合到网络多阶段处理模块中, 通过使用上述的三种注意力模块, 每个通道特征

均可自适应地重新缩放和融合,以在多个表示阶段生成更逼真的高分辨率图像内容。

[0038] 如图1-2所示,本发明基于多阶段注意力增强网络的图像超分辨率方法,通过多阶段注意力增强网络MAAN实现,所述多阶段注意力增强网络包括:

[0039] 浅层特征提取模块,基于增强U-net模块的级联残差组件(CRG),上采样模块(Upscale Module)和图像重构组件。

[0040] 本发明用 I_{LR} 和 I_{SR} 分别表示MAAN的输入和输出。本发明仅用一个卷积层(Conv)来从LR输入中提取浅层特征 F_0

[0041] $F_0 = H_{SF}(I_{LR})$ (1)

[0042] 其中 $H_{SF}(\cdot)$ 表示浅层模块操作。提取到的浅层特征 F_0 被输入到多级CRG模块中,用于提取深层特征。因此,可以进一步得到以下表达式

[0043] $F_{DF} = H_{CRG}(F_0)$ (2)

[0044] 其中, $H_{CRG}(\cdot)$ 表示深层级联残差U-net结构。此主网络包含总数为G的多级CRGs,可以堆叠成很深的网络,从而提供更大的感受野。提取到的深层特征通过一个Upscale Module来进行上采样操作。

[0045] $F_{UP} = H_{UP}(F_{DF})$ (3)

[0046] 生成的高分辨率特征 F_{UP} 通过一个卷积层映射成SR图像

[0047] $I_{SR} = H_{REC}(F_{UP}) = H_{MAAN}(I_{LR})$ (4)

[0048] 其中, $H_{REC}(\cdot)$ 和 $H_{MAAN}(\cdot)$ 分别表示重建层以及整个MAAN网络的函数。

[0049] 本发明中,所述的级联残差组件(CRG),如图1所示,每个CRG主要包含一个增强U-net块(EUB),B个残差块(RBs),一个群池化块(GPB),以及级联残差(RIC)连接。EUB利用了不同分辨率的特征图中互补的多尺度特征和自相似性。此外,它可以在大的视场内找到图像中代表全局的结构,从而可以有效地融合局部和全局特征,以获得更好的上下文表示。在CRG中引入了基本残差块(RBs)以进一步有效地改善特征信息流,因为RB中有多个卷积层可以提取丰富的局部特征。在将RB生成的特征图进行拼接后,进一步在一个CRG内进行全局特征融合,从而自适应地保存内容层级的特征信息。

[0050] 为了进一步减轻CRG的权重参数,便于稳定训练,通过级联几个残差结构实现一种RIC来连接各个模块。

[0051] 本发明采用具有残差结构的CRG作为更深层网络的基本结构。进一步引入CRG内的级联机制来使深层网络的训练保持稳定。将第g个CRG表示为

[0052] $F_g = F_{g-1} + F_{g,P} = F_{g-1} + W_{GPB}H_g(F_{g-1})$ (5)

[0053] 其中 W_{GPB} 表示设置给GPB的卷积层的权重,不失一般性,为简单起见,省略了偏置项。 $H(\cdot)$ 表示第g个CRG的函数。 F_{g-1} , F_g 表示第g个CRG的输入和输出, $F_{g,P}$ 表示GPB的输出。深层特征 F_{DF} 通过如下等式得到

[0054] $F_{DF} = F_0 + W_{LSC}F_g$ (6)

[0055] 其中, W_{LSC} 是在主网络尾部卷积层的权重, F_g 表示通过主干网络全部CRG处理后的特征输出。这种多阶段结构不仅可以简化跨CRG的信息流,还可以为生成HR图像提供非常深的CNN结构。

[0056] 本发明中,所述的增强U-net块(EUB),如图1所示,本发明将多分辨率注意力(MRA)整合到具有残差结构的基本U-net(BUN)中。U-Net结构本质上是一个编码器和解码器网络,

