

网络预出版：

标题：基于卷积神经网络的深度图超分辨率重建

作者：李素梅,雷国庆,范如

收稿日期：2017-06-20

录用日期：2017-08-10

DOI：10.3788/aos201737.1210002

引用格式：

李素梅,雷国庆,范如. 基于卷积神经网络的深度图超分辨率重建[J]. 光学学报, 2017, 37(12): 1210002.

网络预出版文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文：

基于目标提取与引导滤波增强的红外与可见光图像融合

吴一全 王志来

南京航空航天大学电子信息工程学院, 江苏 南京 211106

光学学报, 2017, 37(8): 0810001

基于卷积神经网络与一致性预测器的稳健视觉跟踪

高琳 王俊峰 范勇 陈念年

西南科技大学计算机科学与技术学院, 四川 绵阳 621010

光学学报, 2017, 37(8): 0815003

结合同场景彩色图像的深度图超分辨率重建

王宇 朴燕 孙荣春

长春理工大学电子信息工程学院, 吉林 长春 130022

光学学报, 2017, 37(8): 0810002

航拍图像的路面裂缝识别

王博 王霞 陈飞 贺云涛 李文光 刘莉

北京理工大学光电学院 光电成像技术与系统教育部重点实验室, 北京 100081

光学学报, 2017, 37(8): 0810004

凝视成像降质模型的超分辨率重建

姚烨 乔彦峰 钟兴 于树海 戴路 白杨

中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033

光学学报, 2017, 37(8): 0810003

基于卷积神经网络的深度图超分辨率重建

李素梅 雷国庆* 范如

天津大学电气自动化与信息工程学院, 天津 300072

摘要 针对传统深度图超分辨率重建算法要人工提取特征, 计算复杂度较高, 不容易得到合适表示特征的问题, 提出一种基于卷积神经网络的深度图超分辨率重建算法, 卷积神经网络不需要提前针对特定的任务对图像提取具体的手工特征, 而是模拟人类的视觉系统对原始深度图进行层次化的抽象处理自主提取特征。算法直接进行从低分辨率深度图 (Low resolution depth map:LR) 到高分辨率深度图 (High resolution depth map:HR) 映射的学习。该映射由七个卷积层和一个反卷积层联合实现, 通过卷积操作学习丰富的图像特征, 通过反卷积实现上采样重建高分辨率的深度图。在 Middlebury RGBD 数据集上的实验结果表明, 模型得到的 PSNR 较传统双三次插值算法平均能提高 2.7235dB, RMSE 的值平均降低 0.098; 较经典卷积神经网络算法 (Image Super-Resolution Using Deep Convolutional Networks:SRCNN) PSNR 平均提高 1.5244dB, RMSE 平均降低 0.043。

关键词 图像处理; 超分辨率 (SR); 深度图; 卷积神经网络; 反卷积

中图分类号 TP391.4

文献标识码 A

Depth Map Super-Resolution Based on Deep Convolutional Neural Networks

Li Sumei, Lei Guoqing, Fan Ru

*College of Electrical Automation and Information Engineering, Tianjin University,
Tianjin 300072, China*

Abstract The traditional depth map super-resolution algorithm extract features by hand, It has high computational complexity and not easy to get the appropriate characteristics. In order to solve this problem, we present a depth map super-resolution reconstruction algorithm based on CNN. Unlike the traditional depth map super-resolution reconstruction method, CNN does not need to extract specific manual features for the image in advance for the specific task, but simulate the human visual system to classify the original depth map to get more in line with the characteristics of human visual characteristics. The proposed method directly learns an end-to-end mapping between the low-resolution depth map and high resolution depth map. To be specific the mapping is realized by seven convolution layers and a deconvolution layer. The convolution operation is used to learn the rich image features, and the high resolution depth map is reconstructed by using the deconvolution layer. The experimental results on Middlebury RGBD database show that, compared with the traditional algorithm, the proposed method improves the PSNR with value of 2.7235dB and reduces the RMSE with value

收稿日期: 年-月-日; **收到修改稿日期:** 年-月-日

基金项目: 国家“863”计划资助项目 (2012AA011505, 2012AA03A301); 国家自然科学基金 (61002028)

作者简介: 李素梅 (1975-), 女, 博士, 副教授、主要从事立体信息处理、计算机视觉和模式识别等方面的研究。E-mail: lisumei@tju.edu.cn

***通信联系人.** E-mail: lgq20051118@163.com

of 0.098. Besides, compared with SRCNN, the performance is also improved with 1.5244dB PSNR increment and 0.043 RMSE reduction.

