

双向时域特征流盲去运动模糊方法

赵跃进 刘文龙 刘明 董立泉 惠梅

(北京理工大学光电学院, 北京, 100081)

摘要: 便携式成像设备在日常的生活中已经无处不在, 但是因为拍摄时的抖动或者场景中的快速运动物体, 所拍摄的图像或视频经常存在模糊现象, 造成图像重要细节丢失。为了将模糊的视频图像恢复到清晰的状态, 结合近年来的研究热点——生成对抗网络, 本文提出了一种新颖的端到端的双向时域特征流盲去运动模糊方法。该方法充分利用时空连续特征信息, 在三连帧图像间建立双向的时域特征传输通道。多阶段自编码去模糊网络结构和并行编码-混合解码融合方案能够融合三连帧图像多通道内容信息, 并恢复出更加清晰的视频图像。实验结果表明, 在不牺牲较大时间代价前提下, 本文提出的方法在传统的图像质量评价指标(峰值信噪比和结构相似性)和视觉质量上均优于现有的去模糊算法。

关键词: 盲去运动模糊; 生成对抗网络; 时域特征; 自编码

中图分类号: TP391 **文献标志码:** A

Bidirectional Time-Domain Feature Flow Blind Motion Deblurring Algorithm

Zhao Yuejin, Liu Wenlong, Liu Ming, Dong Liqun, Hui Mei

(School of Optics and Photonics, Beijing Institute of Technology, Beijing, 100081, China)

Abstract: Portable imaging devices are ubiquitous in everyday life. However, as the hand jitter or the fast moving objects in the scene during shooting process, the captured image or video is often blurred, causing important details loss. In order to restore the blurred video and image to a clear state, we combine the recent research hotspots—Generative adversarial network, and propose a novel end-to-end bidirectional time-domain feature flow blind motion deblurring algorithm. The algorithm makes full use of the feature information of spatio-temporal continuity constraint to establish a bidirectional transmission channel of time-domain features between the adjacent frames. The multi-stage autoencoder deblurring network structure and the parallel coding and hybrid decoding fusion solution can fuse the multi-channel content information of a frame triplet and restore a clearer frame for a video. Experimental results show that the proposed algorithm is superior to the existing advanced algorithms on the traditional image quality evaluation indexes, i. e., peak signal to noise ratio (PSNR) and structural similarity (SSIM), and visual quality within acceptable time cost.

Key words: blind motion deblurring; generative adversarial network; time-domain feature; autoencoder

引言

随着时代的发展,视频技术及其相关的成像设备取得了长足的发展,在日常生活中扮演着十分重要的角色,近些年摄像设备数量更是呈指数级增长的态势。但是除了昂贵的高速摄像机,普通成像设备如果在曝光时间范围内发生较为明显的抖动或者场景中存在快速运动的物体,所记录的视频图像往往发生模糊现象,导致观感不佳和重要细节丢失。而且曝光时间越长,这种模糊效应越明显。为了去除这种模糊效应,研究人员很早就开始研究如何恢复模糊的视频图像到清晰状态。一般的模糊模型^[1]可以表达为

$$I_B = K \otimes I_S + N \quad (1)$$

式中: I_S 表示理想清晰图像; K 表示作用在 I_S 上的模糊核; \otimes 表示卷积操作; N 表示在成像过程中环境噪声和传感器噪声等加性噪声; I_B 表示在模糊模型作用下产生的模糊图像。由此模糊模型反推去模糊问题可以被建模为反卷积问题。然而,在诸如相机拍摄视频或图像的实际过程中,模糊核形式完全未知,盲去模糊便成为一个严重的病态问题。