其中许多自上而下和自下而上的块通过一系列密集的跳跃路径连接在一起。这种结构鼓励完全重用信息流并捕获具有不同分辨率的多尺度特征。 F_E^m 表示第m个自上而下块的输出， F_D^{m-1} 表示前面的第m-1个自下而上块的输出。基于EUB的结构，第m个自下而上块生成的深层特征可用如下等式来表示

$$[0057] \quad F_D^m = H_D^m \left(F_D^{m-1}, T_{MRA} \left(F_E^m \right) \right) \quad (7)$$

[0058] 其中 $H_D^m(\bullet)$ 是第m个自下而上块的函数。 $T_{MRA}(\bullet)$ 表示MRA的重新缩放函数。通过从低分辨率特征图 F_D^{m-1} 和低层次特征 F_E^m 产生高分辨率特征 F_D^m ，第m个自下而上阶段的卷积块可以捕获长距离信息。EUB的 T_{MRA} 操作可以使中间特征图在每个图像尺度内在语义上是有区别的。另外，补充信息被提取并融合以突出显著特征。EUB的残差结构以全局方式充分利用层级特征，可以将浅层特征和深层特征合并在一起。

[0059] 本发明，还提出一种多分辨率视觉注意力机制(MRA)，如图2所示，将多分辨率视觉注意力机制(MRA)模块与EUB融合集成在一起。MRA的目的是在全局范围内确认存在互相补充的特征图的并进行重新校准。根据编码器和解码器之间特征图的变化统计量计算通道权重，然后使用得到的权重因子来有选择性增强更多内容信息的特征图并抑制无用特征。将编码器的大小为 $H \times W$ 的C个特征图 $F_E = [f_1^E, \dots, f_C^E]$ 作为输入。给定与 F_E 相同的尺寸，解码器输出的特征图 $F_D = [f_1^D, \dots, f_C^D]$ 作为计算参考模板。池化矢量 $z^E \in R^C$ 的第c个尺度计算如下

$$[0060] \quad z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \left(f_c^E(i, j) - \bar{f}_c^D(i, j) \right)^2 \quad (8)$$

[0061] 其中， \bar{f}_c^D 是 F_D 的平均特征图的第c个通道输出。

[0062] 图2中， F_D 的平均操作 H_{GA} 可计算为 $1/C \cdot \sum f_i^D$ ，全局平均池化 H_{VA} 根据等式(8)在各个特征通道上沿空间维度进行操作。为利用聚合信息，通过全局池化进一步构建注意力关系，使用Sigmoid函数作为权重门函数，得到权重集合为：

$$[0063] \quad w = f(W_U \delta(W_D z^E)) \quad (9)$$

[0064] W_U 和 W_D 表示卷积层的权重集，分别将输出特征的通道维度变为C和C/r。 $f(\bullet)$ 和 $\delta(\bullet)$ 分别表示Sigmoid门和ReLU。最后获得通道注意力权重图w以重新调整输入：

$$[0065] \quad \tilde{f}_c^E = w_c \cdot f_c^E \quad (10)$$

[0066] 其中， w_c 和 f_c^E 分别表示第c通道的权重因子和特征图。通过使用这种通道注意力，可自适应地增强EUB中不同分辨率的特征并进行有效融合。

[0067] 本发明中，还提出提出一种多支流视觉注意力机制(MSA)，如图2所示。将多支流视觉注意力机制(MSA)模块与群池化块(GBP)融合，通过函数 H_{GP} 操作缩放多个网络流的特征图。这种注意力机制不仅减小参数，而且以全局方式利用特征间的相互依赖性关系。多支流视觉注意力机制(MSA)有助于找到对高分辨率图像重构有重要影响的特征图。假设B个特征流被送入GPB中。群池化投票因子 $S \in R^{BC}$ 计算为：

$$[0068] \quad s_c = \sum_{i=1}^H \sum_{j=1}^W y_c(i, j) / \sum_{c=1}^C \sum_{i=1}^H \sum_{j=1}^W Y(i, j, c) \quad (11)$$

