

基于卷积网络的边缘保持滤波方法

石晓红^{1,2,3} 黄钦开⁴ 苗佳欣⁵ 苏 卓^{4,5}

(广州大学数学与信息科学学院 广州 510006)¹ (广州大学计算科技研究院 广州 510006)²
(广东省数学教育软件工程技术研究中心 广州 510006)³
(中山大学数据科学与计算机学院 广州 510006)⁴
(中山大学国家数字家庭工程技术研究中心 广州 510006)⁵

摘 要 边缘保持滤波是计算机视觉、图像处理领域的重要基础理论研究,作为图像预处理操作对后续的处理结果有着重要影响。区别于传统滤波方法,边缘保持滤波方法不仅注重图像的平滑处理,还注重保持边缘细节。卷积神经网络在很多领域得到了应用,并取得显著的成果。本文将卷积神经网络引入边缘保持滤波,利用卷积神经网络的良好扩展性和灵活性来构建深度卷积神经网络模型(Deep Convolutional Neural Network,DCNN),通过 3 种类型的网络堆叠层,采用反向传播迭代更新网络参数,训练残差图像,实现基于 DCNN 的边缘保持滤波方法;还构建了基于梯度域的卷积神经网络模型(Gradient CNN,GCNN),对彩色图像的梯度信息进行学习,通过三层卷积对梯度图进行边缘保持平滑操作,得到边缘保持平滑梯度图,进而利用输入图像引导平滑梯度图进行彩色重建,得到彩色滤波图像。最后通过实验与常见的边缘保持滤波方法进行主观和客观评价对比。DCNN 不仅在视觉上达到了其他滤波的效果,同时也在处理时间上也存在较大优势,表明 DCNN 可以通过大量的数据训练有效地拟合出多种边缘保持滤波算法。与其他边缘保持滤波结果相比,GCNN 在视觉上可以保持颜色风格与输入图像整体一致,而且图像相似度评价指标也更好,表明 GCNN 解决了部分滤波处理出现颜色偏差、梯度反转等问题,而且提高了处理效率。

关键词 滤波,边缘保持,平滑操作,卷积神经网络

中图法分类号 TP391 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.09.042

Edge-preserving Filtering Method Based on Convolutional Neural Networks

SHI Xiao-hong^{1,2,3} HUANG Qin-kai⁴ MIAO Jia-xin⁵ SU Zhuo^{4,5}

(School of Mathematics and Information Science,Guangzhou University,Guangzhou 510006,China)¹
(Institute of Computing Science and Technology,Guangzhou University,Guangzhou 510006,China)²
(Guangdong Provincial Engineering and Technology Research Center for Mathematical Education Software,Guangzhou 510006,China)³
(School of Data and Computer Science,Sun Yat-sen University,Guangzhou 510006,China)⁴
(National Engineering Research Center of Digital Life,Sun Yat-sen University,Guangzhou 510006,China)⁵

Abstract Edge-preserving filtering is a significant basic theory research in the fields of computer vision and image processing. As subsequent operation of pre-processing,edge-preserving filtering has great influence on final results of image processing. Different with traditional filtering,edge-preserving filtering focuses not only on smooth,but also on image edge details. Convolutional neural networks (CNNs) have been applied into a variety of research fields with great success. In this paper,CNN was introduced into edge-preserving filtering. Taking advantages of CNN's excellent extensibility and flexibility,this paper constructed a deep convolutional neural network (DCNN). With three types of cascading network layers,DCNN iteratively updates its parameters by back propagation,produces a residual image and realizes a DCNN-based edge-preserving filtering. Besides,a gradient CNN model (GCNN) was constructed. The gradient of color images is learnt,edge-preserving smoothing operation is conducted for gradient images by three layers of convolution, and edge-preserving filtering gradient images are obtained. Subsequently,the input image is used to guide the filtering gradient image for reconstructing and obtaining color filtering image. Finally,experiments were made to evaluate the

到稿日期:2018-08-30 返修日期:2018-11-19 本文受国家自然科学基金青年基金项目(61502541),2016 年贵州省科技平台及人才团队专项资金项目(黔科合平台人才[2016]5609),2016 年贵州省省级重点支持学科“计算机应用技术”(黔学位合字 ZDXK[2016]20 号),广州大学研究生创新能力培养资助计划(2018GDJC-D03)资助。

石晓红(1978—),女,硕士,主要研究方向为图像处理;黄钦开(1994—),男,硕士生,主要研究方向为图像处理;苗佳欣(1994—),女,硕士生,主要研究方向为图像处理;苏 卓(1985—),男,博士,CCF 会员,主要研究方向为计算机视觉、图形图像处理,E-mail:suzhuo3@mail.sysu.edu.cn (通信作者)。

proposed methods and the proposed methods were compared with popular edge-preserving filtering methods subjectively and objectively. DCNN not only achieves the same visual effects as other methods, but also has big advantages in processing time, which demonstrates that DCNN can effectively and efficiently imitate various filtering methods through training on large amount of data. For GCNN, in terms of visual effects, its output conforms to the input in the color style globally. In terms of image similarity evaluation, it also outperforms other methods. This verifies that GCNN can address the problems of color shift and gradient inversion, as well as improve the filtering efficiency.