Keywords image processing;super-resolution(SR);depth map; convolutional neural network;deconvolution layer

OCIS Codes 100.2000;100.6640;100.6890;100.4996

1 引言

深度信息的获取已成为计算机视觉领域一个非常热门的研究方向,三维深度信息的获取技术在三维重建、机器人导航、手势识别、电影游戏、虚拟场景建模等方面具有举足轻重的地位,广泛的应用于电视娱乐、文物考古、建筑设计等社会生活的多个领域。早期人们采用立体匹配算法得到左右图像之间的视差从而获取深度信息^[1],但该算法在遮挡区域、弱纹理区域、重复纹理区域匹配效果不好,具有一定的局限性。因此,直接获取深度信息的方法受到人们的关注。TOF 摄像机利用一种新型的传感器技术实现动态场景深度信息的实时获取^[2],但所获取深度图像分辨率低,物体边界处深度信息不准确,极易受噪声影响,且 TOF 摄像机成本较高,很难广泛的应用于实际。为了提高深度图像的分辨率,许多学者开展了关于深度图像超分辨重建的研究。

目前,实现深度图像超分辨率算法主要有两种思想:一种是通过搭建 RGB-Depth 系统挖掘同场景彩色图像和深度图像间的相关对应关系,充分利用彩色图像信息,与深度信息一起建立相应的数学模型,以此来实现深度图的超分辨率重建。Kopf 等^[3]和 He 等^[4]分别采用联合双边滤波函数和引导图像滤波函数指导深度图超分辨率重建, Park^[5]等人结合非局部均值(NLM)^{[6][7]}、最小二乘和边缘加权进行深度图重建。Yang 等人^[8]运用引导滤波器^{[4][9]}来指导深度图重建, Liu^[10]等人运用测地距离计算滤波系数指导重建, Ferstl^[11]等人先对校正后的彩色图像进行计算得到一个各向异性扩散张量,并利用它来指导深度图上采样。Li 等人^[12]提出了一种基于稀疏表示过完备字典的学习算法,将稀疏表示理论运用到深度图重建问题中。Kwon 等^[13]、Kiechle 等^[14]也利用稀疏理论构造了重建模型。总体上来说,这类方法所需图像数量少,系统搭建也比较简单,但仍然存在不完善的地方,算法往往会将彩色图像的纹理信息映射到重建的深度图中,对最终结果造成了一定的影响。

另一种是利用多幅深度图像之间的冗余信息,根据图像退化模型进行建模以实现深度图的超分辨率重建。Gevrekei^[15]采用基于凸集投影的深度图序列重建算法,并由 Tekalp 等^[16]对其进行了扩展,该算法是一种基于集合论的方法。Rajagopalan^[17]假设深度图满足 MRF 先

验分布模型,利用最大后验概率构造目标函数。Diebelld 等^[18]也基于马尔科夫随机场进行深度图重建,Aodha 等^[19]利用匹配的方法搜寻最佳的深度图图像块进行重建,Schuon^[20]则利用正则约束项采用极大似然估计构造目标函数,进行深度图的重建。

这类方法只利用了深度图像的内在信息,由单幅低分辨率深度图像进行超分辨率重建,分辨率提升不大。由于初始深度图像分辨率较低,缺失信息较多,重建后的深度图像还很难满足实际应用的需要,所以这类方法一般需要构造序列深度图像才能得到超分辨率的结果。

2006 年,机器学习大师多伦多大学教授 Ceoffrey Hinton 及其学生 Ruslan 发表在世界顶级学术期刊《科学》上的一篇文章引发了深度学习在研究领域和应用领域的发展热潮^[21]。随着深度学习的迅猛发展,卷积神经网络作为一种特殊的深度学习架构,凭借其优秀的表现受到广泛的关注。

2015 年 Wang 等人^[22]将稀疏编码与深度学习相结合应用于图像重建。2016 年 Dong^[23]等人提出了应用于图像重建的端到端的 3 层超分辨率卷积神经网络(SRCNN),该网络结构简单仅用了 3 个卷积层就学习到了丰富的图像特征,取得了较好的重建效果。深度学习作为机器学习的重要发展,能够逐层地自动学习合适的表示特征,因而在众多领域中得到了广泛应用。卷积神经网络(Convolutional Neural Network:CNN)作为深度学习的三大重要模型之一,继承了深度学习技术自动提取特征的优点,并且通过权值共享大大减少了所需要训练的参数,使 CNN 能快速处理高维图像,还具备一定的平移不变性,因而在图像分类、识别、重建等领域取得了较好的效果。

鉴于卷积神经网络在图像重建领域的良好效果,设计了应用于深度图超分辨率重建的 10 层深度卷积神经网络。采用卷积层与反卷积层联合实现深度图超分辨率重建,与传统算法相比,本算法通过网络的学习能力自主提取特征,不需要对图像进行一系列的预处理,模型简单计算复杂度低,更加满足实际应用的需要。

