

(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 111738921 A

(43) 申请公布日 2020. 10. 02

(21) 申请号 202010541526.6

(22) 申请日 2020.06.15

(71) 申请人 大连理工大学

地址 116024 辽宁省大连市甘井子区凌工
路2号

(72) 发明人 叶昕辰 徐睿 孙宝利 王智慧
李豪杰

(74) 专利代理机构 大连理工大学专利中心
21200

代理人 陈玲玉

(51) Int.Cl.

G06T 3/40 (2006.01)

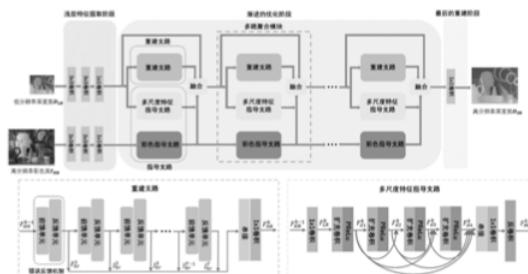
权利要求书2页 说明书5页 附图2页

(54) 发明名称

基于深度神经网络的多信息渐进融合的深
度超分辨率方法

(57) 摘要

本发明公开了基于深度神经网络的多信息渐进融合的深度超分辨率方法，属于图像处理领域。该方法提出了一个多路网络结构分别提取了深度特征信息与彩色信息，并且将多种信息有效地融合并且充分利用。彩色图像及其对应的深度图在同一场景的光度和几何表征都具有很强的结构相似性，使用彩色图像辅助深度图像的恢复，更好的预测深度图上模糊的边界。本发明程序简单，易于实现，能够获得高质量的高分辨率的深度图像，有效地解决了深度图细微结构和小尺度目标区域难以精确地恢复的问题，得到的深度图清晰，目标边界锋利。



CN 111738921 A

CN 111738921 A

权 利 要 求 书

1/2 页

1. 基于深度神经网络的多信息渐进融合的深度超分辨率方法,其特征在于,包括如下步骤:

第一步,准备初始数据:训练数据包括低分辨率与高分辨率的深度图数据对,以及对应的高分辨率的彩色图像;

第二步,重建支路的构建,利用错误反馈机制以及特征注意力机制,构建基于注意力机制的错误反馈网络模块;重建支路由4个基于注意力机制的错误反馈模块串联组成;

第三步,指导支路的构建,指导支路包括多尺度特征指导支路和彩色指导支路;多尺度特征指导支路利用扩充率为1,2,3,4的空洞卷积在不改变特征分辨率的情况下获取多尺度的感受野;彩色指导支路利用了三层普通卷积提取彩色特征;

第四步,由3个多路聚合模块组成整体网络,每个多路聚合模块利用特征通道注意力机制融合重建支路与指导支路的信息;

第五步,构建损失函数并训练网络;

构建损失函数并训练网络:损失函数 \mathcal{L} 衡量训练数据上由彩色图像 I_{HR} 和低分辨率深度图 D_{LR} 预测的高分辨率深度图 D_{SR} 和真实高分辨率深度图 D_{HR} 的差距;

$$\mathcal{L} = \|D_{SR} - D_{HR}\|_1$$

其中, D_{SR} 表示网络预测的结果; $\|\cdot\|_1$ 表示1范数;网络训练过程即是在训练数据不断优化网络参数使 \mathcal{L} 收敛得到最终的网络模型。

2. 根据权利要求1所述的基于深度神经网络的多信息渐进融合的深度超分辨率方法,其特征在于,第二步,重建支路的构建,利用错误反馈机制以及特征注意力机制,构建了基于注意力机制的错误反馈网络模块;包括以下步骤:

