



(12)发明专利申请

(10)申请公布号 CN 109118431 A  
(43)申请公布日 2019.01.01

(21)申请号 201811031483.6

(22)申请日 2018.09.05

(71)申请人 武汉大学

地址 430072 湖北省武汉市武昌区珞珈山  
武汉大学

(72)发明人 王中元 易鹏 江奎 韩镇

(74)专利代理机构 武汉科皓知识产权代理事务  
所(特殊普通合伙) 42222

代理人 魏波

(51) Int. Cl.

G06T 3/40(2006.01)

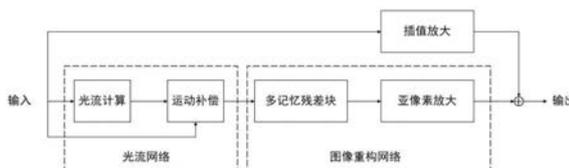
权利要求书2页 说明书4页 附图1页

(54)发明名称

一种基于多记忆及混合损失的超分辨率重建方法

(57)摘要

本发明公开了一种基于多记忆及混合损失的超分辨率重建方法,包括光流网络和图像重构网络两个部分。在光流网络中,对于输入的多帧,计算当前帧与参考帧之间的光流,并利用光流作运动补偿,将当前帧尽可能补偿到与参考帧相似。在图像重构网络中,将补偿后的多帧依次输进网络,网络采用多记忆残差块提取图像特征,使得后面输入帧能接收到前面帧的特征图信息。最后,将输出的低分辨率特征图进行亚像素放大,并与双立方插值放大后的图像相加,得到最终的高分辨率视频帧。训练过程采用一种混合损失函数,对光流网络和图像重构网络同时进行训练。本发明极大地增强了帧间信息融合的特征表达能力,能够重建出细节真实丰富的高分辨率视频。



1. 一种基于多记忆及混合损失的超分辨率重建方法,其特征在于,包括以下步骤:

步骤1:选取若干视频数据作为训练样本,从每个视频帧中相同的位置截取大小为 $N \times N$ 像素的图像作为超分辨率学习目标,将其下采样 $r$ 倍,得到大小为 $M \times M$ 的低分辨率图像,作为网络的输入,其中, $N = M \times r$ ;

步骤2:将 $2n+1$ 张时间连续的低分辨率视频图像输入光流网络,作为低分辨率输入帧,而处于中心位置的低分辨率图像帧作为低分辨率参考帧;依次计算每个低分辨率输入帧与低分辨率参考帧之间的光流,并使用光流对每个低分辨率输入帧作运动补偿,获得低分辨率补偿帧;其中, $n \geq 0$ ;

步骤3:将低分辨率补偿帧输入图像重构网络,利用多记忆残差块进行帧间信息融合,得到残差特征图;

步骤4:采用混合损失函数,对光流网络和图像重构网络同时进行约束,并进行反向传播学习;

步骤5:将步骤3中得到的残差特征图放大,获得超分辨率残差图像,并将参考帧放大,获得超分辨率插值图像;

步骤6:将步骤5中得到的超分辨率插值图像与超分辨率残差图像相加,得到超分辨率视频帧。

2. 根据权利要求1所述的基于多记忆及混合损失的超分辨率重建方法,其特征在于:步骤2中,采用从粗粒度到细粒度的方法提取光流,并使用运动补偿算子对输入帧进行运动补偿。

3. 根据权利要求1所述的基于多记忆及混合损失的超分辨率重建方法,其特征在于:步骤3中,采用多记忆残差块,存储当前帧的特征信息,以便与下一帧进行特征信息融合;

$$I_{n+1} = \{I_n, 0_n\} = \{I_n, \text{ConvLSTM}_n(I_n)\} \quad (1)$$

其中, $\text{ConvLSTM}_n(\cdot)$ 表示多记忆残差块中第 $n$ 个卷积记忆块, $I_n$ 表第 $n$ 个卷积记忆块的输入, $0_n$ 表示对应的输出;将 $I_n$ 与 $0_n$ 作连结,得到 $I_{n+1}$ ,即第 $n+1$ 个卷积记忆块的输入。