2 基于卷积神经网络的深度图像超分辨率重建模型

基于卷积神经网络的超分辨率技术可以直接在高、低分辨率图像之间建立 " 端到端 " 的映射,并以轻型结构来达到较好的性能效果。因此设计了利用卷积神经网络的 10 层深度图超分辨率重建模型,网络结构如图 1 所示:

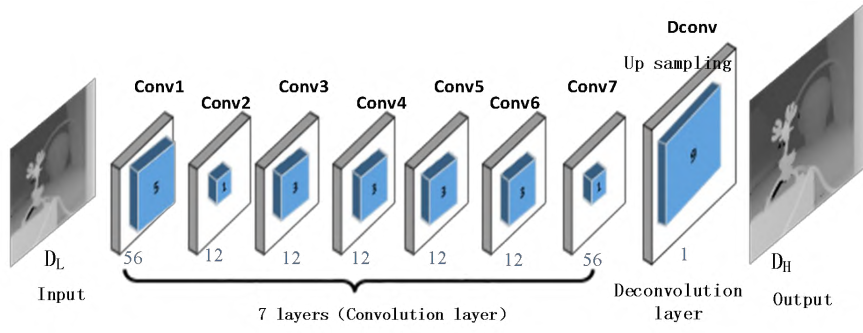


图 1 基于卷积神经网络的深度图超分辨率重建网络

Fig.1 Depth map Super - resolution Network Based on Convolution Neural Network

第一层为输入层，输入一个低分辨率的深度图 D_L ，第二到第八层为卷积层，低分辨率的深度图经过卷积层获得一系列的特征图，这些特征图经过非线性映射成高分辨率的图像块。第九层为反卷积层，经过反卷积操作实现上采样进一步提高图像块的分辨率，最终由第十层输出重建的高分辨率的深度图 D_H 。

2.1 卷积操作

卷积层是卷积神经网络的核心组成部分，其具有局部连接和权值共享特征，卷积层所完成的操作为前一层的一个或者多个特征图作为输入与一个或者多个卷积核进行卷积操作，产生一个或者多个输出。卷积操作可由公式 (1)，(2) 来表示：

$$F^1 = \sigma(W_{c(1)} * D_L + b_1) , \quad (1)$$

$$F^i = \sigma(W_{c(i)} * F^{i-1} + b_i) , \quad (2)$$

F^1 代表第一个卷积层的输出， F^i 代表第 i 个卷积层的输出， D_L 是第一层输入的低分辨率的深度图。 F^{i-1} 代表第 $i-1$ 个卷积层的输出（上一层的输出作为下一层的输入）， $W_{c(1)}$ 代表第一个卷积层的权重， $W_{c(i)}$ 代表第 i 个卷积层的权重。卷积权重对应的是一个数量为 n ，大小为 $f \times f$ 的滤波器组， n 和 f 的值需要具体进行设置。 $*$ 代表卷积操作， b_1 代表第一层的偏置， b_i 代表第 i 层的偏置。其中，偏置的维数始终与该层卷积核的数量保持一致。 σ 代表激活函数，使用参数整流线性单元（PReLU）作为激活函数。其中

$\delta(y) = \max(0, y) + a \min(0, y)$ ， a 是负 y 的可学习的斜率系数。PReLU 是对 ReLU（修正线性单元）的进一步改进，其增加了参数的修正，在一定程度上能够起到正则的效果，并且能够提高模型的泛化能力。

卷积操作可以很好的提取特征，通过 BP（Back Propagation）的误差反向传播^[24]，可以根据不同任务，得到对于这个任务最好的一个参数，学习出相对于这个任务最好的卷积核。对于每个卷积层需要设置的几个参数为：卷积核的大小(Size)，卷积核的数目(Number)，卷积操作的步长以及零填充 (Pad) 的大小。下面简单说明卷积操作的过程。