Keywords Filtering, Edge-preserving, Smoothing, Convolutional neural networks

1 引言

边缘保持滤波常用于计算机视觉、图形图像处理的预处理操作,其结果的好坏直接影响了许多后续操作。传统的图像平滑算法侧重于平滑效果,导致图像经过滤波操作后边缘细节丢失。为了解决边缘细节丢失的问题,很多边缘保持的滤波方法被提出,并被广泛应用于计算机动画、数码摄影等领域,但这些方法都需要针对特定的图片人工设定参数,以获得较好的处理效果。目前边缘保持滤波模型主要包括基于均值滤波和基于最优化滤波两种。基于均值滤波的边缘保持滤波算法输出的边缘保持图像的各个位置像素值取决于输入原始图像对应位置的像素与其邻域内其他像素的密切程度,而这种密切程度往往取决于中心点像素和邻域内其他像素在空间域和亮度域的关系。基于最优化滤波以加权最小二乘(WLS)滤波^[1]为代表,经过滤波处理后,结果图像尽量平滑,在边缘部分与原始图像尽量保持相似。除了上述两大类滤波算法外,还有其他边缘保持滤波算法,包括基于拉普拉斯金字塔改良得到的具有边缘感知效果的局部拉普拉斯滤波^[2]、基于域变换思想的加速策略^[3]等,这些方法也在图像滤波预处理领域得到广泛的应用。

随着深度学习理论的发展,深度卷积神经网络在很多领域得到了应用,并取得显著成果。本文试图将深度卷积神经网络引入图像边缘保持滤波处理,利用卷积神经网络良好的扩展性和灵活性,通过学习、分析已有的边缘保持滤波算法框架,结合深度学习,提出了基于卷积神经网络的边缘保持滤波方法,以改善边缘保持滤波的效果,提高处理效率。

本文的主要工作和贡献包括:

- 1)构建了一种深度卷积神经网络模型(DCNN),提出了基于深度卷积神经网络模型(DCNN)的边缘保持滤波方法,实现高效拟合多种边缘保持滤波。
 - 2)提出基于梯度域卷积神经网络(GCNN)的图像边缘保持滤波方法,通过使卷积神经网络在梯度域上学习图像边缘保持滤波的处理过程,以及利用输入的彩色原图引导学习梯度域向像素(彩色)域的彩色图恢复过程,来达到相似度很高的滤波模拟效果,解决了部分滤波处理出现颜色偏差、梯度反转等问题,提高了处理效率。
 - 3)通过实验,将以上两种方法和主流滤波方法进行对比,从主观和客观评价标准的角度分析和验证了以上两种方法的有效性。
- 本文第 2 节简要回顾了边缘保持滤波方法的发展及卷积神经网络在图像领域的应用;第 3 节和第 4 节是本文的主要工作,分别阐述了基于深度卷积神经网络模型(DCNN)的边

缘保持滤波方法及基于梯度域卷积神经网络(GCNN)的边缘保持滤波方法;第 5 节对本文提出的方法进行了实验分析,并与现有主流滤波算法进行了比较;最后总结全文。

2 相关工作

研究人员在边缘保持的图像平滑处理领域做了大量的工作,提出了众多边缘保持滤波方法。

2.1 边缘保持滤波

以高斯核函数为基础的双边滤波模型^[4-6]是均值滤波方法的一种,也是最经典的边缘保持滤波之一,目前已经被广泛应用于计算机图像和计算机视觉领域,如图像去噪^[7]、HDR 压缩^[8]、多尺度细节分解^[9]和图像抽象^[10]等。与高斯滤波类似,双边滤波模型的每个位置对应的输出像素值由输入图像对应位置像素的邻域内像素值的加权和所决定。它们的不同之处在于:高斯滤波只使用了以像素间欧氏距离为基础的空间域高斯核函数来计算邻域内各个位置的像素权重,并没有考虑到邻域内各个像素值之间的数值关系;而双边滤波结合使用了以像素空间位置为基础的空间域高斯核函数和以像素亮度差为基础的亮度域高斯核函数来形成邻域内各个像素的权值信息。双边滤波模型由于兼顾了邻域各个像素的位置信息和亮度像素,从而实现了图像边缘感知的非线性自适应平滑操作。

虽然双边滤波在边缘保持问题上取得了比较好的效果,但是这种滤波方式也有一定的局限性和问题。首先这种滤波方式会产生梯度反转的伪影问题;同时双边滤波还有效率上的问题,直接实现这种算法的时间复杂度高,针对这一问题,Porikli^[7]和 Yang 等^[8]均提出了可以在常数时间内实现双边滤波的算法。

为了改进双边滤波时间复杂度高、产生梯度反转的伪影等问题,He 等提出了引导图像滤波^[13],与双边滤波不同的是,引导图像滤波在于引用引导图像进行滤波,参考了引导图像的结构信息,并且算法的时间复杂度大大降低。

各向异性扩散^[14]也被称为 P-M 扩散,是基于均值的滤波方式之一,广泛应用于图像处理和计算机视觉领域中的图像细节特征保持和去噪问题。各向异性的思想在于,由于图像边缘往往在梯度值较大的点处,如果扩散方程在梯度值比较大的区域减速扩散,在梯度值较小的区域加速扩散,那么就可以在去噪的同时保护图像的边缘细节。基于这种思想,Perona 和 Malik 提出了两种扩散系数方程,也就是有名的 P-M 方程。