目前,大多数的研究根据去模糊对象将去模糊问题大致分为单帧图像去模糊和视频图像去模糊。给定单张模糊图像,最直接的方法是维纳滤波^[2],但是维纳滤波对噪声特别敏感,恢复的图像几乎不可用。因此传统的大多数研究将该问题转为迭代优化问题。Lam 等人^[3]从数理统计的角度利用先验知识在迭代过程中估计和更新反卷积公式以增强反卷积效果。基于稀疏和冗余表征理论, Ma 等人^[4]首先在冗余字典上表征不同模糊和图像,并分别对其表征系数施加稀疏约束后迭代优化。迭代算法计算量巨大,因此 Xu 等人^[5]建立了传统优化方法与神经网络架构之间的联系,所提出的可分离反卷积结构对伪盲去模糊有很好的鲁棒性,但此方法仍限于非盲图像反卷积。随着机器学习特别是卷积神经网络(Convolutional neural networks, CNN)的兴起, Sun 等人^[6]设计一套运动模糊核模板,然后通过分类识别模糊图像块的模糊核类型建立一个非均匀模糊场用于估计潜在清晰图像。而 Nah 等人^[7]最先将生成对抗网络(Generative adversarial network, GAN)引入到图像去模糊问题中,并提出一个多尺度 CNN 结构由粗到精地逐级提升去模糊效果。Kupyn 等人^[1]将 Wasserstein 生成对抗网络(WGAN-GP)^[8]引入到去模糊问题,并提出 DeblurGAN 去模糊网络。视频图像去模糊因为考虑到时间维度信息往往可以取得更好的结果。Kim 等人^[9]提出一动态 CNN 网络用于配准视频帧和一时空循环去模糊网络,可基本实现实时去运动模糊。而后为了处理由相机抖动、移动物体和场景中深度变化等原因引起的空间局部变化的模糊,他们用双向光流近似像素级模糊核提出基于能量的模型,可以同时估算视频光流和潜在清晰帧^[10]。为避免繁琐的人工视频图像配准过程, Su 等人^[11]提出一种累积连续帧之间信息的去模糊网络。

生成对抗网络^[12]近些年来在图像去噪、图像超分辨、图像生成等领域取得了显著的成果。所生成的图像不仅能在视觉感知方面更贴近于人眼的观测,达到以假乱真的效果,而且能从真实数据库中学习到先验知识为图像补充更多真实的内容细节。因此,为了消除视频图像的运动模糊,本文提出一种基于生成对抗网络的双向时域特征流盲运动去模糊网络,其原理图如图 1 所示。本文先提取当前帧的隐含特征,并建立一双向时域特征传输通道,当前帧的隐含特征信息能被传递到前一帧及其后一帧。而在后一阶段,结合当前帧特征信息的前一帧和后一帧的混合编码将会被联合解码以产生与当前帧相对应的清晰图像。此外,本文也

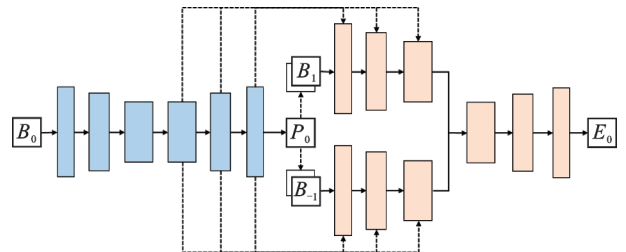


图 1 双向时域特征流盲去运动模糊网络的原理图

Fig. 1 Basic architecture of the bidirectional time-domain feature flow blind motion deblurring network

把感知损失^[13]作为多组分损失函数的一部分,与传统的均方误差(Mean squared error, MSE)或平均绝对误差(Mean absolute deviation, MAE)作为优化目标相比,感知损失促使网络选择偏于视觉质量的解决方案,使得恢复后图像细节更丰富锐利。基于此,本文去模糊模型可以取得非常显著的效果。

1 模糊数据集

深度学习是由大数据驱动,去模糊网络的学习同样依赖于一个高质量的模糊数据集。但是去模糊不像图像超分辨、图像去噪和图像绘色等任务数据集制作方便快捷。为了得到模糊-清晰的样本对,Nah等人^[7]用GoPro高速摄像机拍摄高帧率高清晰的短视频,然后平均相邻帧合成出原始图像的模糊版本,但可供下载的数据只适用于单帧图像去模糊。另外一种随机轨迹生成法^[7]能仿真空间变化的运动模糊核,将清晰图像转为运动模糊效果。但这种方法前后仿真出的模糊帧没有时间上的联系,所以对于视频盲去模糊问题而言,仿真而成的模糊视频序列与真实运动模糊序列存在较大差异。