假设输入图像是一个如图 2 的 4×4 的图像块，卷积核的大小为 2×2 ，偏置初始化为全 0 的矩阵，步长为 1。

第一行元素的计算过程如下：

$$2 = 1 \times 1 + 1 \times 1 + 0 \times 0 + 0 \times 1 + \text{偏置 } 0$$

$$3 = 1 \times 1 + 1 \times 1 + 0 \times 0 + 1 \times 1 + \text{偏置 } 0$$

$$3 = 1 \times 1 + 1 \times 1 + 1 \times 0 + 1 \times 1 + \text{偏置 } 0$$

同理可以计算得到其他行的值。

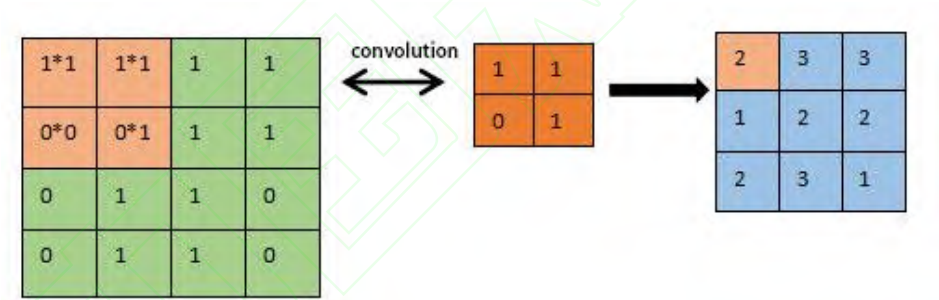


图 2 卷积过程示意
Fig.2 Examples of convolution processes

2×2 的卷积核以步长 1 在原图像块上滑动进行卷积操作。于是 4×4 的图像块经过卷积操作之后得到的大小为 3×3 的图像块，得到的结果经过一个激活函数之后作为卷积操作最终的输出。由以上的过程可以看出，图像块经过卷积操作之后大小会减小，为了保证输出图像的大小不变，论文对所有卷积操作的输出都进行对应的零填充操作。各卷积层的参数设置如表 1 所示，在本模型中所有卷积操作的步长都为 1，其中 Name 为卷积层的名称。

表 1 各卷积层的参数设置

Tab.1 Parameter settings for each convolution layer

Name	Number (n)	Size (f)	Pad
Conv1	56	3×3	1
Conv2	12	1×1	0
Conv3	12	3×3	1
Conv4	12	3×3	1
Conv5	12	3×3	1
Conv6	12	3×3	1
Conv7	56	1×1	0

2.2 反卷积操作

论文通过一个反卷积层实现了最终的重建过程，该层在这里相当于上采样操作。在文献[25]中提出当步长大于1时卷积相当于下采样，而反卷积相当于上采样。当步长为1时，卷积和反卷积操作本质上是相同的，且步长 $s > 1$ 时，整个网络的复杂度会降低 s^2 ，但步长也不能过大，否则会严重影响深度图的重建质量。同时该文献还提出适当增加反卷积核的大小可以取得更好的重建质量。受该文献启发，论文应用反卷积操作来得到最终的重建结果，通过调整步长来实现对采样因子的调整，采用相对较大的卷积核提升重建质量。该过程可以由公式3来表示：

$$F = \sigma(W_d \bullet F^7 + B) \quad , \quad (3)$$

F 代表反卷积层的输出， W_d 代表反卷积层的权重参数， \bullet 代表反卷积操作， F^7 代表最后一个卷积层的输出， B 为偏置。步长根据网络采样因子的大小进行相应的调整，但始终大于1。值得注意的一点是网络最后一层输出图像块大小的计算，假设输入反卷积层的图像大小为 I ，反卷积层参数为：核的大小 $R \times R$ 、步长 s 、填充大小为 p ，那么经过反卷积之后的输出图像的大小为：

$$o = s(I-1) + R - 2p \quad , \quad (4)$$

在本网络模型中反卷积层的各参数设置为 $R=9$ （即核的大小为 9×9 ）， $p=4$ ， $s=3, 4$ 。

2.3 训练

网络通过最小化重建得到的结果和真实视差图之间的代价，来不断调整网络参数 $\Theta = \{W_{c(i)}, W_d, b_i, B\}$ ，对于一组高分辨率的深度图 X_j 和一组由网络重建得到的深度图 $F^j(Y; \Theta)$ ，

论文采用均方误差（MSE）作为代价函数：

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|F^j(Y; \Theta) - X_j\|^2, \quad (5)$$

其中 n 代表训练样本的数量。利用随机梯度下降法和网络的反向传播^[26]来最小化 MSE 以调整网络的参数，网络权重的更新过程为：

$$\Delta_{k+1} = 0.9 \times \Delta_k - \eta \times \frac{\partial L}{\partial W_k^\ell}, W_{k+1}^\ell = W_k^\ell + \Delta_{k+1}, \quad (6)$$

Δ_k 代表上一次的权重更新值， ℓ 代表层数， k 代表网络的迭代次数。 η 是学习率， W_k^ℓ 代表第 ℓ 层的第 k 次迭代时的权重， $\frac{\partial L}{\partial W_k^\ell}$ 代表对代价函数中相应的权重求偏导。权重采用均值为 0，方差为 0.001 的高分布进行随机初始化，模型在训练过程中采用固定学习率。

3 实验结果与分析

从 MPI Sintel 深度数据集[27]和 Middlebury[28]数据集中收集了 61 幅深度图用于网络的训练与测试，其中 52 幅作为原始训练样本，9 幅作为测试样本。