基于最优化的滤波模型有加权最小二乘(WLS)滤波^[1]、L0 平滑^[5]、相对全变分(RTV)^[11]等,这类方法是针对边缘保

持的特性提出目标函数和约束条件,然后通过最优化的方式对方程进行求解得到输出图像,求解的方式有梯度下降等。

2.2 卷积神经网络在图像领域的应用

深度学习在图像处理领域中较早引起人们关注的是图像识别方向,无论是手写体识别^[16],还是基于深度卷积神经网络的 ImageNet 识别^[17],都表明了深度学习在图像识别上的优势。当然,目前深度学习不仅在图像识别领域取得了成果,在图像视觉领域也取得了突破性的成果。Dong 等^[12]基于 CNN 提出的深度卷积神经网络在图像超分辨率问题上有较好的解决能力;Xie 等^[13]提出的深度神经网络模型在图像去噪和去水印问题上均取得较好效果;Liu 等^[14]提出的基于 CNN 和 RNN 的混合神经网络被应用于一系列的低层图像视觉问题。

3 基于深度卷积神经网络模型(DCNN)的边缘保持滤波

卷积神经网络(CNN)作为神经网络的模型之一,在许多图像处理问题上都表现出了强大的能力。相对于其他深度神经网络,CNN 在深层网络方面表现出良好的扩展性和灵活性,而且更加适用于使用 GPU 进行并行编程,提高运行速度。因此,本文利用卷积神经网络(CNN)良好的扩展性和灵活性,构建了深度卷积神经网络(DCNN),提出了基于 DCNN 的边缘保持的滤波方法,可以有效拟合多种滤波效果,过程如图 1 所示。

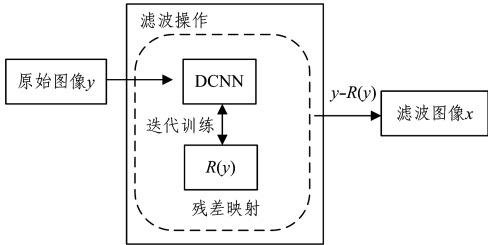


图 1 基于 DCNN 的边缘保持滤波

Fig.1 DCNN-based edge-preserving filtering

残差学习的提出是为了解决神经网络由于学习的深度而导致的退化问题。残差学习的思想在于利用多层网络拟合一个残差映射,而不是利用每组少数的几个层的网络直接拟合出所期望的实际映射关系。用 $H(x)$ 表示期望得到的实际映射,使用堆叠的非线性多层网络拟合另一个映射关系 $F(x)$: $H(x)-x$,那么实际的映射关系即可表示为 $F(x)+x$ 。与一般的残差网络不同,DCNN 并没有使用多个残差单元,而只使用了一个残差单元来映射出残差图像。

批规范化是深度学习中的重要处理方式之一,广泛的应用已经证明了它的有效性和重要性。在深层网络训练时,由于模型参数不断修改,各层的输入数据的概率分布也不断变化,因此必须使用较小的学习速率及较好的权重初值,这将导致网络的训练速度很慢,同时容易出现激活函数饱和而导致训练困难的现象。批规范化就是通过一定的规范化手段,把每层神经网络任意神经元的输入数据分布强行拉回到均值为 0、方差为 1 的标准正态分布,而不是简单的正态分布,从而使

输入数据的小变化将导致损失函数出现较大的变化,避免了梯度消失问题,也大大加快了网络的训练速度。

标注输入的彩色图片为 y ,边缘保持滤波操作为 $L(y)$, $L(y)$ 可以是线性的或非线性的、全局的或局部的操作,设其输出为 x ;同时标记 y 与 $L(y)$ 的差值为 $r=y-x$ 。DCNN 网络的输入是原始图片 y ,其输出是残差 r ,而不是边缘保持图片 x 。在 DCNN 中,我们利用残差学习的思想,让输入通过训练网络得到一个残差映射模型 $R(y)=r$,然后根据 $R(y)$ 和 y ,就可以得到边缘保持图片 $x=y-R(y)$ 。

在深度神经网络的训练过程中,使用损失函数来估计模型的预测值与真实值之间的不一致程度,同时结合迭代策略对网络的参数进行更新。在 DCNN 网络训练的过程中,使用的损失函数是平方损失函数,表达式为:

$$C(w)=\frac{1}{2N}\sum_{i=1}^N\|R(y_i;w)-(y_i-x_i)\|^2$$
 (1)

其中, w 表示神经网络的参数; $\{(y_i, x_i)\}_{i=1}^N$ 表示 N 对数据,其中 y_i 表示第 i 个输入图片, x_i 表示对应的第 i 个边缘保持图片; $R(y_i;w)$ 表示在参数 w 下的网络得到的残差映射函数针对输入 y_i 得到的结果。DCNN 网络就是希望通过大量数据,迭代更新参数 w ,最终得到残差映射函数 $R(y)$ 。

图 2 为本方法构建的深度卷积神经网络(DCNN)结构。根据 DCNN 结构图,首先给定网络的深度为 D ,网络中有 3 种类型的堆叠层。

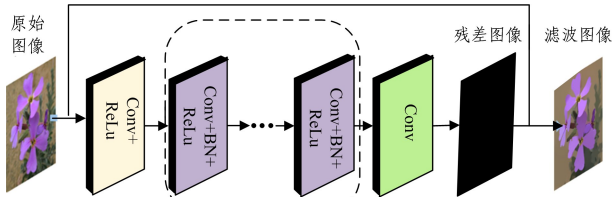


图 2 深度卷积神经网络结构

Fig.2 Structure of deep convolutional neural network

Conv+ReLU 为网络的第一个堆叠层,其主要针对每一个彩色输入图(3 通道)生成 64 个特征图。其中 Conv 使用 64 个 $3\times3\times3$ 的滤波器,生成 64 个输出特征图,然后使用非线性激活函数 ReLU 对结果进行激活操作。