4. 根据权利要求1所述的基于多记忆及混合损失的超分辨率重建方法,其特征在于:步骤4中,采用混合损失函数,同时约束光流网络和图像重构网络,并进行训练;

$$\mathcal{L}_{JR} = \sum_{i=-T}^T \lambda_i |I_0^H - SR(J_i)| \quad (2)$$

$$\mathcal{L}_{OF} = \sum_{i=-T}^T |I_i^L - \widetilde{I_{0 \rightarrow i}^L}|_1 + \alpha \|\nabla F_{i \rightarrow 0}\|_1 \quad (3)$$

$$\mathcal{L} = \mathcal{L}_{JR} + \beta \mathcal{L}_{OF} \quad (4)$$

其中, $\mathcal{L}_{JR}$ 与 $\mathcal{L}_{OF}$ 分别表示图像重构网络与光流网络的损失函数;公式(2)中, $i$ 表示时间步, $T$ 代表时间步的最大范围; $SR(\cdot)$ 代表超分辨率这个过程, $J_i$ 表示输入的第 $i$ 个补偿帧; $I_0^H$ 表示未下采样的高分辨率参考帧, $\lambda_i$ 是第 $i$ 个时间步长的权重;公式(3)中, $I_i^L$ 是第 $i$ 个低分辨率帧, $\widetilde{I_{0 \rightarrow i}^L}$ 表示根据光流场 $F_{i \rightarrow 0}$ 作用而成的补偿帧 $I_0^L$ , $\nabla F_{i \rightarrow 0}$ 表示光流场 $F_{i \rightarrow 0}$ 的全变分, $\alpha$ 是一个惩罚项约束参数;最后将 $\mathcal{L}_{JR}$ 与 $\mathcal{L}_{OF}$ 结合起来,得到公式(4)中的混合损失函数 $\mathcal{L}$ , $\beta$ 表示参数。

5. 根据权利要求1所述的基于多记忆及混合损失的超分辨率重建方法,其特征在

于:步骤5中,对输出的残差特征图采用亚像素放大,对低分辨率参考帧采用双立方插值放大;

其中,亚像素放大的过程表示如下:

$$\begin{aligned}\text{Dim}(I) &= H \times W \times N_0 \\ &= H \times W \times r \times r \times N_1 \\ &= H \times r \times W \times r \times N_1\end{aligned}\quad (5)$$

其中, $\text{Dim}(\cdot)$ 表示一个张量的维度, $I$ 代表输入张量, $H$ 与 $W$ 分别为张量 $I$ 的高和宽, $N_0$ 则是张量 $I$ 的特征图数量, $r$ 表示放大倍数;对该张量进行公式(5)所示的变形操作,便得到高和宽各放大了 $r$ 倍后的张量;其中, $N_0 = N_1 \times r \times r$ 。

## 一种基于多记忆及混合损失的视频超分辨率重建方法

### 技术领域

[0001] 本发明属于数字图像处理技术领域,涉及一种视频超分辨率重建方法,具体涉及一种多记忆的混合损失函数约束的超分辨率重建方法。

### 背景技术

[0002] 近年来,随着高清显示设备(如HDTV)的出现以及4K(3840×2160)和8K(7680×4320)等超高清视频分辨率格式的出现,由低分辨率视频重建出高分辨率视频的需求日益增加。视频超分辨率是指从给定的低分辨率视频重建高分辨率视频的技术,广泛应用于高清电视、卫星图像、视频监控等领域。

[0003] 目前,应用最广泛的超分辨率方法是基于插值的方法,如最近邻插值,双线性插值以及双三次插值。这种方法通过将固定的卷积核应用于给定的低分辨率图像输入,来计算高分辨率图像中的未知像素值。因为这种方法只需要少量的计算,所以它们的速度非常快。但是,它们的重建效果也欠佳,特别是在重构高频信息较多的图像区域。近年来,为了找到更好的方式来重建丢失的信息,研究人员们开始致力于研究基于样本的方法,也称为基于学习的方法。最近,Dong等人率先提出基于卷积神经网络的超分辨率方法,该方法具有从众多多样化图像样本中学习细节的能力,因而备受关注。