训练样本：卷积神经网络的训练需要大量的数据样本，为了提升训练数据的有效性，本文对原始训练样本进行了扩大。最终训练样本由训练集和验证集两部分组成。训练集通过对原始的 52 幅训练样本进行 90° 、 180° 、 270° 旋转，之后进行 6、7、8、9、10 倍的扩大得到，即一共 1040 ($52 \times 4 \times 5$) 幅图像作为最终的训练集。验证集由从 52 幅深度图中任意选取的 10 幅深度图进行与训练集同样的变换得到，即 200 ($10 \times 4 \times 5$) 幅图像作为训练过程中的验证集。

测试样本：为了进一步验证本模型的有效性，利用 Middlebury RGBD 数据集中的 9 幅高分辨率的深度图像进行网络的测试，利用均方根误差（RMSE）和峰值信噪比（PSNR）作为结果的评价标准。

将训练集中的深度图进行相应倍数的下采样，得到降质的图像，之后采用大小为 7 的步长将降质后的图像分割成 7×7 的图像块，将此图像块作为网络的输入 (D_L)。利用公式 3 计算得到标签的大小，同样利用大小为 7 的步长对训练集（高分辨率未降质的图像块）进行分割，得到网络的标签 (D_H)，对验证集也进行同样的操作。本论文利用 caffe 工具箱进行网络的训练与验证^[29]，训练好网络之后在 MATLAB 上输入测试样本，用来评估不同放大因子时的

网络重建深度图质量。

图 3 展示了相同迭代次数时 7 层、10 层、12 层卷积神经网络针对同一幅图 (Art) 在 PSNR 上收敛曲线的比较, 红色代表 10 层的网络, 蓝色代表 12 层的网络, 黄色代表 7 层的网络。

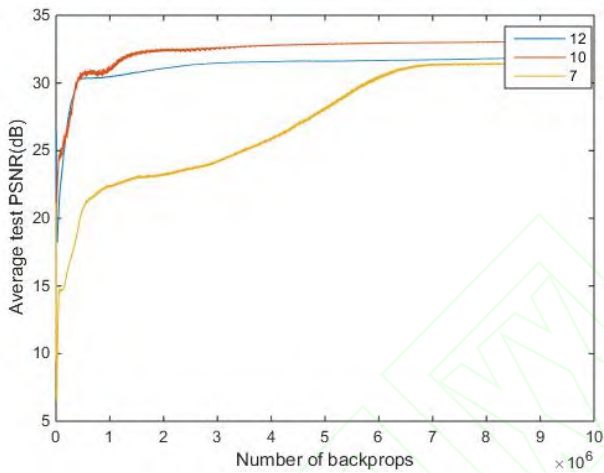


图 3 不同网络层数获得的重建结果对比

Fig.3 The Comparison of reconstruction results obtained by different layers

表 2 进一步给出了在相同条件下不同网络层数重构的高分辨率深度图像的具体 PSNR 值, 以及迭代相同次数时不同层数的网络所需要的总的测试时间。图 3 和表 2 的结果表明, 深度图超分辨率重建网络并不是“越深越好”, 这可能是由于训练困难所致。本网络不包含池化层或全连接层, 因此对初始化参数和学习速率比较敏感。当加深网络层数时, 很难设定保证收敛的学习率。即使它收敛, 网络也可能会陷入不利的局部最小值。为什么不是“越深越好”需要进一步研究更深层网络的梯度和训练动态。同时随着网络层数的增加需要训练的参数增多, 这必然会导致时间复杂度增加, 付出相应的时间代价。最终, 在权衡时间代价以及重建质量之后本文采用了 10 层卷积神经网络。

表 2 不同网络层数的比较 (time 和 PSNR)

Tab.2 Quantitative comparison of different layers (time and PSNR)

	Time(s)	PSNR(dB)
7 layers	92.702	31.5321
10 layers	113.076	33.4809
12 layers	127.040	31.8638

图 4 给出了运用大小为 3 的采样因子得到 D_L , 之后利用网络进行训练的结果。通过与经典双三次插值算法和经典卷积神经网络算法 (SRCNN) ^[23] 在相同采样因子下重构得到的深度图