Conv+BN+ReLU 是网络的第 2 ~ ($D-1$) 层,每个层的 Conv 都使用 64 个 $3\times3\times64$ 的滤波器对 64 个输入特征图进行卷积,总共得到 64 个特征图;Conv 和 ReLU 之间使用 BN 层进行批规范化,调整数据的分布;最后使用 ReLU 对结果特征图进行非线性的激活操作。

Conv 是网络的最后一层,将 64 个输入特征图作为输入,使用 3 个 $3\times3\times64$ 的滤波器,生成 3 个输出特征图,也就是残差三通道图像。

在 DCNN 的训练过程中,首先从图片数据集中选取一个原始图片块 y ,然后将其作为神经网络的输入,把对应的残差图片块 $R(y)$ 作为神经网络的输出,通过反向传播的方式更新神经网络的参数,最后通过迭代的方式逐步学习到所需的模型。在通过 DCNN 得到的模型求解出边缘保持图片时,只需要使用输入图像 y 进行前向传播得到残差图像 $R(y)$,通过 $x=y-R(y)$,便可以得到边缘保持图像 x 。

4 基于梯度域卷积神经网络(GCNN)的边缘保持滤波

本文利用卷积神经网络对边缘保持滤波的梯度信息进行学习,提出基于梯度域卷积神经网络(GCNN)的图像边缘保持滤波方法,通过模拟联合滤波结构的二重卷积神经网络结果使梯度图复原成彩色图,使得最终的图像边缘保持平滑效果,其过程如图3所示。

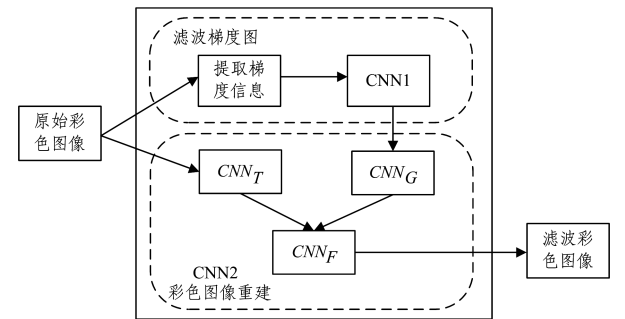


图3 基于GCNN的边缘保持滤波

Fig.3 GCNN-based edge-preserving filtering

4.1 基于梯度域的边缘特征提取

对图像滤波器进行近似模拟时,在彩色域直接进行训练存在着一些问题。

如图4所示,通过绘制图形的一维信号图可以看到,在图像彩色域上的学习结果存在一些波动,而在梯度域上训练波动则大幅度削弱。可能的原因是颜色值表示了大量信息,对于区分边缘和细节的任务来说,很多像素颜色值的绝对信息是多余的,从而对训练产生了干扰。边缘或是细节的判断取决于其与相邻像素值的差异,也就是相对信息。

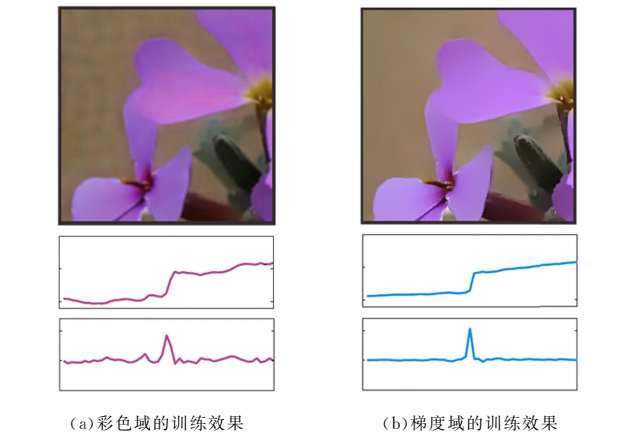


图4 梯度域和彩色域的训练效果对比

Fig.4 Training effect comparison in gradient domain and color domain

本方法采用梯度图作为网络的输入。因为梯度图本身的一阶差分已经对边缘信息进行了一定的处理和提取,所以可以摒除颜色本身对滤波结果的影响。

除此之外,因为绝大多数边缘保持滤波器对于同一幅图像旋转之后得到的滤波结果是相同的,所以对于每张彩色图,经过旋转翻转再求差分可以得到一组梯度图,全部作为输入。令这一组图片为 $I=\{I_0, I_1, I_2, \dots, I_i\}$,每组共有 D 张图片,即 $i=\{0, 1, \dots, D-1\}$ 。从而使目标变为最小化式(2):

$$\frac{1}{D} \sum_i \{ \frac{1}{2} \| \mathcal{F}_w(\partial I_i) - \partial \mathcal{L}(I_i) \|^2 \}$$
 (2)

边缘保持平滑滤波器作为平滑滤波器的一种,在大面积的区域应该保持平滑,即相邻像素值的差别总体处在一个比较低的值,也就是差分总体比较小。因此再引入一个约束,得到式(3):

$$\frac{1}{D} \sum_i \{ \frac{1}{2} \| \mathbf{F}_w(\partial I_i) - \partial \mathcal{L}(I_i) \|^2 + \lambda \phi(\mathbf{F}_w(\partial I_i)) \}$$
 (3)