2-1) 基于注意力机制的错误反馈网络模块由前馈单元与反馈单元组成;在第 t 个基于注意力机制的错误反馈网络模块中, $t=1,2,3,4$;前馈单元以上一个基于注意力机制的错误反馈网络模块的输出特征 F_{hr}^{t-1} 作为输入,当 $t=1$ 时,第一个基于注意力机制的错误反馈网络模块的输入为经过浅层特征提取的深度特征;经过一个卷积操作进行下采样得到特征 F_{lr1}^t ;再经过一个平均池化操作得到特征 $F_{pooling}^t$,其中 $F_{pooling}^t$ 的尺寸为 F_{lr1}^t 的 $\frac{1}{4}$;将平均池化后的特征 $F_{pooling}^t$ 通过反卷积上采样得到 F_{lr2}^t , F_{lr2}^t 与 F_{lr1}^t 的尺寸相同;对 F_{lr2}^t 与 F_{lr1}^t 进行相减再通过激活函数得到上采样误差 W_{Res1}^t :

$$W_{Res1}^t = PReLU(F_{lr2}^t - F_{lr1}^t)$$

式中,PReLU(\cdot)为激活函数;最后在前馈单元中,将误差 W_{Res1}^t 与 F_{lr1}^t 进行相乘,使网络更加关注误差大的特征区域,前馈单元的输出为:

$$F_{lr3}^t = F_{lr1}^t + \gamma(F_{lr1}^t * W_{Res1}^t)$$

式中, γ 是一个平衡系数,本发明中 $\gamma=0.1$;*代表对应位置的元素相乘;

2-2) 反馈单元,首先将前馈单元的输出特征 F_{lr3}^t 通过反卷积上采样到目标尺寸得到特征 F_{hr1}^t ;经过一个卷积操作进行下采样得到特征 F_{lr4}^t ;再通过一个反卷积操作上采样得到

CN 111738921 A

权 利 要 求 书

2/2 页

F_{hr2}^t ; 对 F_{hr2}^t 与 F_{hr1}^t 进行相减再通过激活函数得到上采样误差 W_{Res2}^t :

$$W_{Res2}^t = PReLU(F_{hr2}^t - F_{hr1}^t)$$

最后在反馈单元中, 将误差 W_{Res2}^t 与 F_{hr1}^t 进行相乘, 反馈单元的最后输出为:

$$F_{hr}^t = F_{hr1}^t + \gamma(F_{hr1}^t * W_{Res2}^t)。$$

3. 根据权利要求1所述的基于深度神经网络的多信息渐进融合的深度超分辨率方法, 其特征在于, 第三步, 多尺度特征指导支路的构建, 包括以下步骤:

3-1) 多尺度特征指导支路利用扩充率为1,2,3,4的空洞卷积在不改变特征分辨率的情况下获取多尺度的感受野; 经过每层空洞卷积得到特征 $F_{d_i}^k$:

$$F_{d_i}^k = DilatConv([F_{d_1}^k, \dots, F_{d_{i-1}}^k], i), i = 2, 3, 4.$$

式中, $DilatConv(\cdot)$ 表示空洞卷积操作, i 即表示第 i 个空洞卷积, d_i 表示对应的空洞率; 最后串接每个空洞卷积的输出, 通过一个卷积减少特征通道, 再通过一个反卷积操作上采样到目标尺寸, 得到多尺度特征指导支路的输出 F_{MB}^k :

$$F_{MB}^k = DeConv(Conv_{1 \times 1}([F_{d_1}^k, F_{d_2}^k, F_{d_3}^k, F_{d_4}^k]))$$

式中, $DeConv(\cdot)$ 表示反卷积操作;

3-2) 彩色指导支路利用了三层普通卷积提取了彩色特征; 在每个彩色指导支路使用3层连续的卷积操作, 得到彩色指导特征 F_{CB}^k 。

4. 根据权利要求3所述的基于深度神经网络的多信息渐进融合的深度超分辨率方法, 其特征在于, 第四步, 多路聚合模块, 融合重建支路与指导支路的信息; 重建支路由4个基于注意力机制的错误反馈模块串联组成, 将每个错误反馈机制的输出串接, 再通过一个卷积操作进行初步融合, 得到第 k 个, $k=1, 2, 3$, 多路融合模块中重建支路的输出 F_{RB}^k :

$$F_{RB}^k = Conv_{1 \times 1}^k([F_{hr}^{1,k}, F_{hr}^{2,k}, \dots, F_{hr}^{T,k}])$$

式中, $Conv_{1 \times 1}$ 表示卷积核大小为 1×1 的卷积操作, $[\cdot]$ 为特征图的串接; T 为基于注意力机制的错误反馈网络模块的最大串联个数;