进行对比，由局部放大图可以看出本模型训练得到的网络在一定程度上解决了图像模糊问题，得到的深度图像视觉效果有所提高。

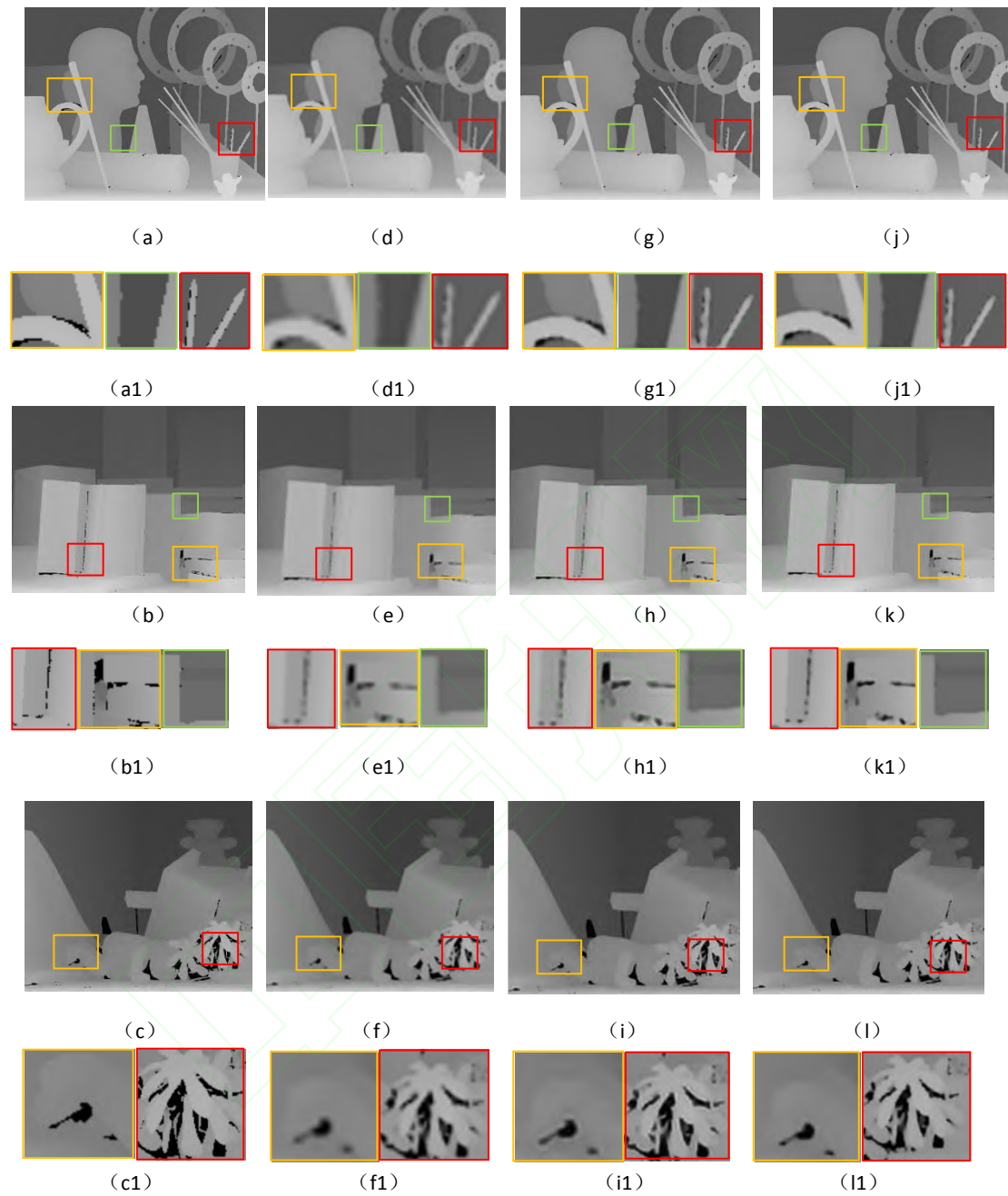


图 4 不同方法获得的重建结果对比

Fig.4 The Comparison of reconstruction results obtained by different methods

图 4 中 a-c 分别为 Art、Books、Teddy 的原始高分辨率的深度图，a1-c1 分别为它们局部放大的结果，d-f 是双三次插值重建得到的深度图，d1-f1 为局部放大的结果，g-i 为 SRCNN 重建得到的深度图，g1-i1 为重建结果的局部放大，j-1 是本模型得到的结果，j1-11 是结果的局部放大。

表 3 给出了重构的高分辨率深度图像在峰值信噪比 (PSNR) 和均方根误差 (RMSE) 上的客观比较, 计算 PSNR 的方式来自文献[23], 计算 RMSE 的方式来自文献[13]。由表 2 中的值可以看出与经典的算法相比较本算法具有明显的优势, 在所有测试图像上都能取得较好的效果, PSNR 的值对比双三次插值算法平均能提高 2.7235dB, 对比文献[23]能平均提高 1.5244dB。RMSE 的值对比双三次插值算法能平均降低 0.098, 对比文献[23]平均降低 0.043。从该实验结果可知, 本模型针对不同的输入较经典算法均能取得较好的重建结果。