其中, $\phi(z) = (z^2 + \epsilon^2)^{1/2}$, 该惩罚项接近但不等于0,是正则化权重。

CNN1网络负责训练原图的梯度图到滤波图像的梯度图的转换过程,其结构如图5所示。

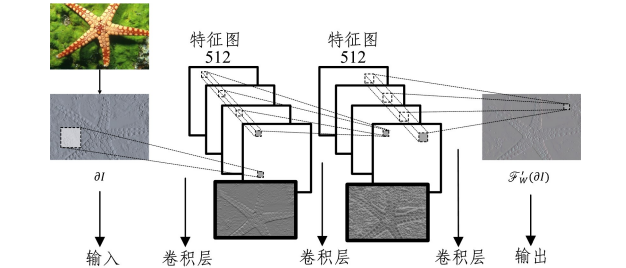


图5 CNN1的结构

Fig.5 Structure of CNN1

梯度信息 ∂I 作为输入,通过第一个核为 $16 \times 16 \times 3$ 的卷积层得到512张特征图,该步骤是对各个局部块进行重新映射,并且放在向量中。第二个核的卷积层为 $1 \times 1 \times 2$,产生像素向量的加权平均值,该操作对应的边缘保持滤波的平滑过程同样得到512张特征图。最后一个卷积层采用 $8 \times 8 \times 3$ 卷积核,从而获得对图像的重建输出。该操作通过对边缘区域相邻像素值差的控制,来提升边缘锐度,从而获得与原始图像相似的边缘保持效果。

4.2 基于引导图的彩色图重建

将图片输入网络时,彩色图已经转化为梯度图,网络训练得到的模型也用于进行输入原图梯度图和滤波结果梯度图之间的转换。由于最终的目的是得到图像的彩色图而不是梯度图,因此下一步工作是将梯度图转化成彩色图。在梯度图到彩色图的重建过程中,原图作为已知信息进行引导。

已知原图,通过输入图片的垂直和水平方向的梯度,采用最小化式(4)进行图像重建:

$$\| S - I \|^2 + \phi$$
 (4)

其中, $\phi = \beta \{ \| \partial_x S - \mathcal{F}_w'(\partial_x I) \|^2 + \| \partial_y S - \mathcal{F}_w(\partial_y I) \|^2 \}$, S 表示最终重建得到的彩色图, $\| S - I \|^2$ 是颜色置信项,用输入图片引导平滑的图像重建。 ϕ 是损失函数,利用网络模型得到的梯度结果来引导重建。 β 是平衡两项损失函数的参数,这个参数依赖滤波器的种类,可以通过简单的搜索来确定合适的 β 。

式(4)的能量方程可以通过稀疏五点拉普拉斯变成最小化的一个线性系统,例如采用Xu等^[11]的预处理共轭梯度算法(PCG)来加速该稀疏线性系统。

本方法采用学习的框架将梯度图恢复成彩色图。梯度图重建过程可以看作以原图的彩色图为基础,经过滤波的梯度图包含边缘信息,可以作为引导图,引导原图向滤波图像转化。该思想类似于联合滤波,Li等提出了一种基于卷积神经

网络的模型^[15],对目标输入和引导输入分别进行训练,然后将其合并,在特征提取网络中得到最后的输出。该方法的梯度图重建任务中,原图是目标图像,而上一个网络模型得到的梯度结果为引导图,用滤波图像作为标签训练。

引导滤波的网络结构设计,最基本的思路是将目标图和引导图拼接在一起作为网络的输入,但是实验表明这种网络效果并不理想。

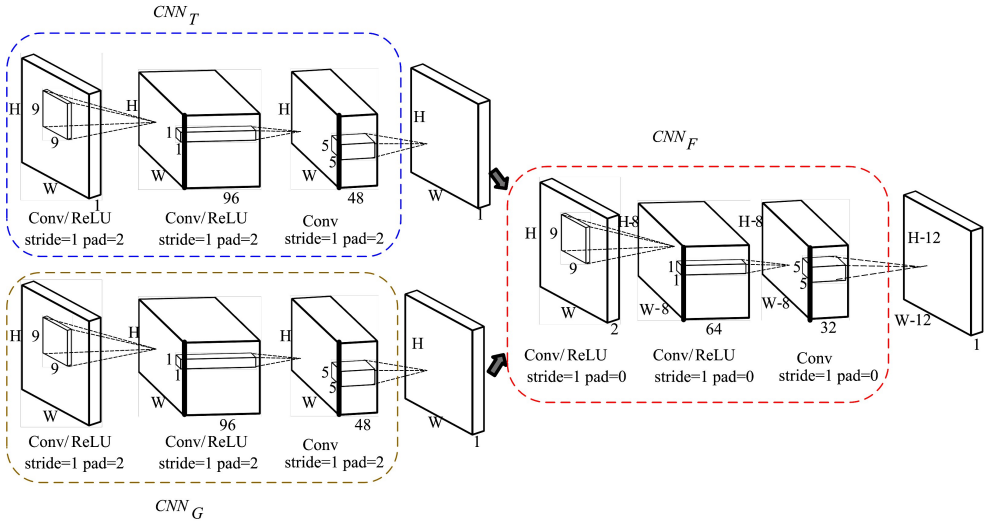


图 6 CNN2 的结构
Fig. 6 Structure of CNN2

假设目标图是 I^T ,引导图是 I^G ,目标输出是 I^{σ} ,对于 N 组训练图像样本 $\{I^T, I^G, I^{\sigma}\}_{i=1}^N$,网络的目标是最小化式(5):

$$\|I^{\sigma} - \phi(I^T, I^G)\|_2^2 \quad (5)$$

其中, ϕ 指代该网络表示的引导滤波算子。式(5)在本质上符合图像相似度判断的一个重要指数——均方根误差(MSE),因此 CNN2 的网络可以输出在相似度上可以得到近似标准滤波结果的图像。