得到融合后的特征 F_{HR}^k :

$$F_{HR}^k = Fusion([F_{RB}^k, F_{MB}^k, F_{CB}^k, F_{HR}^{k-1}])$$

式中 $Fusion(\cdot)$ 表示特征通道注意力机制, F_{HR}^{k-1} 为上一个多路聚合模块的输出。

基于深度神经网络的多信息渐进融合的深度超分辨率方法

技术领域

[0001] 本发明属于图像处理及计算机视觉领域,涉及采用深度卷积神经网络对低分辨率的深度图像进行超分辨率恢复,并使用与深度图像对应的彩色图像指导进行恢复以求得高质量高分辨率深度图,具体涉及一种基于深度神经网络的多信息渐进融合的场景深度超分辨率的方法。

背景技术

[0002] 场景深度作为基础的信息广泛用于三维(3D)建模,姿态估计以及辅助驾驶等。然而,在实际环境中,深度图的获取仍然是一项具有挑战性的任务,需要投入大量的精力和高成本的专业设备,如时间飞行(Time of flight,ToF)深度传感器,微软的Kinect相机或双目相机设备。由于实际场景的环境复杂性和深度传感器的限制,高精度和高分辨率的深度图几乎无法直接从深度传感器中获取,这也限制了其在实际应用中的使用。例如,广泛使用的微软的第二代Kinect相机(Kinect2)获取的深度图像的分辨率为 512×424 ,而对应的彩色图像的分辨率为 1920×1080 。因此,由低分辨率的深度图获取高质量的高分辨率深度图的超分辨率技术对于实际应用显得十分重要。

[0003] 随着深度学习的发展,利用深度卷积神经网络可以学习一个从低分辨率的深度图获取高分辨率的深度图的映射模型。通常,基于深度学习获取高分辨率深度图的方法根据输入数据的不同分为两种,一种方法是只使用低分辨的深度图作为输入(Gernot Riegler, Matthias Rüther, and Horst Bischof. Atgv-net: Accurate depth super-resolution. In ECCV, pages 268-284, 2016);另一种方法是使用低分辨率的深度图和对应的高分辨率的彩色图像,彩色图像作为辅助信息恢复高分辨率的深度图(Tak-Wai Hui, Chen Change Loy, and Xiaou Tang. Depth map super-resolution by deep multi-scale guidance. In ECCV, pages 353-369, 2016)。彩色图像及其对应的深度图在同一场景的光度和几何表征都具有很强的结构相似性。因此,现有的深度图超分辨率的方法大多以颜色信息为指导来恢复退化深度图。虽然上述基于卷积神经网络(convolutional neural network,CNN)的方法表现出了令人印象深刻的性能,但是在处理深度细节的恢复时,结果并不令人满意。而且对于彩色图像信息的使用也不够充分。基于以上问题,本发明设计了一个基于深度神经网络的高质量的深度图超分辨率框架。

发明内容

[0004] 本发明旨在克服现有技术的不足,提出了一个基于深度神经网络进行高质量深度图超分辨率的多信息渐进融合框架。该框架是一个多信息渐进融合的多路结构,由多个多路融合模块组成,本发明中使用3个多路融合模块,每个多路融合模块包括2个部分:重建支路和指导支路。重建支路利用基于错误反馈机制(Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Deep back-projection networks for super-resolution. In IEEE CVPR, pages 1664-1673, 2018.)搭建的网络结构模拟图像恢复过程,

使网络关注于恢复误差大的图像区域,逐渐地减小恢复误差。指导支路可以学习多种先验知识用于辅助深度图细节的恢复。最后,设计融合模块用于融合重建支路和指导支路,抽取具有选择判别性的特征。

[0005] 本发明的技术方案为,一种基于深度神经网络的多信息渐进融合的场景深度超分辨率方法,所述方法包括下列步骤:

[0006] 第一步,准备初始数据:训练数据包括低分辨率与高分辨率的深度图数据对,以及对应的高分辨率的彩色图像;

[0007] 2)重建支路的构建:

[0008] 2-1)利用错误反馈机制 (Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Deepback-projection networks for super-resolution. In IEEE CVPR, pages 1664-1673, 2018.) 以及特征注意力机制,构建了基于注意力机制的错误反馈网络模块。