表 3 不同算法的在数据集上的定量比较 (RMSE 和 PSNR)

Tab.3 Quantitative comparison of different algorithms on the dataset (RMSE and PSNR)

	Bicubic		SRCNN[23]		ours	
	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR
Art	1.5313	31.3049	1.3814	32.9155	1.3464	33.4809
Books	1.1334	31.2227	1.1324	31.7996	1.0385	33.7587
Moebius	1.1412	31.9929	1.1334	32.6431	1.0513	34.6130
Dolls	1.1052	32.6389	1.1043	33.1548	1.1017	34.7382
Laundry	1.1572	33.6228	1.0848	34.8337	1.0744	36.1265
Reindeer	1.1758	32.6560	1.0806	34.2127	1.0373	35.5582
Cones	1.4905	29.6001	1.4425	30.4998	1.3820	32.1335
Teddy	1.4659	28.8076	1.3950	29.6820	1.3430	31.3920
Venus	0.5610	43.8180	0.5076	45.7149	0.5014	48.3748

为了进一步验证本模型的有效性, 与现有的先进的算法进行了对比实验。表 4 到表 6 给出了采样倍数为 4 时本算法与不同优秀算法在 RMSE 上的比较。通过对比可以看出, 本算法对于不同的测试图片均能取得较好的重建结果, 表中用下划线表示位列第二的重建结果, 用粗体表示最好的重建结果。结果表明, 本模型优于只利用深度信息的重建算法[5][18][19]和传统的与彩色图像相结合的算法[8][9][11]。与现有的基于实例的算法[12-14]相比, 本模型无需明确学习字典, 没有复杂的预处理过程, 所有的特征均通过网络自主提取。由表 4 的 Dolls 数据看出, 本模型在 Dolls 这幅图上得到的 RMSE 值低于文献[22]采用的将稀疏编码与深度学习相结合的算法。可能的原因是 Dolls 这幅图像纹理丰富, 需要提取的细节信息较多, 因此, 可以推测, 如果针对纹理相对丰富的图像, 或许需要建立更为深层的网络来提取更多的细节特征, 以提高重建质量。尽管如此, 本模型对比很多算法仍取得了较好的重建结果, 而且网络模型简单, 具有一定的实际应用价值。

表 4 不同算法的在数据集上的定量比较 (RMSE)

Tab.4 Quantitative comparison of different algorithms on the dataset (RMSE)

	Art	Books	Moebius
Bilinear	4.147	1.673	1.449
MRFs[18]	3.794	1.546	1.439
Park et al.[5]	<u>3.498</u>	<u>1.530</u>	<u>1.349</u>
Guided[9]	3.788	1.572	1.434
Ferstl et al[11]	3.785	1.603	1.458
ours	1.692	1.248	1.257

表 5 不同算法的在数据集上的定量比较 (RMSE)

Tab.5 Quantitative comparison of different algorithms on the dataset (RMSE)

	Dolls	Laundry	Reindeer
Bicubic	1.305	2.408	2.809
Park et al.[5]	1.301	2.132	2.407
Aodha et al.[19]	1.977	2.969	3.178
Ferstl et al.[11]	1.355	2.511	2.712
AP[8]	1.350	2.255	2.431
Wang et al.[22]	0.989	<u>1.630</u>	<u>1.914</u>
ours	<u>1.274</u>	1.311	1.275

表 6 不同算法的在数据集上的定量比较 (RMSE)

Tab.6 Quantitative comparison of different algorithms on the dataset (RMSE)

	Venus	Teddy	Cones
Park et al.[5]	1.8	4.89	5.64
Li et al.[12]	3.55	4.92	6.34
Ferstl et al.[11]	2.52	3.3	4.45
Kiechle et al.[14]	<u>0.82</u>	1.82	2.97
Wang et al.[22]	1.19	2.03	3.08
Kwon et al[13]	1.14	<u>1.80</u>	<u>2.13</u>
ours	0.61	1.52	1.57

4 结论

目前的深度图的超分辨率重建主要是通过手工提取特征,对于不同的图像重建问题,往往需要采用不同的特征提取方法。卷积神经网络理论的出现给深度图超分辨率重建带来了新思路,其充分考虑了图像目标在空间中的平移、旋转和缩放不变性,将传统图像重建技术上的特征提取和聚合结合在一起,利用网络本身的学习能力主动提取特征。通过深入研究卷积神经网络在深度图像超分辨率重建上的应用,提出了一种基于深度卷积神经网络的深度图超