Su等人^[11]提供了一高帧率视频数据集,该数据集也成为了诸多图像去模糊研究中的通用数据集。本文尝试利用这些高帧率视频制作出一模糊数据集。为了达到满意的运动模糊效果,本文先采用双向光流插帧法在原始的两帧图像之间插入 N 帧子帧,然后再利用邻近子帧合成出当前帧的对应模糊版本。具体操作如图2所示:给定一段视频, f_t 表示在时间 t 的采样帧, f_{t-1} 和 f_{t+1} 表示 f_t 的前一帧和后一帧,通过OpenCV可以建立 $f_{t-1} \leftarrow f_t$ 和 $f_t \rightarrow f_{t+1}$ 的双向光流场,再在每一个方向的光流中计算出 N 帧子帧,所得到的图像序列如式(2)所示。

$$(f_{t-1}, f_{t-1}^N, \dots, f_{t-1}^1, f_t, f_t^1, \dots, f_{t+1}^N, f_{t+1}) \quad (2)$$

最后本文利用式(3)计算出 f_t 对应的模糊帧 b_t

$$b_t = \frac{1}{2L+1} \sum_{i=1}^L (f_t + f_{t-1}^i + f_{t+1}^i) \quad (3)$$

式中 L 表示在 f_t 前后各取 L 帧子帧。

通过在两帧图像中插入子帧,一段低帧率的视频可以转变为任意高帧率,再不受摄像机帧率限制。而通过平均子帧合成的模糊视频与原始清晰视频具有相同的帧率。在实际的转换过程中,考虑到视频转换速度和所合成的模糊效果,本文设置 $N=20$, $L=10$ 。以三连续模糊帧作为负样本,而以中间模糊帧对应的清晰帧作为正样本,由正负样本可以组成一组模糊-清晰样本对。为了保证数据集的质量,本文经过仔细挑选,舍弃那些模糊程度不足、高相似和无明显对象的样本对,最后得到7 728组高质量样本,再将这些样本对分成5 710对训练样本和2 018对测试样本。

2 多阶段编码-解码结构

为了实现视频图像去运动模糊的目标,本文采用多帧图像序列去模糊的方案。实验表明三帧连续图像序列基本能达到去模糊的质量要求,而且处理时间在接受范围内。假设当前帧为 B_0 ,本文考虑到当前帧的前后一帧,分别表示为 B_{-1} 和 B_1 ,组合成一个 (B_{-1}, B_0, B_1) 三连帧图像序列。在最小二乘生成对抗网络(Least square GAN, LSGAN)^[14]的基础上,本文设计了一个双向时域特征流盲去运动模糊生成

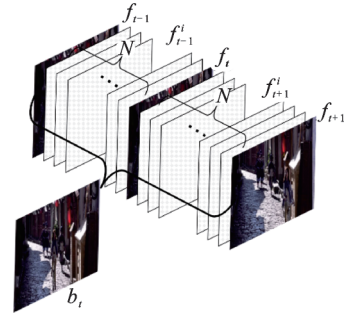


图2 双向光流插帧平均法创建模糊图像
Fig. 2 Blurred image creation by bidirectional optical flow interpolation frame average method

网络(图1)。该网络训练一个端到端的全卷积生成器 G 。 G 采用多阶段的编码-解码结构,接收该三连帧模糊图像序列作为网络的输入。但这三连帧模糊图像在网络中并不是在输入端以通道叠合方式同时输入,而是先后输入到网络中,所以本文网络整体上可以细分为两个去模糊阶段。前一阶段的输出可作为后一阶段的输入,所以去模糊的结果也是结合了三连帧的全部内容信息。考虑到相邻帧之间的时域特征信息,在第一阶段先提取当前帧的内部特征信息,并在第二阶段通过双向时域特征传输通道将其分别转移到前一帧和后一帧的编码器中,最后前后帧编码会被混合解码,生成的清晰图像 E_0 将作为 B_0 最终的去模糊结果。