[0009] 2-2)将4个相同的基于注意力机制的错误反馈模块串联组成重建支路。

[0010] 3)指导支路的构建:指导支路由2个不同功能的指导支路构成,即多尺度特征指导支路和彩色指导支路。

[0011] 3-1)多尺度特征指导支路利用扩充率为1,2,3,4的空洞卷积 (Fisher Yu and Vladlen Koltun, "Multi-scale context aggregation by dilated convolutions," CoRR, vol. abs/1511.07122, 2015.) 在不改变特征分辨率的情况下获取获得多尺度的感受野。

[0012] 3-2)彩色指导支路利用了三层普通卷积提取彩色图上的对应特征。

[0013] 4)由3个多路聚合模块组成整体网络,每个多路聚合模块利用特征通道注意力机制 (Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Learning a discriminative feature network for semantic segmentation. In IEEE CVPR, pages 1857-1866, 2018.) 融合了重建支路与指导支路的信息。

[0014] 5)构建损失函数并训练网络:损失函数 \mathcal{L} 衡量训练数据上由彩色图像 I_{HR} 和低分辨率深度图 D_{LR} 预测的高分辨率深度图 D_{SR} 和真实高分辨率深度图 D_{HR} 的差距。

[0015] $\mathcal{L} = \|D_{SR} - D_{HR}\|_1$

[0016] 其中, D_{SR} 表示网络预测的结果。 $\|\cdot\|_1$ 表示1范数。网络训练过程即是在训练数据不断优化网络参数使 \mathcal{L} 收敛得到最终的网络模型。

[0017] 本发明的有益效果是:

[0018] 本发明基于深度神经网络搭建的一个多信息渐进融合的深度图超分辨率的框架,最终得到高质量的高分辨率的深度图像,具有以下特点:

[0019] 1、程序简单,易于实现,能够获得高质量的高分辨率的深度图像;

[0020] 2、本方法将场景深度图像恢复当作一个渐近地恢复的过程,由多个多路聚合模块逐渐地恢复高质量的深度图;

[0021] 3、本方法使用多尺度特征学习的结构,有效地解决了深度图细微结构和小尺度目标区域难以精确地恢复的问题。

[0022] 4、本方法结合彩色图像与深度图内部结构一致性的先验指导信息,得到的深度图清晰且边界锋利。

附图说明

- [0023] 图1是具体实施流程图。
- [0024] 图2是基于特征注意力的错误反馈机制的框架图。
- [0025] 图3是初始数据样例。其中：(a) 低质量深度图；(b) 高分辨率真实深度图；(c) 高分辨率彩色图。
- [0026] 图4是深度恢复结果以及与其他方法的比较，其中：(a) 高分辨率深度图 (b) 低分辨率深度图；(c) 真实的高分辨率深度图；(d) Shuhang Gu等人的结果 (Shuhang Gu, WangmengZuo, Shi Guo, Yunjin Chen, Chongyu Chen, and Lei Zhang. Learning dynamic guidance for depth image enhancement. In IEEE CVPR, pages 712-721, 2017)；(e) Xinchen Ye等人的结果 (Xinchen Ye, XiangyueDuan, and Haojie Li. Depth super-resolution with deep edge-inference network and edge-guided depth filling. In IEEE ICASSP, pages 1398-1402, 2018)；(f) 本发明的结果。

