

基于双通道卷积神经网络的深度图超分辨率研究

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 目前直接获取的深度图受其成像原理及硬件设备等因素的限制, 存在分辨率低、边缘信息丢失等缺点, 极大地影响了深度图的应用。针对这一问题, 提出基于双通道卷积神经网络的深度图超分辨率重建模型。该模型由深、浅两个通道组成, 21 层的深层通道通过联合卷积与反卷积, 结合跳跃连接与多尺度理论, 实现深度图细节特征的快速学习; 3 层的浅层通道用于学习深度图的轮廓特征; 最后融合深、浅两个通道, 将细节与轮廓相结合, 实现由低分辨率深度图到高分辨率深度图的端到端的学习。该模型充分利用卷积神经网络的学习能力自主提取深度图的有效特征, 避免了手工提取特征的不准确性。在 Middlebury RGBD 数据集上的实验结果表明, 本文模型在大采样因子 8 时仍能取得较好的效果, 具有很高的实际应用价值。

关键词 图像处理; 超分辨率; 深度图; 卷积神经网络; 残差网络

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/AOS201838.1010002

Depth Map Super-Resolution Based on Two-Channel Convolutional Neural Network

Li Sumei, Lei Guoqing*, Fan Ru

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract The depth map obtained directly is limited by the disadvantages such as low resolution and missing edge information, it greatly affects the application of depth map. In order to solve this problem, a two-channel convolutional neural network for depth map super-resolution is proposed. It consists of two channels, deep and shallow, and there are 21 layers in the deep network. Through joint convolution and deconvolution, combining skip connection and multi-scale theory, the deep channels can quickly learn the detailed features of depth map. Shallow network of 3 layers are used to learn the rough features of depth maps. Finally, the two channels are combined with details and outlines to realize end-to-end mapping from low resolution depth map to high resolution one. The model makes full use of the learning ability of the convolutional neural network to independently extract the effective features of the depth map and avoid the inaccuracy of manually extracting features. The experimental results on the Middlebury RGBD dataset show that the proposed model can achieve good results at a large sampling factor of 8, and has a high practical value.

Key words image processing; super-resolution; depth map; convolutional neural network; residual network

OCIS codes 100.2000; 100.6640; 100.6890; 100.4996

1 引 言

三维技术在航空航天、远程医疗、军事训练、教育教学、三维电视, 以及游戏娱乐等领域具有广阔的应用前景和商业价值, 深度信息的精确获取是三维研究所需要的关键技术。目前, 深度图可由硬件设备直接获取^[1-5], 也可由算法生成。受硬件本身系统的影响, 直接获取的深度图噪声大、分辨率低, 且成

本较高; 但由算法生成的深度图, 不仅分辨率高, 而且成本低。因此, 通过算法实现深度图质量的提高已成为该领域的研究热点。

现有算法主要有两种思路: 1) 基于深度图序列的重建方法。Schuon 等^[6]利用双边全变差正则化项, 限制图像总能量进行深度图重建。Gevrekci 等^[7]基于凸集投影理论进行序列深度图超分辨率重建。Rajagopalan 等^[8-9]结合马尔可夫随机场, 采用

收稿日期: 2018-02-06; 修回日期: 2018-04-26; 录用日期: 2018-05-16

基金项目: 国家自然科学基金(61002028)、国家 863 计划(2012AA011505, 2012AA03A301)