该方法梯度图重建彩色图的过程中,目标图是彩色图原图,引导图由 3 层梯度信息组成,分别是原图水平方向梯度、CNN1 输出的水平方向梯度以及垂直方向梯度,训练的目标输出是滤波算法得到的标准滤波图。

通过 Li 等^[16]的实验观察以及之前 Dong 等的工作^[12]发现,对于高层语义任务十分有效的更深的网络结构,其对于底层图像处理任务的效果并不明显,加深网络并没有使训练效果得到明显提高。因此本文方法采用两个神经网络模型(CNN1 和 CNN2),其相对于 VGG-16^[16],GoogleNet^[17]等适用于图像分类的网络结构来说比较浅。基于梯度域的边缘保持滤波模型网络 CNN1 只有 3 个卷积层,梯度图引导彩色图重建模型网络 CNN2 同样采用 3 个卷积层,总体训练只需 6 层网络即可完成。这样的网络结构对于提高滤波算法效率更加适用,网络层数的减少意味着在测试阶段花费的时间也更短,能够有效提高算法效率。

5 实验结果分析

对于所提出的两种方法,本文进行了大量的实验,主要拟合多种边缘保持滤波算法,包括但不限于双边滤波(BLF)、基于加权最小二乘的边缘保持滤波(WLS)、L0 平滑等。当然,

彩色图重建的网络由以下 3 部分组成: CNN_T , CNN_G 和 CNN_F 。其中, CNN_T 和 CNN_G 分别用来提取目标图及引导图的特征,然后将提取的特征一起放进 CNN_F 进行训练,以得到最终结果。 CNN_T 和 CNN_G 被用作非线性特征提取器,去除局部细节信息。 CNN_F 作为非线性回归方法将目标图和引导图的特征映射到最终的结果图像。该网络称为 CNN2,其结构如图 6 所示。

本文提出的模型并不只限于这些方法,而是对几乎所有的边缘保持平滑滤波具有适应性,本文只选择部分有代表性的边缘保持滤波算法进行实验分析。

5.1 基于 DCNN 的边缘保持滤波的实验结果分析

从图 7 可以看出,当网络的深度为 5 时,DCNN 可以较好地拟合出各种边缘保持滤波算法的效果。

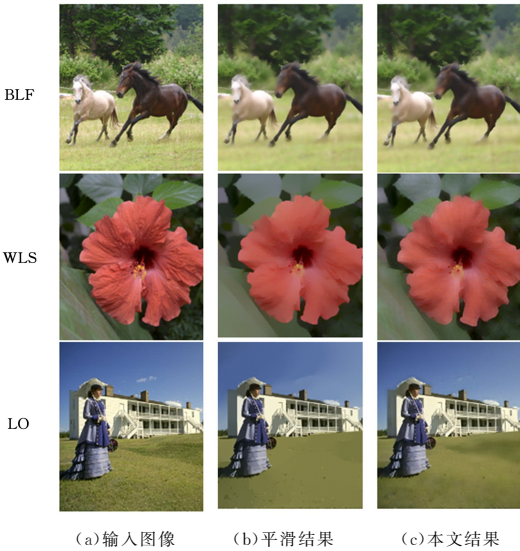


图 7 边缘保持滤波与 DCNN(深度 $D=5$)的结果对比
Fig. 7 Results comparison between popular filtering methods and DCNN (depth $D=5$)

由于 DCNN 网络的深度变小,其前向传播的速度得到了提升,表 1 是在不同分辨率图片下 DCNN 模型运行时间与各种算法的运行时间的对比。

表 1 不同分辨率的彩色图片下各模型的运行时间对比
Table 1 Time cost comparison of models under images with different resolutions

(单位:s)			
分辨率	QVGA	VGA	720P
BLF	1.27	5.03	14.82
WLS	0.43	2.08	7.11
L0	0.24	1.16	3.89
本文方法	0.32	1.36	4.04

综合图 7 和表 1 可以看出,网络深度为 5 的 DCNN 模型达到了现有的边缘保持滤波的效果,同时在运行时间,其速度与 L0 平滑算法相当,比基于加权最小二乘的边缘保持滤波(WLS)和双边滤波(BLF)要快得多。

神经网络的收敛速度往往取决于多种因素的影响,为了使 DCNN 有较快的收敛速度,在 DCNN 网络中我们使用了批规范化的技术,同时激活函数使用了 ReLu。批规范化在加网络的收敛速度上有较好的效果,图 8 是 DCNN 在模拟双边滤波和 L0 平滑滤波时各个周期的误差值(为了突出变化,每个周期的迭代次数为 200)。