具体实施方式

- [0027] 下面结合实施例和附图对本发明的基于深度神经网络的多信息渐进融合的场景深度超分辨率方法做出详细说明。
- [0028] 一种基于深度神经网络的多信息渐进融合的场景深度超分辨率方法，如图1所示，包括下列步骤：
- [0029] 1) 准备初始数据；
- [0030] 初始数据包括低分辨率和高分辨率深度图，以及相对应的高分辨率彩色图，其中一组数据如图3所示。为了训练网络，数据集使用Middlebury官方数据 (<http://vision.middlebury.edu>)，其中38张彩色-深度图像作为训练数据集，6张彩色-深度图像作为测试数据集。首先预处理训练数据集，从训练彩色-深度图像对中以截取 256×256 的彩色-深度图像块。通过双三次插值算法下采样 256×256 的深度图到 $128 \times 128, 64 \times 64, 32 \times 32$ 和 16×16 ，分别用于2倍，4倍，8倍，16倍的超分辨率任务，最终形成15500组图像对用于训练。
- [0031] 2) 重建支路的构建：
- [0032] 2-1) 利用错误反馈机制 (Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Deepback-projection networks for super-resolution. In IEEE CVPR, pages 1664-1673, 2018.) 以及特征注意力机制，构建了基于注意力机制的错误反馈网络模块，如图2所示。
- [0033] 2-1-1) 如图2所示，基于注意力机制的错误反馈网络模块由前馈单元与反馈单元组成。在第t个基于注意力机制的错误反馈网络模块中 ($t=1, 2, 3, 4$)，前馈单元以上一个基于注意力机制的错误反馈网络模块的输出特征 F_{hr}^{t-1} 作为输入（当 $t=1$ 时，第一个基于注意力机制的错误反馈网络模块的输入为经过浅层特征提取的深度特征，如图1所示），经过一个卷积操作进行下采样得到特征 F_{lr1}^t ；再经过一个平均池化操作得到特征 $F_{pooling}^t$ ，其中 $F_{pooling}^t$ 的尺寸为 F_{lr1}^t 的 $\frac{1}{4}$ ；将平均池化后的特征 $F_{pooling}^t$ 通过反卷积上采样得到 F_{lr2}^t, F_{lr2}^t

CN 111738921 A

说 明 书

4/5 页

与 F_{lr1}^t 的尺寸相同;对 F_{lr2}^t 与 F_{lr1}^t 进行相减再通过激活函数得到上采样误差 W_{Res1}^t :

$$[0034] \quad W_{Res1}^t = PRelu(F_{lr2}^t - F_{lr1}^t)$$

[0035] 式中,PRelu(\cdot)为激活函数(Parametric rectified linear unit,PRelu)。最后在前馈单元中,将误差 W_{Res1}^t 与 F_{lr1}^t 进行相乘,使网络更加关注误差大的特征区域,前馈单元的输出为:

$$[0036] \quad F_{lr3}^t = F_{lr1}^t + \gamma(F_{lr1}^t * W_{Res1}^t)$$

[0037] 式中,γ是一个平衡系数,本发明中 $\gamma=0.1$.*代表对应位置的元素相乘。

[0038] 2-1-2)反馈单元,首先将前馈单元的输出特征 F_{lr3}^t 通过反卷积上采样到目标分辨率得到特征 F_{hr1}^t ;经过一个卷积操作进行下采样得到特征 F_{hr4}^t ;再通过一个反卷积操作上采样得到 F_{hr2}^t ;对 F_{hr2}^t 与 F_{hr1}^t 进行相减再通过激活函数得到上采样误差 W_{Res2}^t :

$$[0039] \quad W_{Res2}^t = PRelu(F_{hr2}^t - F_{hr1}^t)$$

[0040] 最后在反馈单元中,将误差 W_{Res2}^t 与 F_{hr1}^t 进行相乘,反馈单元的最后输出为:

$$[0041] \quad F_{hr}^t = F_{hr1}^t + \gamma(F_{hr1}^t * W_{Res2}^t)$$

[0042] 2-2)重建支路由4个迭代的基于注意力机制的错误反馈模块串联组成,将每个错误反馈机制模块的输出串接再一起,再通过一个卷积操作进行初步融合,得到第k个多路融合模块中重建支路的输出 F_{RB}^k :

$$[0043] \quad F_{RB}^k = Conv_{1 \times 1}^k([F_{hr}^{1,k}, F_{hr}^{2,k}, \dots, F_{hr}^{T,k}])$$

[0044] 式中,Conv_{1×1}^k表示卷积核大小为1×1的卷积操作,[\cdot]为特征图的串接。T为基于注意力机制的错误反馈网络模块的最大串联个数,取值为4。本发明中k=1,2,3。

[0045] 3)指导支路的构建:指导支路由2个不同功能的指导支路构成:多尺度特征指导支路和彩色指导支路。

[0046] 3-1)多尺度特征指导支路利用扩充率为1,2,3,4的空洞卷积(Fisher Yu and Vladlen Koltun,“Multi-scale context aggregation by dilated convolutions,”CoRR, vol.abs/1511.07122,2015.)在不改变特征分辨率的情况下获取获得多尺度的感受野。经过每层空洞卷积得到特征 $F_{d_i}^k$:

$$[0047] \quad F_{d_i}^k = DilatConv([F_{d_1}^k, \dots, F_{d_{i-1}}^k], i), i = 2, 3, 4.$$