* E-mail: lgq20051118@163.com

最大后验概率方法实现深度图超分辨率重建。Timofte 等^[10]改进稀疏编码实现重建。此类方法仅利用深度图信息,致使最终分辨率仅提高 2~3 倍。由于初始深度图分辨率本身就很低,所以利用该方法重建后的深度图很难满足实际应用要求。2) 基于信息融合技术的重建方法。Kopf 等^[11-13]利用图像滤波函数指导深度图超分辨率重建,但重建结果存在边缘细节丢失。为了优化图像滤波过程,Liu 等^[14]利用测地距离计算滤波系数。Park 等^[15]将非局部均值(NLM)^[16-17]、最小二乘和边缘加权相结合来恢复边缘细节。Ferstl 等^[18]通过计算得到的各向异性扩散张量来指导深度图上采样。Li 等^[19]、Kwon 等^[20]、Kiechle 等^[21]和 Lu 等^[22]搭建了基于稀疏字典的彩色图引导的深度图重建模型。Diebel 等^[23]则利用马尔可夫随机场(MRF)实现深度图重建。总体来说,这类方法所需图像数量少,系统搭建比较简单,但通常会将彩色图像的纹理信息映射到重建的深度图中,影响深度图的质量。

上述传统算法通常需要人工提取特征,计算复杂度较高。而深度学习作为机器学习的重要发展,能够逐层自动学习合适的表示特征,已在众多领域中得到广泛应用。卷积神经网络(CNN)作为深度学习的代表,能快速处理高维图像,在图像分类、识别、重建等领域展示出较好的效果。Dong 等^[24]提出了应用于平面图像超分辨率重建的卷积神经网络(SRCNN),SRCNN 先对低分辨率的图像块进行双三次插值,然后将插值后的图像作为网络的输入,SRCNN 开启了深度学习在图像重建领域的研究热

潮。Wang 等^[25]将稀疏编码与深度学习相结合应用于图像重建,也取得了较好的效果。

SRCNN 在图像重建领域取得了很好的效果,但仍然存在不足:1) 网络较浅仅有 3 层不能充分学习图像特征;2) 模型不能实现端到端的上采样,需要先进行双三次插值再送入网络;3) 卷积核过于单一,不能有效提取图像的多尺度特征。针对 SRCNN 的不足,本文搭建了基于双通道卷积神经网络的深度图超分辨模型。模型由深层、浅层两个通道组成,在 21 层的深层网络中通过引入残差单元加速网络的收敛;3 层的浅层网络与 SRCNN 类似。本文算法在深、浅两个通道中采用反卷积实现上采样,避免了网络外的预处理实现了端到端的上采样;为了提取不同尺度的特征,在深层网络中采用大小不同的 4 个卷积核并行提取图像特征。结果表明,相比其他算法,本文算法在深度图的重建任务上取得了更好的表现。

2 网络模型

本文算法的网络模型如图 1 所示,黑色线代表深层网络的实现,绿色线代表浅层网络的实现。相同颜色的矩形块代表这些层输出的特征图的数目相同,具有红色边框的矩形块代表反卷积层,图中每个块右下角的数字为该层输出特征图的数目,块中的数字代表该层卷积核的大小。图 1 中具有相同输出特征图数目和卷积核大小的卷积层,参数设置相同。因此,本文模型共有 11 种卷积层,2 种反卷积层,表 1 给出了这些层的具体参数,其中“ C_k ”为第 k 种