2.1 第一阶段去模糊结构

编码器-解码器结构^[15]是图像处理网络中最常用的网络结构,它在图像超分辨、图像去噪等方面均得到广泛应用。这种结构通常是由一个编码器和一个解码器组成对称形式。编码器用于提取输入图像的特征并在高维空间中表达。为了增大网络的感受野,编码器一般会逐级将图像下采样,与此同时扩大特征层的通道数以尽量减少特征信息的损失。而解码器是编码器的逆过程,通过上采样操作将编码器在高维空间中的编码特征逐级恢复到与原始输入大小相同的图像。

在第一阶段,在参考Tao等人的SNR-DeblurNet^[16]后,本文设计的编码器-解码器结构如图3所示。本文在编码阶段先使用32个 5×5 大小的卷积核从原始输入图像中提取32通道特征层,然后使用两个步长为2的卷积操作将输入图像下采样到原始大小的1/4,特征层通道数也扩大到初始特征层的4倍。He等人^[17]研究表明残差块能在增加网络深度的同时提高网络的收敛速度,所以本文在每个卷积层后添加3个残差块,每个残差块由两个卷积层所构成并形成残差连接。解码器采用与编码器对称的网络结构,本文使用两个步长为2的转置卷积层将高维空间特征上采样。同样在每个转置卷积层前也放置3个相同形式的残差块。所有的卷积层和转置卷积层均采用 5×5 大小卷积核。下采样编码操作提取了语义信息但也损失重要的空间信息,为了补充空间信息,本文在编码器和解码器相对层间添加跳连接^[18],使得网络学习输入图像与目标图像的残差信息,如此可以大大提高网络的收敛速度。

为了充分利用当前帧的隐含特征信息和相邻帧的时间维度信息联系,本文在第一阶段去模糊网络中设立1个特征提取器,提取当前帧在编码器的隐含特征。因为提取的特征分别传递给前一帧和后一帧的编码器,所以本文将这个特征传递通道称为“双向时域特征流”。双向时域特征流和第一阶段初步去模糊图像 P_0 均传递到第二阶段作为前后帧编码器的公共输入。

2.2 第二阶段去模糊结构

在第一阶段后,本文得到一初步的去模糊结果 P_0 。在第二阶段,本文将 P_0 分别与前一帧和后一帧以通道叠合方式作为第二阶段网络的输入端的输入。第二阶段去模糊结构如图4所示,其主体结构与

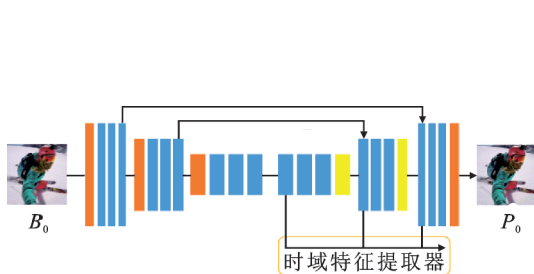


图3 第一阶段去模糊网络结构图

Fig.3 Structure of the first-stage deblurring network

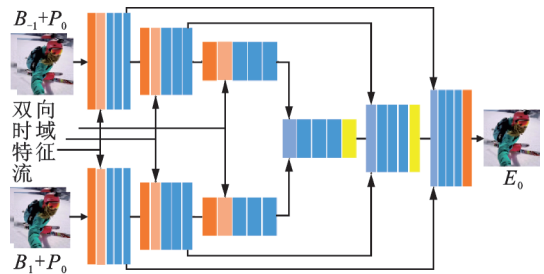


图4 第二阶段去模糊网络结构图

Fig.4 Structure of the second-stage deblurring network

第一阶段相同,但是其编码器由两子编码器组成,这两子编码器分别对输入端 $(B_{-1}+P_0)$ 和 (B_1+P_0) 并行编码,同时吸收来自第一阶段的时域特征信息。而后两高维编码信息由一个共同解码器联合解码。为了吸收时域特征流信息,本文采用 1×1 大小卷积核对时域特征与当前子编码器前向输出特征叠加后的混合特征进行卷积,由此两子编码器分别建立不同时域方向当前帧与相邻帧的特征映射。因此,虽然两子编码器具有相同的结构和状态,但不共享参数。