[0048] 式中,DilatConv(\cdot)表示空洞卷积操作,i即表示第i个空洞卷积,d_i表示对应的空洞率。最后我们串接每个空洞卷积的输出,通过一个卷积减少特征通道,再通过一个反卷积操作上采样到目标尺寸,得到多尺度特征指导支路的输出 F_{MB}^k :

$$[0049] \quad F_{MB}^k = DeConv(Conv_{1 \times 1}([F_{d_1}^k, F_{d_2}^k, F_{d_3}^k, F_{d_4}^k]))$$

[0050] 式中,DeConv(\cdot)表示反卷积操作。

CN 111738921 A

说 明 书

5/5 页

[0051] 3-2) 彩色指导支路利用了三层普通卷积提取了彩色特征。在每个彩色指导支路使用3层连续的卷积操作,卷积核的大小分别是 3×3 , 3×3 和 1×1 ,得到彩色指导特征 F_{CB}^k 。

[0052] 4) 融合模块的构建:利用特征通道注意力机制 (Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Learning a discriminative feature network for semantic segmentation. In IEEE CVPR, pages 1857-1866, 2018.) 融合了重建支路与指导支路的信息。得到融合后的特征 F_{HR}^k :

$$[0053] F_{HR}^k = Fusion([F_{RB}^k, F_{MB}^k, F_{CB}^k, F_{HR}^{k-1}])$$

[0054] 式中Fusion(\bullet)表示特征通道注意力机制, F_{HR}^{k-1} 为上一个多路聚合模块的输出。

[0055] 5) 构建损失函数并训练网络:损失函数 \mathcal{L} 衡量训练数据上由彩色图像 I_{HR} 和低分辨率深度图 D_{LR} 预测的高分辨率深度图 D_{SR} 和真实高分辨率深度图 D_{HR} 的差距。

$$[0056] \mathcal{L} = \|D_{SR} - D_{HR}\|_1$$

[0057] 其中, D_{SR} 表示网络预测的结果。 $\|\cdot\|_1$ 表示1范数。网络训练过程即是在训练数据不断优化网络参数使 \mathcal{L} 收敛得到最终的网络模型。

[0058] 本方法对一组数据的恢复结果以及与其他先进的方法的比较如图4所示,其中(a)高分辨率深度图(b)低分辨率深度图;(c)真实的高分辨率深度图;(d)Shuhang Gu等人(Shuhang Gu, Wangmeng Zuo, Shi Guo, Yunjin Chen, Chongyu Chen, and Lei Zhang. Learning dynamic guidance for depth image enhancement. In IEEE CVPR, pages 712-721, 2017)的结果;(e)Xinchen Ye等人(Xinchen Ye, Xiangyue Duan, and Haojie Li. Depth super-resolution with deep edge-inference network and edge-guided depth filling. In IEEE ICASSP, pages 1398-1402, 2018)的结果;(f)本发明的结果。结果标明本发明的性能优于其他方法。