[0004] 单张图像超分辨率是指利用一张低分辨率的图像,重构出其对应的高分辨率图像。与之相比,视频超分辨率则是利用多张有关联性的低分辨率视频帧,重建出它们对应的高分辨率视频帧。除了利用单张图像内部的空间相关性,视频超分辨率更重视利用低分辨率视频帧之间的时间相关性。

[0005] 传统的视频超分辨率算法利用图像先验知识,来进行像素级的运动补偿和模糊核估计,以此重建高分辨率视频。然而,这些方法通常需要较多计算资源,并且难处理高倍率放大倍数或大幅帧间相对运动的情况。

[0006] 最近,基于卷积神经网络的视频超分辨率方法已经出现,这种方法直接学习从低分辨率帧到高分辨率帧之间的映射关系。Tao等人提出了细节保持的深度视频超分辨率网络,他们设计出了一种亚像素运动补偿层,将低分辨率帧映射到高分辨率栅格上。然而,亚像素运动补偿层需要消耗大量显存,其效果却十分有限。Liu等人设计了一个时间自适应神经网络,来自适应地学习时间依赖性的最优尺度,但目前只是设计了一个简单的三层卷积神经网络结构,从而限制了性能。

### 发明内容

[0007] 为了解决上述技术问题,本发明提供了一种基于多记忆残差块和混合损失函数约束的超分辨率重建方法,在图像重构网络中插入多记忆残差块,更有效地利用帧间的时间相关性和帧内的空间相关性。并利用混合损失函数,同时约束光流网络和图像重构网络,进一步提高网络的性能,提取更真实丰富的细节。

[0008] 本发明所采用的技术方案是:一种基于多记忆及混合损失的视频超分辨率重建方

法,其特征在于,包括以下步骤:

[0009] 步骤1:选取若干视频数据作为训练样本,从每个视频帧中相同的位置截取大小为 $N \times N$ 像素的图像作为高分辨率学习目标,将其下采样 $r$ 倍,得到大小为 $M \times M$ 的低分辨率图像,作为网络的输入,其中, $N = M \times r$ ;

[0010] 步骤2:将 $2n+1$  ( $n \geq 0$ )张时间连续的低分辨率视频图像输入光流网络,作为低分辨率输入帧,而处于中心位置的低分辨率图像帧作为低分辨率参考帧。依次计算每个低分辨率输入帧与低分辨率参考帧之间的光流,并使用光流对每个低分辨率输入帧作运动补偿,获得低分辨率补偿帧;

[0011] 步骤3:将低分辨率补偿帧输入图像重构网络,利用多记忆残差块进行帧间信息融合,得到残差特征图;

[0012] 步骤4:采用混合损失函数,对光流网络和图像重构网络同时进行约束,并进行反向传播学习;

[0013] 步骤5:将步骤3中得到的残差特征图放大,获得高分辨率残差图像,并将参考帧放大,获得高分辨率插值图像;

[0014] 步骤6:将步骤5中得到的高分辨插值图像与高分辨率残差图像相加,得到超分辨率视频帧。

[0015] 本发明使用了多记忆残差块,极大的增强了网络的特征表达能力,同时采用混合损失函数约束网络训练,因而不仅能重构出逼真丰富的图像细节,而且网络训练过程收敛速度快。

## 附图说明

[0016] 图1为本发明的网络整体框架简图。

## 具体实施方式

[0017] 为了便于本领域普通技术人员理解和实施本发明,下面结合附图及实施例对本发明作进一步的详细描述,应当理解,此处所描述的实施例仅用于说明和解释本发明,并不用于限定本发明。

[0018] 请见图1,本发明提供一种卫星影像超分辨率重建方法,其特征在于,包括以下步骤:

[0019] 一种基于多记忆及混合损失的超分辨率重建方法,其特征在于,包括以下步骤:

[0020] 步骤1:选取若干视频数据作为训练样本,从每个视频帧中相同的位置截取大小为 $N \times N$ 像素的图像作为高分辨率学习目标,将其下采样 $r$ 倍,得到大小为 $M \times M$ 的低分辨率图像,作为网络的输入,其中, $N = M \times r$ ;

[0021] 步骤2:将 $2n+1$  ( $n \geq 0$ )张时间连续的低分辨率视频图像输入光流网络,作为低分辨率输入帧,而处于中心位置的低分辨率图像帧作为低分辨率参考帧。依次计算每个低分辨率输入帧与低分辨率参考帧之间的光流,并使用光流对每个低分辨率输入帧作运动补偿,获得低分辨率补偿帧;

[0022] 本发明采用一种采用现有的从粗粒度到细粒度的方法提取光流,并使用现有的运

动补偿算子对输入帧进行运动补偿。

[0023] 以四倍超分辨率为例。首先计算粗粒度光流,将双线性放大四倍后的当前帧与参考帧输入网络,使用两次步长为2的卷积,此时光流的尺寸为目标高分辨率图像的四分之一,再用亚像素放大将计算的光流放大到目标高分辨率,并进行运动补偿。然后计算细粒度光流,将双线性放大四倍后的当前帧与参考帧,以及粗粒度计算得到的光流与补偿帧输入网络,但这次只使用一次步长为2的卷积,此时光流的尺寸为目标高分辨率的二分之一,再用亚像素放大将计算的光流放大到目标高分辨率,并进行运动补偿。

[0024] 步骤3:将低分辨率补偿帧输入图像重构网络,利用多记忆残差块进行帧间信息融合,得到残差特征图;

[0025] 本发明采用一种多记忆残差块,存储当前帧的特征信息,以便与下一帧进行特征信息融合。

$$[0026] \quad I_{n+1} = \{I_n, O_n\} = \{I_n, \text{ConvLSTM}_n(I_n)\} \quad (1)$$

[0027] 其中,ConvLSTM<sub>n</sub>表示多记忆残差块中第n个卷积记忆块,I<sub>n</sub>表第n个卷积记忆块的输入,O<sub>n</sub>表示对应的输出。将I<sub>n</sub>与O<sub>n</sub>作连结,得到I<sub>n+1</sub>,即第n+1个卷积记忆块的输入。

[0028] 步骤4:采用混合损失函数,对光流网络和图像重构网络同时进行约束,并进行反向传播学习;

[0029] 本发明采用两种损失函数,同时约束光流网络和图像重构网络,并进行训练;

$$[0030] \quad \mathcal{L}_{JR} = \sum_{i=-T}^T \lambda_i |I_0^H - SR(J_i)| \quad (2)$$

$$[0031] \quad \mathcal{L}_{OF} = \sum_{i=-T}^T |I_i^L - \widetilde{I}_{0 \rightarrow i}^L|_1 + \alpha \|\nabla F_{i \rightarrow 0}\|_1 \quad (3)$$

$$[0032] \quad \mathcal{L} = \mathcal{L}_{JR} + \beta \mathcal{L}_{OF} \quad (4)$$

[0033] 其中, $\mathcal{L}_{JR}$ 与 $\mathcal{L}_{OF}$ 分别表示图像重构网络与光流网络的损失函数;公式(2)中,i表示时间步,T代表时间步的最大范围;SR(·)代表超分辨率这个过程,J<sub>i</sub>表示输入的第i个补偿帧; $I_0^H$ 表示未下采样的高分辨率参考帧, $\lambda_i$ 是第i个时间步长的权重;公式(3)中, $I_i^L$ 是第i个低分辨率帧, $\widetilde{I}_{0 \rightarrow i}^L$ 表示根据光流场 $F_{i \rightarrow 0}$ 作用而成的补偿帧 $I_0^L$ , $\nabla F_{i \rightarrow 0}$ 表示光流场 $F_{i \rightarrow 0}$ 的全变分, $\alpha$ 是一个惩罚项约束参数;最后将 $\mathcal{L}_{JR}$ 与 $\mathcal{L}_{OF}$ 结合起来,得到公式(4)中的混合损失函数 $\mathcal{L}$ , $\beta$ 表示参数。