两子编码器的编码并没有独立解码,而是采用混合解码的方式。本文将两编码通道叠合输入到公有解码器。为了融合两者之间的信息,本文采用 1×1 大小卷积核学习到两子编码特征的相应融合权值。而后与第一阶段解码器类似,两个步长为2的转置卷积层将高维空间混合编码特征上采样恢复出与 B_0 相同大小的去模糊图像。同样,为了加速网络收敛,并遵循由粗到精的优化策略,本文在初步估计结果 P_0 与输出层之间引入了全局跳连接,以便网络学习到初始估计 P_0 的残差校正 I_R ,而不是从头开始生成清晰的图像。因此,整个去模糊网络相对于当前帧 B_0 的去模糊结果 E_0 表示为 $E_0=I_R+P_0$ 。

2.3 损失函数设定

本文的去模糊网络以LSGAN^[14]为基本框架。LSGAN的主要思想就是在判别器 D 中使用更加平滑和非饱和梯度的损失函数。为此,Mao等人选用最小二乘损失也就是 L_2 损失惩罚与真实分布偏远的的数据。在优化过程中,促使生成器 G 逐渐拟合真实数据 P_{data} 。所以LSGAN训练的鉴别器损失和生成器损失分别如式(4,5)所示。

$$L_{adv-d} = \frac{1}{2} E_{x \sim p_{data}} [D(x) - 1]^2 + \frac{1}{2} E_{\hat{x} \sim p_G} [D(\hat{x})]^2 \quad (4)$$

$$L_{adv-g} = \frac{1}{2} E_{\hat{x} \sim p_G} [D(\hat{x}) - 1]^2 \quad (5)$$

式中: P_{data} 表示真实的数据分布; P_G 表示生成数据分布; x 表示真实样本; \hat{x} 表示生成器生成样本。除了基本的对抗生成损失之外,为了使生成的图像更接近于真实的图像,减少内容上的差异,本文又在生成损失基础上添加像素级损失 L_{pixel} 和特征级损失 L_{feat} 作为内容损失。

像素级损失一般使用MSE或MAE衡量,本文选择MSE,所以本文的 L_{pixel} 定义为每阶段去模糊结果与目标图像的Euclidean距离之和

$$L_{pixel} = \frac{1}{W_i H_i C_i} \|I_E^{(i)} - I_S\|_2^2 \quad (6)$$

式中: $I_E^{(i)}$ 表示第 i 个去模糊阶段的复原图像; I_S 表示当前帧 B_0 的目标图像 S_0 。

近些年感知损失^[13](Perception loss)在图像复原领域研究得到广泛应用,目的是在特征层级上缩小与目标图像的差异。本文定义 L_{feat} 为去模糊图像 E_0 与目标图像 S_0 在VGG19网络^[19]中relu3_3层上特征的 L_1 距离。所以 L_{feat} 可以表示为

$$L_{feat} = \frac{1}{W_j H_j C_j} \|\phi_j(y) - \phi_j(\tilde{y})\|_1 \quad (7)$$

式中 ϕ_j 表示输入图像在VGG19第 j 层的特征图。综上所述,本文的总生成损失可以表示为

$$L_g = L_{adv-g} + \lambda_1 L_{pixel} + \lambda_2 L_{feat} \quad (8)$$

式中 λ_i 分别表示像素级损失和特征损失相对对抗生成损失的权值。在本文的实验中,设置 $\lambda_1=10, \lambda_2=5$ 基本能满足训练要求。

2.4 训练细节

对抗网络训练是一个生成器和鉴别器相互竞争相互协作的过程。本文在一块NVIDIA Tian V100显卡上实施所有训练和测试实验。实验设置batch_size为1,生成器除了最后卷积层使用Tanh激活外,