[0069] 其中, $Y(\cdot)$ 表示与GPB的所有通道连接中的特征图。通过使用式 (9) 和 (10) 中类似方法, 可获取第 c 个通道中全局缩放因子和特征图以重新缩放特征流。

[0070] 本发明中, 还提出一种对内容重要性排序的注意力机制 (CRA), 如图2所示, 利用 H_{CS} 操作通过测量不同特征图之间的内容相似度来实现CRA模块。这种自注意力特征图包含了较多的重要内容信息, 增强了中频和高频段的图像特征, 这些特征通常代表边缘、纹理和其他细节, 会显著提高重建图像的视觉质量。 $T = [t_1, \dots, t_c]$ 表示大小为 $H \times W$ 的 C 个特征图。通过 $1/C \cdot \sum t_i$ 计算平均特征图 \bar{t} 。用 t_i^M 和 \bar{t}^M 表示每个特征图的最大池化过程, 通常选择将特征图的大小减小至 $1/4$ 。

[0071] 特征图的相似度计算如下:

$$[0072] \quad q_c = \frac{(t_c^M)^T \cdot \bar{t}^M}{\|t_c^M\| \|\bar{t}^M\|} \quad (12)$$

[0073] 将 t_c^M 和 \bar{t}^M 向量化以计算特征图得分 q_c , 可根据式 (9) 和 (10) 获得加权特征图。最后, 更多信息以粗糙至精细的方式传输到网络重建层。

[0074] 以上所述仅是本发明的优选实施方式, 应当指出的是, 对于本技术领域的普通技术人员来说, 在不脱离本发明原理的前提下, 还可以做出若干改进和润饰, 这些改进和润饰也应视为本发明的保护范围。