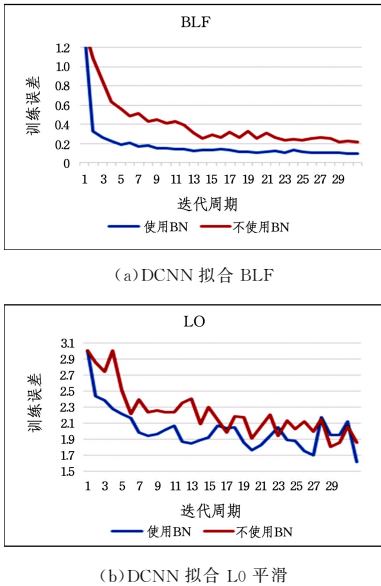


图 8 BN 对 DCNN 迭代训练误差的影响
Fig. 8 Effect of batch normalization on training error of DCNN

从图 8 中可以看出,在使用了批规范化的技术后,DCNN 在前几个周期的训练中收敛速度明显提高,而且其收敛效果也比不使用批规范化技术要好。

表 1 和图 8 表明,DCNN 可以通过大量的数据训练有效地拟合出多种边缘保持滤波的效果。本文算法使用了批规范化和残差学习等新技术,训练速度和收敛速度较快;不用设置任何图像参数,通过前向传播即可得到效果;可以有效拟合多种边缘保持滤波的效果;选择较小的网络深度时,模型的运行时间较短。当然使用神经网络也存在一定的不足,例如,不像 L0 平滑算法等边缘保持滤波算法,基于深度学习得到的 DCNN 模型难以扩展到其他领域应用,诸如结构-纹理分解、细节增强等;模型的效果依赖于数据集,而数据集难以精确、有效地制备。

5.2 基于梯度域 CNN 的边缘保持滤波的实验结果分析

评判图像的相似程度除了通过肉眼对视觉效果进行主观

评判外,还可以采用几种常见的经典图像相似度评判标准,包括均方误差(MSE)、峰值信噪比(PSNR)、结构相似度指数(SSIM)。

本节将本文方法与 Xu 的深度边缘保持滤波方法^[18]进行对比。实验中采用同样的训练数据量,用 20 万张训练集图片和 1 万张验证集图片,在 CNN1 中进行梯度图的训练,全部训练迭代 40 次。在梯度图重建彩色图的步骤中,对于 Xu 的深度边缘保持滤波方法依旧沿用其优化的方法操作,而在本文方法中,用 8 万 5 千张训练集图片和 5 千张验证数据集图片进行 CNN2 的训练,总共迭代 100 次。

在测试环节中,使用“柯达真彩数据集”^[25]的 24 张图片对 5 种图像滤波方法进行测试,包括滚动引导滤波(RGF)^[19]、L0 平滑(L0)^[5]、相对全变分(RTV)^[11]、双边滤波(BLF)^[4]、加权最小二乘滤波(WLS)^[1]。分别用 Xu 的方法^[11]和本文方法得到的模拟结果与原始滤波算法得到的标准结果进行对比,计算平均的均方误差(MSE)、峰值信噪比(PSNR)和结构相似度指数(SSIM),结果如表 2 所列。

表 2 滤波方法模拟结果的相似指数对比
Table 2 Similarity evaluation comparison of filtering methods

方法对比	评价指标	RGF ^[19]	L0 ^[5]	RTV ^[11]	BLF ^[4]	WLS ^[1]
深度边缘保持滤波 ^[11]	MSE	0.003654	0.00463	0.001258	0.001029	0.001207
	PSNR	73.28976	72.13671	77.67794	78.70042	77.80019
	SSIM	0.937727	0.927152	0.967844	0.977528	0.968143
本文方法	MSE	0.000250	0.001089	0.000322	0.000184	0.000359
	PSNR	84.78351	77.86603	83.41769	86.10438	82.87494
	SSIM	0.986578	0.945302	0.983216	0.991342	0.981074

MSE 值越小,相似度越高;PSNR 值越大,相似度越高;SSIM 值越大,相似度也越高。由表 2 可见,本文方法得到的边缘保持滤波效果与实际更加接近,在所有的指标上都优于 Xu 的方法。

除了在图像相似度指数上的评估外,图 9 展示了实验中用到的 5 种边缘保持平滑滤波方法的模拟效果。

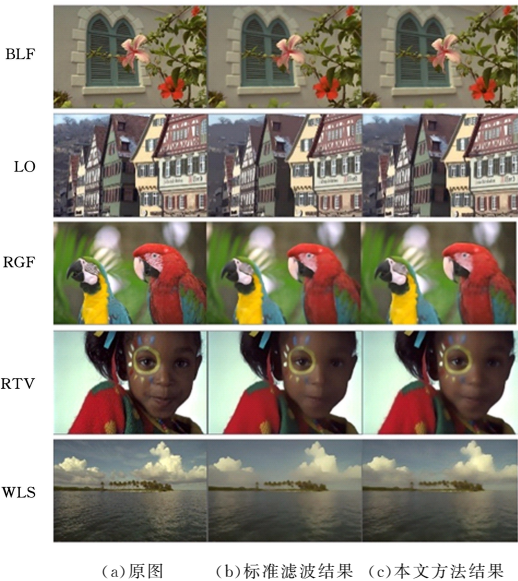


图 9 本文方法与常见滤波方法的结果对比
Fig. 9 Results comparison of proposed method and common filtering methods

在视觉上,本文方法可以保持颜色风格与原图整体一致,不会在梯度图重建彩色图的过程中出现可能会发生的颜色偏差、梯度反转等问题。但文本方法仍有不足之处,比如从最后一栏结果中的天空部分能够看出云彩的痕迹,证明平滑程度不够好。

表 3 展示了本文方法与几种图像边缘保持平滑滤波方法处理“柯达数据集”^[25]得到的平均时间的实验对比结果。

表 3 不同方法的滤波处理效率
Table 3 Efficiency of different methods
(单位:s)