其余卷积和转置卷积操作激活函数均选择 ReLU, 使用 Instance Normalization 正则化^[20], 而鉴别器的卷积层激活函数均选择 Leaky_ReLU, 使用 Layer Normalization 正则化^[21]。而因为训练样本大小为 $1\ 280$ 像素 \times 720 像素, 所以从训练样本中随机剪切出 256 像素 \times 256 像素大小的图像块作为网络的输入。本文选择 Adam^[22] 作为优化器, 初始的学习率为 $1e^{-5}$, 并在训练的过程中呈指数级衰减, 衰减率 $\alpha = 0.3$ 。在训练阶段, 生成器与鉴别器交替更新, 在其中一个更新时保持另一个参数不变。本文设置最大的 epoch 数为 300。整个训练过程大约需要 3~5 d 时间。

3 实验评估

为了从定性和定量评价本文算法, 从统计度量和视觉质量两个方面分别将本文算法与当前先进的去模糊算法进行比较。

3.1 统计度量

本文模型的目的主要是提高去模糊的视觉质量, 所以不太关注传统的定性评价指标。但是为了进一步说明本文模型的优势, 本文将模型与当前一些优秀的去模糊算法进行对比。基于本文所创建的测试数据集, 用传统的图像质量评价标准: 峰值信噪比 (Peak signal to noise ratio, PSNR) 和结构相似性 (Structural similarity, SSIM) 评估并计算单个测试样本平均处理时间。本文选择 Multi-scale CNN^[7], RDN^[23] 和 DeblurGAN^[7] 作为比较对象, 因为它们的网络具有独特的结构并且效果优异。为了更体现公平性, 其中 RDN 的序列长度设置为 3, 而 Multi-scale CNN 的尺度 K 值也设为 3。所有测试均在 2 018 个测试样本基础上进行, 测试样本大小为 256 像素 \times 256 像素。实验结果如表 1 所示。

表 1 统计度量评价结果

Tab. 1 Results of statistical evaluation

算法	Multi-scale CNN	RDN	DeblurGAN	本文模型
PSNR	28.032	27.384	28.338	28.695
SSIM	0.886	0.878	0.894	0.897
时间/s	4.405	0.102	0.033	0.054

显然, 与优秀的去模糊算法相比, 本文提出的模型在 SSIM 和 PSNR 方面均表现优异。一方面, 本文模型的 PSNR 不仅远超过 RDN 算法 1.31 dB 和 Multi-scale CNN 算法 0.66 dB, 而且在 DeblurGAN 的基础上提高 0.36 dB。并且本文的 SSIM 也高于其他竞争对手。更重要的是, 在处理速度方面, Multi-scale CNN 几乎 80 倍率慢于本文模型。但也注意到因为本文模型的多阶段编码-解码去模糊结构, 本文模型接近 1.6 倍率慢于 DeblurGAN。但综合而言, 本文模型可以更好地实现性能与速度均衡。

3.2 视觉质量

高 PSNR 和 SSIM 并不等同于好的去模糊效果。本文追求的目标则是更好的视觉质量。仍然用 Multi-scale CNN 和 DeblurGAN 与本文模型作比较。本文从测试结果中随机选择多个去模糊的图像, 直观的视觉对比如图 5 所示。由图 5 可以看出, 与其他算法相比, 本文模型的去模糊结果展现出更好的视觉质量, 去模糊后图像包含更好的纹理细节和较少的噪声, 这表明在模糊核形式未知条件下, 本文模型不需要预先估计运动模糊图像的模糊核便能盲处理各种情况下造成的运动模糊问题。此外, 对于图像中那些不模糊的区域, 本文的去模糊模型仍然能较好地保持其原始清晰度, 而不会人为引入过多伪影。