CN 111738921 A

说 明 书 附 图

1/2 页

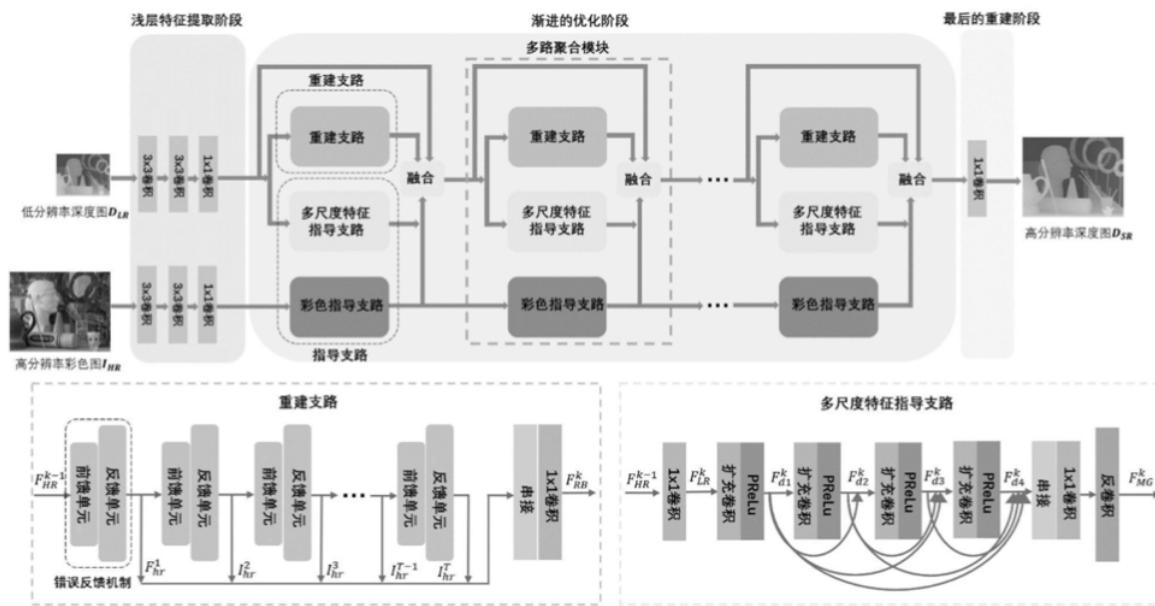


图1

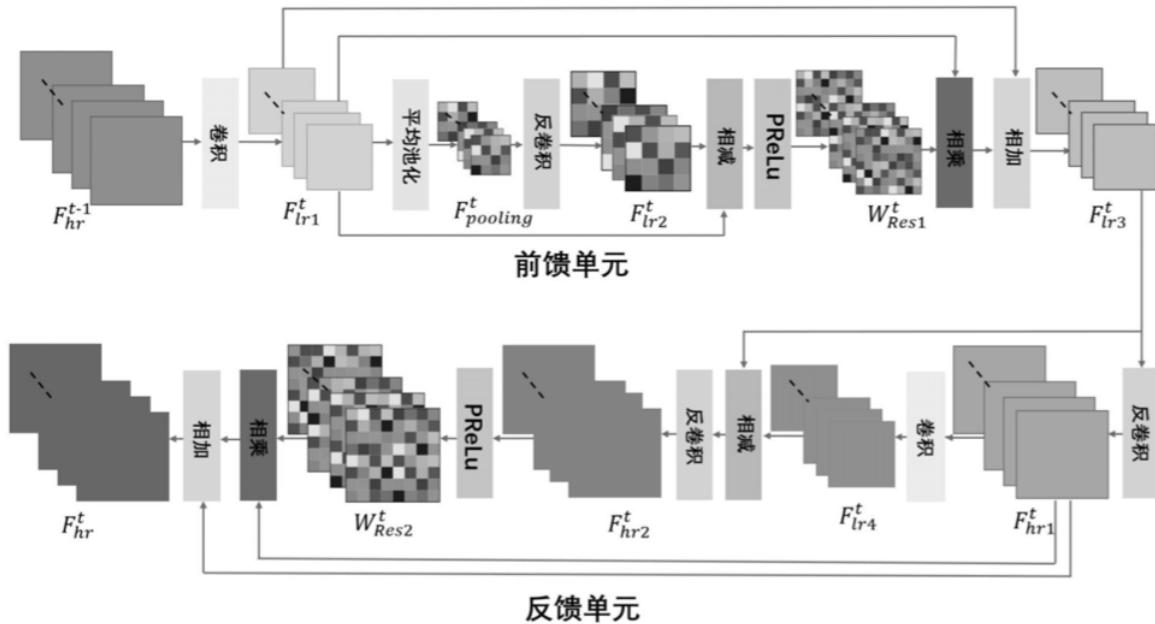


图2

Q

CN 111738921 A

说 明 书 附 图

2/2 页