滤波方法	平均时间
WLS ^[1]	3.215 6
RTV ^[11]	3.018 8
RGF ^[19]	0.828 1
L0 ^[5]	1.907 7
BLF ^[4]	28.850 0
RegCov ^[20]	195.063 9
SDFilter ^[21]	59.032 0
LLF-Matlab ^[2]	397.491 9
本文方法	2.288 2

由表 3 可见,本文方法在速度上虽然比滚动引导滤波(RGF)^[19]慢,但仍能够达到与其他图像边缘保持平滑滤波算法相当的水平,例如对于加权最小二乘方法(WLS)^[1]、L0 平滑(L0)^[5]和相对全变分(RTV)^[11]。另外,本文方法对复杂的滤波算法具有明显的加速效果,例如双边滤波(BLF)^[4]、区域协方差方法(RegCov)^[20]、联合静态和动态引导滤波(SDFilter)^[21]和局部拉普拉斯方法的 Matlab 实现(LLF-Matlab)^[2],本文方法可以达到接近 200 倍的加速效果。

结束语 本文结合深度学习,利用卷积神经网络良好的扩展性和灵活性,构建了深度卷积神经网络(DCNN)模型,提出了基于 DCNN 的边缘保持滤波方法,实现了高效拟合多种滤波效果。另外,本文还利用卷积神经网络对边缘保持滤波的梯度信息进行学习,提出基于梯度域卷积神经网络(GC-NN)模型的边缘保持滤波方法,解决了部分滤波处理出现的颜色偏差、梯度反转等问题,提高了处理效率。但本文提出的两种方法主要是针对静态图像的显示优化处理,而动态视频空间上的图像处理有更大的实际应用需求,如何将深度神经网络与动态视频的显示优化问题联系起来,是一个未来值得深入研究的课题。

参 考 文 献

[1] FARBMAN Z,FATTAL R,LISCHINSKI D. Edge-preserving decompositions for multi-scale tone and detail manipulation [J]. ACM Transactions on Graphics (TOG),2008,27(3):1-10.

[2] PARISS,HASINOFFSW,KAUTZJ. LocalLaplacianfilters;edge-aware image processing with a Laplacian pyramid [J]. ACM Transactions on Graphics (TOG),2015,58(3):81-91.

[3] GASTAL E S L,OLIVEIRA M M. Domain transform for edge-aware image and video processing[J]. ACM Transactions on Graphics (TOG),2011,30(4):Article No. 69.

[4] PARIS S,DURAND F. A fast approximation of the bilateral filter using a signal processing approach [J]. International Journal of Computer Vision,2009,81(1):24-52.

[5] XU L,LU C,XU Y,et al. Image smoothing via L 0 gradient mini-

mization [J]. ACM Transactions on Graphics (TOG),2011,30(6):174.

[6] TOMASI C,MANDUCHI R. Bilateral filtering for gray and color images [C]//1998 IEEE Sixth International Conference on Computer Vision (ICCV'98). Piscataway,NJ:IEEE,1998:839-846.

[7] PORIKLI F. Constant time O (1) bilateral filtering [C]//2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08). Piscataway,NJ:IEEE,2008:1-8.

[8] YANG Q,TAN K,AHUJA N. Real-time O (1) bilateral filtering [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09). Piscataway,NJ:IEEE,2009:557-564.

[9] HE K,SUN J,TANG X. Guided image filtering [C]//European Conference on Computer Vision (ECCV'10). New York,NY:Springer,2010:1-14.

[10] PERONA P,MALIK J. Scale-space and edge detection using anisotropic diffusion [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1990,12(7):629-639.

[11] XU L,YAN Q,XIA Y,et al. Structure extraction from texture via relative total variation [J]. ACM Transactions on Graphics,2012,31(6):Article No. 139.

[12] DONG C,LOY C C,HE K et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2016,38(2):295-307.

[13] XIE J,XU L,CHEN E. Image denoising and inpainting with deep neural networks [C]//25th International Conference on Neural Information Processing Systems. Red Hook,NY:Curran Associates,Inc.,2012:341-349.

[14] LIU S,PAN J,YANG M. Learning recursive filters for low-level vision via a hybrid neural network [C]//European Conference on Computer Vision (ECCV'16). New York,NY:Springer,2016:560-576.

[15] LI Y,HUANG J,AHUJA N,et al. Deep joint image filtering [C]//European Conference on Computer Vision (ECCV'16). New York,NY:Springer,2016:154-169.

[16] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//International Conference on Learning Representations 2015 (ICLR2015). San Diego,CA,2015.

[17] SZEGEDY C,LIU W,JIA Y,et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15). Piscataway,NJ:IEEE,2015:1-9.

[18] XU L,REN J,YAN Q,et al. Deep edge-aware filters [C]//International Conference on Machine Learning. 2015:1669-1678.

[19] ZHANG Q,SHEN X,XU L et al. Rolling guidance filter [C]//European Conference on Computer Vision (ECCV'14). New York,NY:Springer,2014:815-830.

[20] KARACAN L,ERDEM E,ERDEM A. Structure-preserving image smoothing via region covariances [J]. ACM Transactions on Graphics,2013,32(6):Article No. 176.

[21] HAM B,CHO M,PONCE J. Robust image filtering using joint static and dynamic guidance [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15). Piscataway,NJ:IEEE,2015:4823-4831.