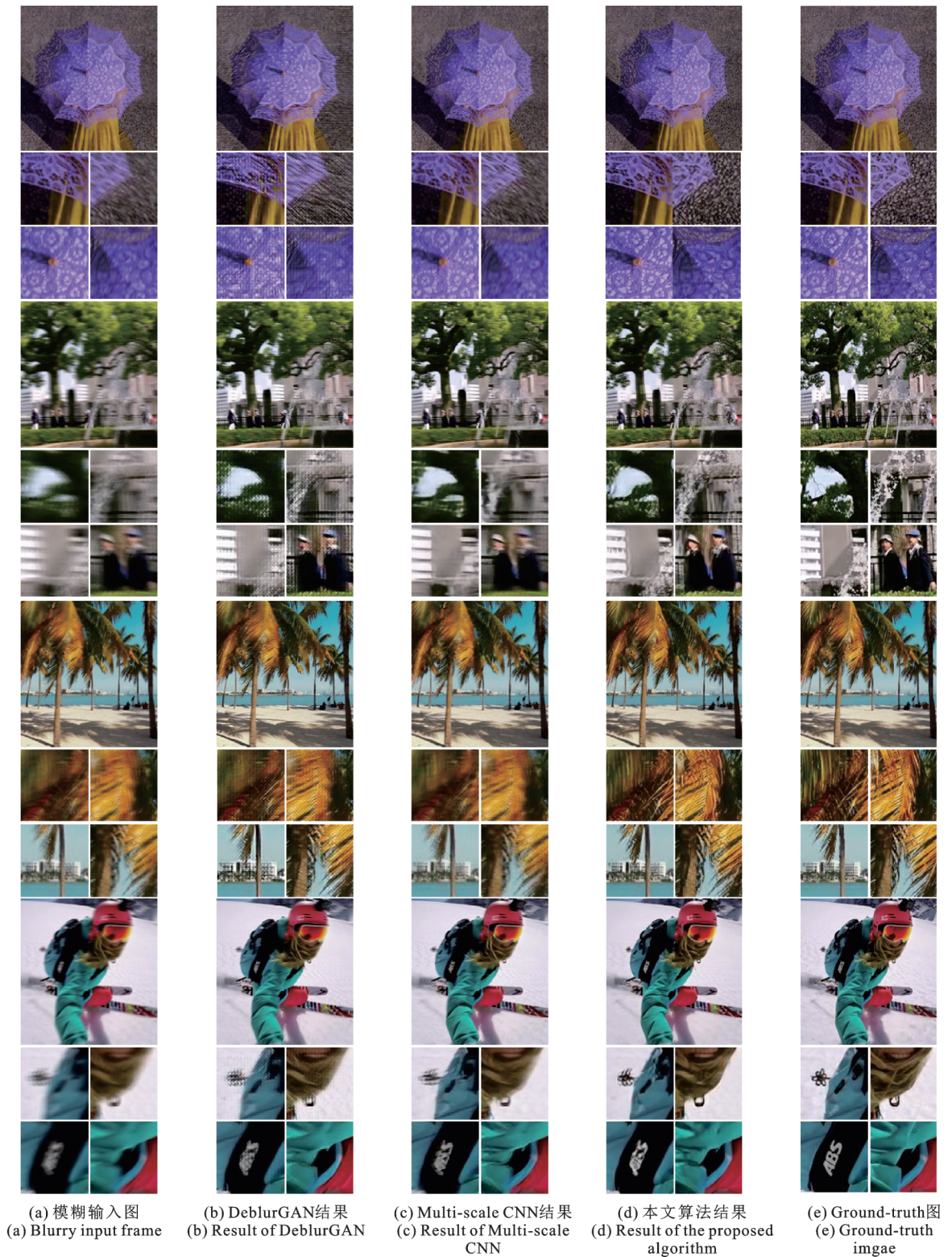


图5 视觉质量评价对比结果

Fig.5 Results of visual quality comparison

4 结束语

本文先采用双向光流插帧合成法创建一个高质量的运动模糊数据集,该数据集不仅适用于单帧图像去模糊研究而且也适用于视频图像去模糊研究。进一步,对于视频图像去模糊问题,本文提出了一种双向时域特征流盲运动去模糊生成网络,该网络取三连帧图像序列作为输入,采用多阶段的编码-解码结构。在其中为了充分利用相邻帧时域信息,本文提出一种并行编码和混合解码的解决方案。实验测试结果表明,本文提出的模型在统计度量评价和图像质量评价方面均优于当前一些优秀的算法。但本文方法也存在处理速度较慢、对硬件性能要求高等问题,距离实际应用还需要进一步优化。

参考文献:

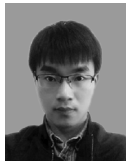
- [1] Kupyn O, Budzan V, Mykhailych M, et al. DeblurGAN: Blind motion deblurring using conditional adversarial networks[EB/OL]. <https://arxiv.org/abs/1711.07064>, 2017.
- [2] Proakis J G, Manolakis D G. 数字信号处理——原理、算法与应用[M]. 4版. 北京:电子工业出版社, 2013: 863-871.
Proakis J G, Manolakis D G. Digital signal processing—Principles, algorithms and applications[M]. 4th ed. Beijing: Electronic Industry Press, 2013: 863-871.
- [3] Lam E Y, Goodman J W. Iterative statistical approach to blind image deconvolution[J]. Journal of the Optical Society of America: A Optics Image Science & Vision, 2000, 17(7):1177.
- [4] Ma L, Zhang R, Qu Z, et al. Blind image deconvolution using sparse and redundant representation[J]. Optik-International Journal for Light and Electron Optics, 2014, 125(23): 6942-6945.
- [5] Xu L, Ren J S J, Liu C, et al. Deep convolutional neural network for image deconvolution[C]// NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2014: 1790-1798.
- [6] Sun J, Cao W, Xu Z, et al. Learning a convolutional neural network for non-uniform motion blur removal[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2015: 769-777.
- [7] Nah S, Kim T H, Lee K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, Hawaii, USA: IEEE, 2017, 2: 3.
- [8] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs[EB/OL]. <https://arxiv.org/abs/1704.00028v3>, 2017.
- [9] Kim T H, Lee K M. Generalized video deblurring for dynamic scenes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2015: 5426-5434.
- [10] Hyun Kim T, Ahn B, Mu Lee K. Dynamic scene deblurring[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]:IEEE, 2013: 3160-3167.
- [11] Su S, Delbracio M, Wang J, et al. Deep video deblurring for hand-held cameras[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, Hawaii, USA: IEEE, 2017, 2: 6.
- [12] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[EB/OL]. <https://arxiv.org/abs/1406.2661>, 2014.
- [13] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution[C]//European Conference on Computer Vision. Cham, Switzerland: Springer, 2016: 694-711.
- [14] Mao X, Li Q, Xie H, et al. Least squares generative adversarial networks[C]//Computer Vision (ICCV), 2017 IEEE International Conference on. [S.l.]: IEEE, 2017: 2813-2821.
- [15] Mao X, Shen C, Yang Y B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections[EB/OL]. <https://arxiv.org/abs/1603.09056>, 2016.
- [16] Tao X, Gao H, Wang Y, et al. Scale-recurrent network for deep image deblurring[EB/OL]. <https://arxiv.org/abs/1802.01770>, 2018.
- [17] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks[C]//European Conference on Computer Vision. Cham, Switzerland: Springer, 2016: 630-645.

- [18] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2016: 770-778.
- [19] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. <https://arxiv.org/abs/1409.1556>, 2014.
- [20] Ulyanov D, Vedaldi A, Lempitsky V S. Instance normalization: The missing ingredient for fast stylization[EB/OL]. <https://arxiv.org/abs/1607.08022>, 2016.
- [21] Ba J L, Kiros J R, Hinton G E. Layer normalization[EB/OL]. <https://arxiv.org/abs/1607.06450>, 2016.
- [22] Kingma D P, Ba J. Adam: A method for stochastic optimization[EB/OL]. <https://arxiv.org/abs/1412.6980>, 2014.
- [23] Wieschollek P, Hirsch M, Schölkopf B, et al. Learning blind motion deblurring[C]// 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 231-240

作者简介:



赵跃进(1958-),男,教授,博士生导师,研究方向:基于MEMS的红外成像技术、太赫兹成像技术、空间光学技术等, E-mail: yj-zhao@bit.edu.cn。



刘文龙(1979-),男,硕士研究生,研究方向:基于深度学习的图像处理技术。



刘明(1979-),男,副教授,研究方向:智能图像处理、物体检测与识别、光电仪器。



董立泉(1979-),男,副教授,研究方向:光电仪器和检测技术、波前编码技术、视频和图像处理、红外夜视技术。



惠梅(1963-),女,副教授,研究方向:相移干涉测量、分段镜相位测试、红外MEMS系统成像和机械设计。

(编辑:张黄群)