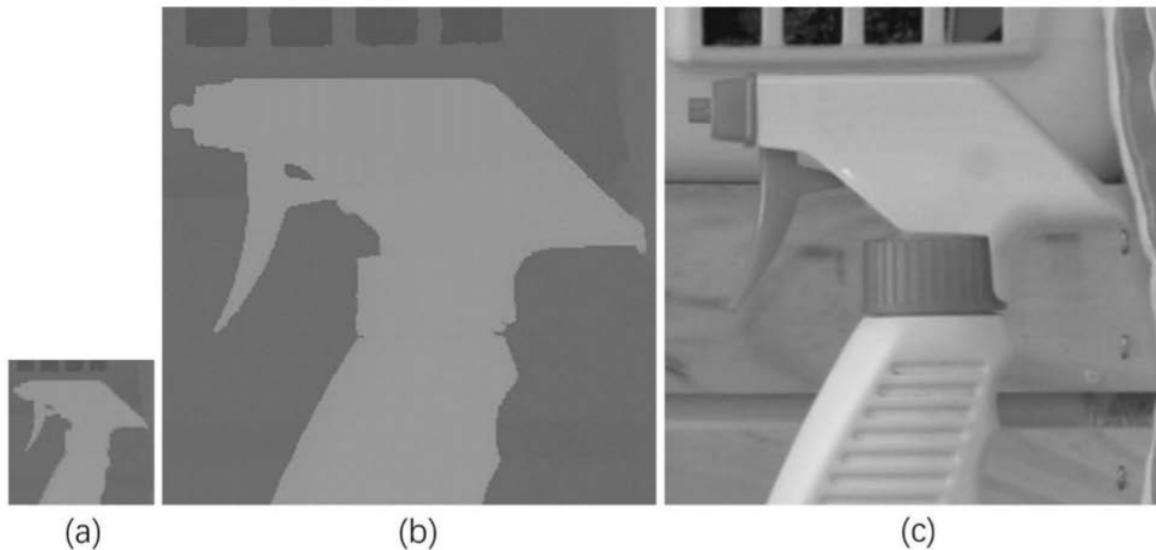


图3

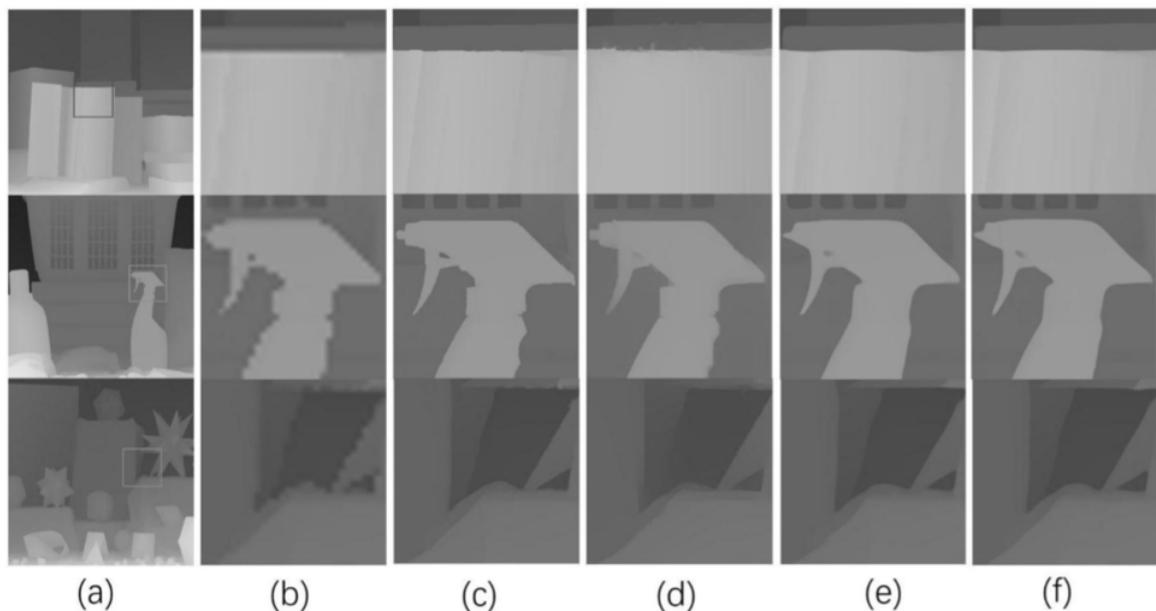


图4

A blank white page with a few small black specks, likely scanning artifacts or dust.