分辨率重建算法。利用卷积神经网络庞大的学习能力解决了传统算法计算复杂度高,不能有效提取特征,实际应用代价大的缺点。本模型在整个重建过程中没有进行任何网络之外的预处理过程,实验结果表明,本算法可重构得到高分辨率高质量的深度图像,能够应用于实际。下一步,研究内容应重点考虑加深网络的深度,或者采取其他更有效的网络模型,获得更优的重建效果。

参考文献

- 1 Jiang Hongzhi,Zhao Huijie,Liang Xiaoyue.Fast phase stereo matching based on pole correction[J].Optics and Precision Engineering, 2011, (10):2520-2525.
姜宏志,赵慧洁,梁宵月,李冬. 基于极线校正的快速相位立体匹配[J]. 光学精密工程, 2011, (10):2520-2525.
- 2 Liu jiaoli.Li sumei,Li yongda,Liu fuyan.High-resolution depth maps based on TOF-Stereo fusion[J]. Information Technology, 2016, (12):190-193.
刘娇丽,李素梅,李永达,刘富岩. 基于 TOF 与立体匹配相融合的高分辨率深度获取[J]. 信息技术, 2016, (12):190-193.
- 3 Kopf J, Cohen M F, Lischinski D, et al. Joint bilateral upsampling[C],ACM, 2007:96.
- 4 He K, Sun J, Tang X. Guided image filtering[C], European Conference on Computer Vision. Springer-Verlag, 2010:1-14.
- 5 Park J, Kim H, Tai Y W, et al. High quality depth map upsampling for 3D-TOF cameras[C],IEEE International Conference on Computer Vision. IEEE, 2012:1623-1630.
- 6 Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one[J]. Siam Journal on Multiscale Modeling & Simulation, 2005, 4(2):490--530.
- 7 Buades A, Coll B, Morel J M. Image denoising methods. A New Nonlocal Principle[J]. Siam Review, 2010, 52(1):113-147.
- 8 Yang Y, Wang Z. Range image super-resolution via guided image filter[C],International Conference on Internet Multimedia Computing and Service. 2012:200-203.
- 9 He K, Sun J, Tang X. Guided image filtering. [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(6):1397-1409.
- 10 Liu M Y, Tuzel O, Taguchi Y. Joint geodesic upsampling of depth images[C], IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2013:169-176.
- 11 Ferstl D, Reinbacher C, Ranftl R, et al. Image guided depth upsampling using anisotropic total generalized variation[C],IEEE International Conference on Computer Vision. IEEE Computer Society, 2013:993-1000.

- 12 Li Y, Xue T, Sun L, et al. Joint example-based depth map super-resolution[C], IEEE International Conference on Multimedia and Expo. IEEE, 2012:152-157.
- 13 Kwon HH, Tai YW, Lin S. Data-driven depth map refinement via multi-scale sparse representation[J]. 2015:159-167.
- 14 Kiechle M, Hawe S, Kleinsteuber M. A joint intensity and depth Co-sparse analysis model for depth map super-resolution[C], IEEE International Conference on Computer Vision. IEEE Computer Society, 2013:1545-1552.
- 15 Gevrekci M, Pakin K. Depth map super resolution[C], Signal Processing and Communications Applications. IEEE, 2011:502-505.
- 16 Patti A J, Sezan M I, Tekalp / M. High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur[M]. 1994.0
- 17 Rajagopalan A N, Bhavsar A, Wallhoff F, et al. Resolution enhancement of PMD range maps. [C], Pattern Recognition, Dagm Symposium, Munich, Germany, June 10-13, 2008, Proceedings. DBLP, 2008:304-313.
- 18 Diebel J, Thrun S. An application of markov random fields to range sensing[J]. Advances in Neural Information Processing Systems, 2005:291--298.
- 19 Aodha OM, Campbell NDF, Nair A, et al. Patch based synthesis for single depth image super-resolution[C], European Conference on Computer Vision. Springer-Verlag, 2012:71-84.
- 20 Schuon S, Theobalt C, Davis J, et al. LidarBoost: Depth superresolution for ToF 3D shape scanning[C], Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE Xplore, 2009:343-350.
- 21 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. [J]. Science, 2006, 313(5786):504-7.
- 22 Wang Z, Liu D, Yang J, et al. Deep networks for image super-resolution with sparse prior[C], IEEE International Conference on Computer Vision. IEEE, 2015:370-378.
- 23 Dong C, Chen C L, He K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(2):295.
- 24 Durbin R, Rumelhart D E. Product Units: A computationally powerful and biologically plausible extension to backpropagation networks[J]. Neural Computation, 1989, 1(1):133-142.
- 25 Yu K, Dong C, Chen C L, et al. Deep convolution networks for compression artifacts reduction[J]. 2016.
- 26 Lécun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- 27 <http://sintel.is.tue.mpg.de/>
- 28 <http://vision.middlebury.edu/stereo/>