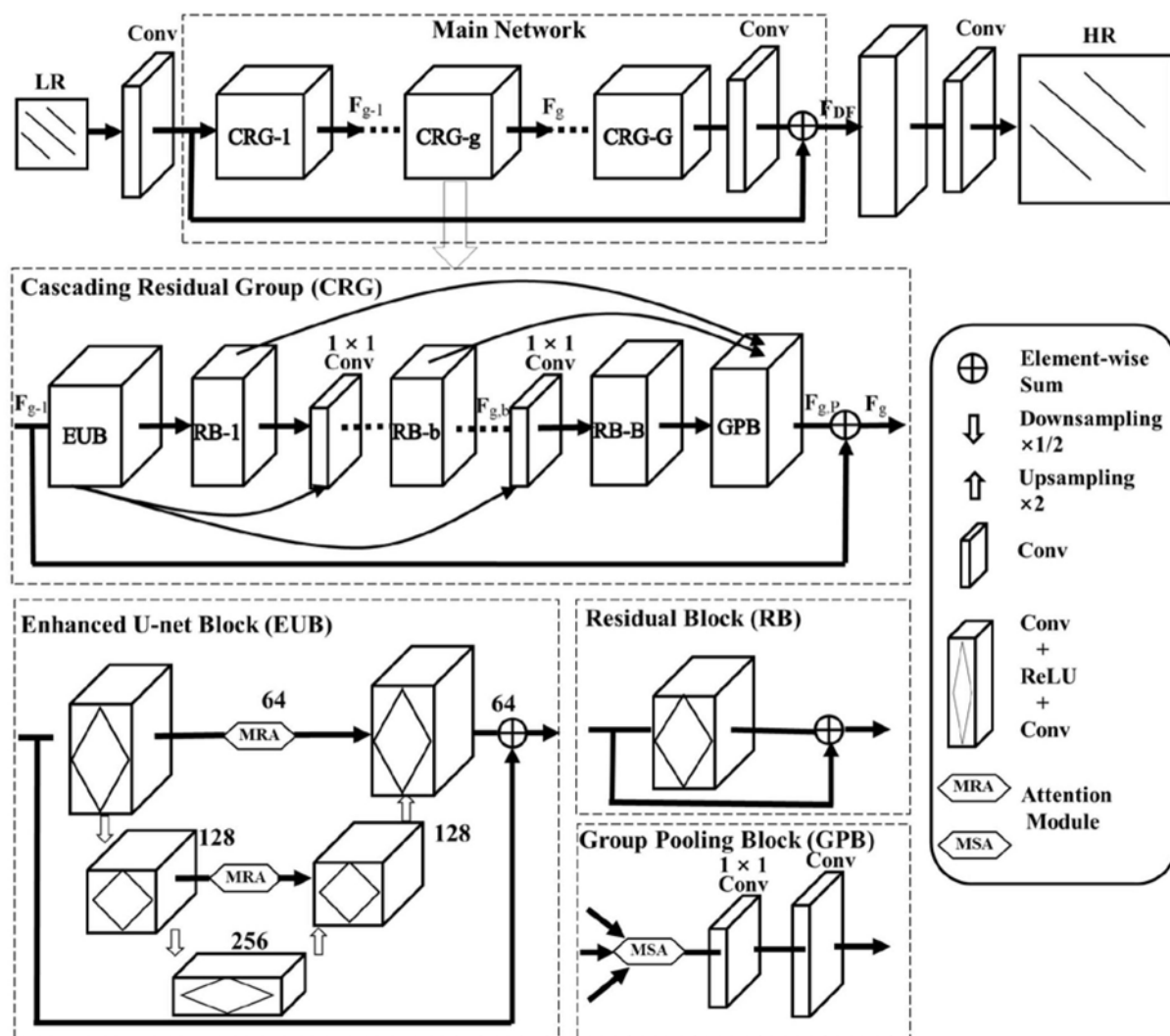


图1

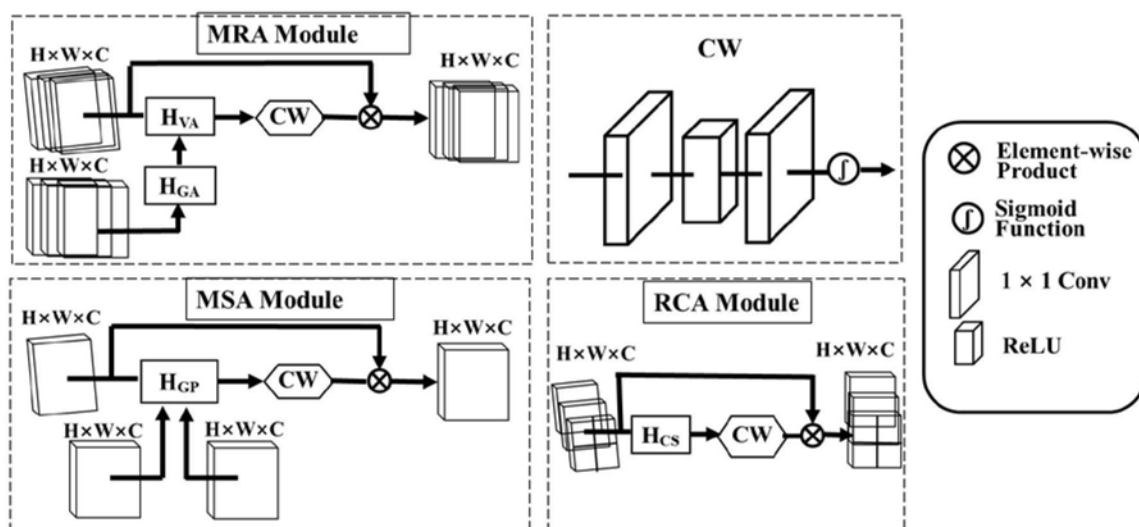


图2