[0034] 步骤5:将步骤3中得到的残差特征图放大,获得高分辨率残差图像,并将参考帧放大,获得高分辨率插值图像;

[0035] 本发明采用亚像素放大,利用特征图的深度信息重构高分辨率图像的空间信息,不同于传统的转置卷积,能提取更丰富的图像细节;将低分辨率参考帧用双立方插值放大,获得高分辨率插值图像。

[0036] 亚像素放大的过程表示如下:

$$[0037] \quad \text{Dim}(I) = H \times W \times N_0$$

$$[0038] \quad = H \times W \times r \times r \times N_1$$

$$[0039] \quad = H \times r \times W \times r \times N_1 \quad (5)$$

[0040] 其中,Dim(·)表示一个张量的维度,I代表输入张量,H与W分别为张量I的高和宽, $N_0$ 则是张量I的特征图数量,r表示放大倍数。对该张量进行公式(5)所示的变形操作,便可

得到高和宽各放大了 $r$ 倍后的张量。其中,  $N_0 = N_1 \times r \times r$ 。

[0041] 步骤6: 将步骤5中得到的高分辨插值图像与高分辨率残差图像相加, 得到超分辨率视频帧。

[0042] 本发明在光流网络中, 对于输入的多帧, 计算当前帧与参考帧之间的光流, 并利用光流作运动补偿, 将当前帧尽可能补偿到与参考帧相似。在图像重构网络中, 将补偿后的多帧依次输进网络, 网络采用多记忆残差块提取图像特征, 使得后面输入帧能接收到前面帧的特征图信息。最后, 将输出的低分辨率特征图进行亚像素放大, 并与双立方插值放大后的图像相加, 得到最终的高分辨率视频帧。训练过程采用一种混合损失函数, 对光流网络和图像重构网络同时进行训练。本发明极大地增强了帧间信息融合的特征表达能力, 能够重建出细节真实丰富的高分辨率视频。

[0043] 本发明能够同时利用帧内空间相关性和帧间时间相关性来保证超分辨率重建效果。

[0044] 应当理解的是, 本说明书未详细阐述的部分均属于现有技术。

[0045] 应当理解的是, 上述针对较佳实施例的描述较为详细, 并不能因此而认为是对本发明专利保护范围的限制, 本领域的普通技术人员在本发明的启示下, 在不脱离本发明权利要求所保护的范围情况下, 还可以做出替换或变形, 均落入本发明的保护范围之内, 本发明的请求保护范围应以所附权利要求为准。

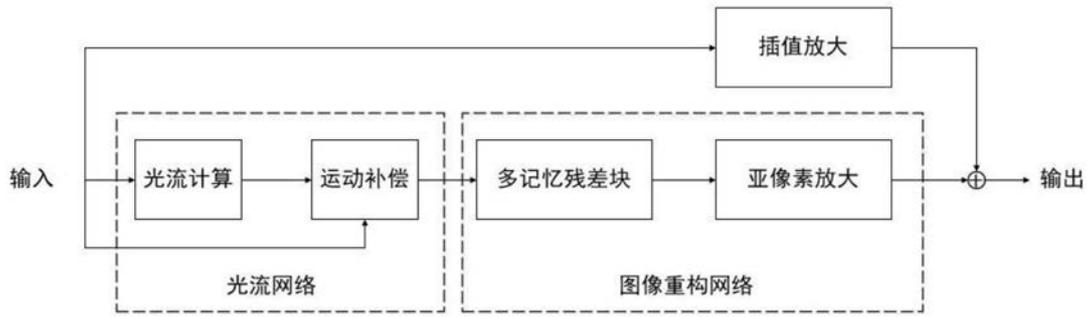


图1