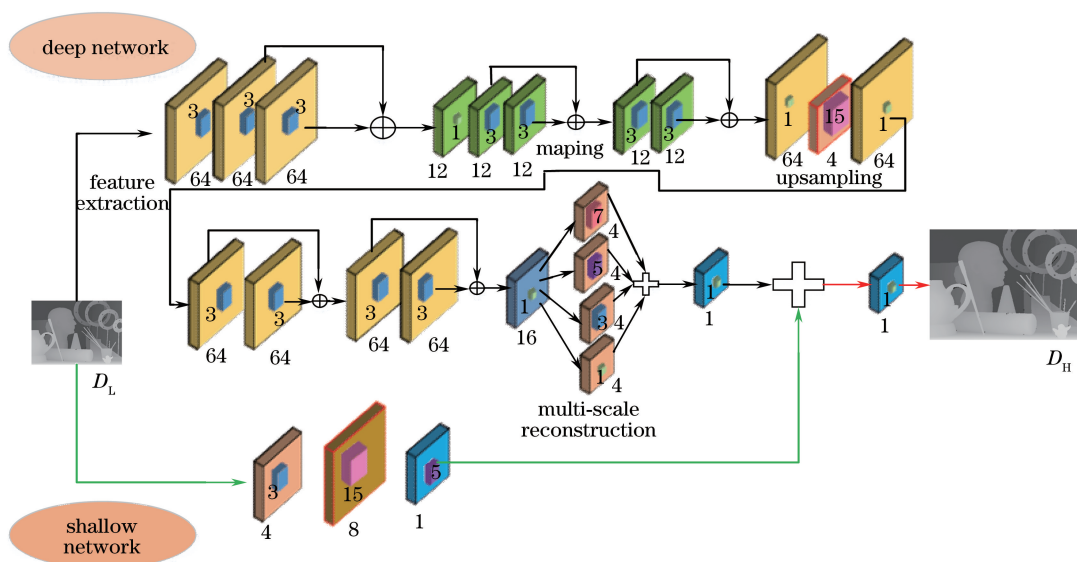


图 1 双通道卷积神经网络的深度图超分辨率模型

Fig. 1 Two-channel convolutional neural network for depth map super-resolution model

表 1 模型参数设置

Table 1 Parameter settings of the model

Name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	DC1	DC1
Number	64	64	16	12	12	4	4	4	4	1	1	4	8
Size	3×3	1×1	1×1	3×3	1×1	7×7	5×5	3×3	1×1	5×5	1×1	15×15	15×15
Pad	1	0	0	1	0	3	2	1	0	2	0	7	7
Step	1	1	1	1	1	1	1	1	1	1	1	4/8	4/8

卷积层,“DC k ”为第 k 种反卷积层,“Number”为该层输出特征图的数目,“Size”为该层滤波器的大小,“Pad”为该层填充的大小,“Step”为该层步长的大小。

2.1 深层通道

深层通道的实现包括 4 步:特征提取、非线性映射、上采样,以及多尺度变换。

1) 特征提取。特征提取阶段包含 3 个卷积层,

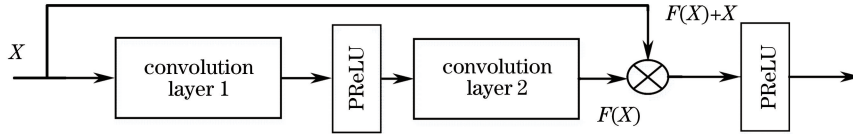


图 2 残差单元

Fig. 2 Residual unit

传统的线性整流函数(ReLU)激活函数随着网络训练次数的增加可能会落入硬饱和区,导致权重无法更新。本文模型利用 PReLU 作为激活函数,其函数负半轴增加了参数修正,避免了神经元死亡,同时在一定程度上起到正则的效果,提高了模型的泛化能力。PReLU 函数表示为

$$\delta(y) = \max(0, y) + a \min(0, y), \quad (1)$$

式中 y 为输入信号, a 为函数负轴的可学习系数,当 $a=0$ 时为 ReLU 的表达式。

卷积层是卷积神经网络的核心组件,具有局部连接和权值共享特征,卷积层将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,产生一个或者多个输出。卷积操作可表示为

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

式中: F^i 为第 i 个卷积层的输出,当 $i=1$ 时 $F^0 = D_L$, D_L 为第一层输入的低分辨率的深度图; $W_{c(i)}$ 为第 i 个卷积层的权重; $*$ 为卷积操作, b_i 为第 i 层的偏置。

2) 非线性映射。该阶段利用 12 个 1×1 卷积核将特征提取阶段输出的 64 维向量映射成 12 维的低维向量,然后利用 4 个卷积层,每层用 12 个 3×3 的滤波核增加映射的非线性,这个阶段也同样利用了跳跃连接加速网络收敛。

每层有 64 个 3×3 大小的滤波器组。为了加速网络收敛,在模型中使用跳跃连接,形成残差单元^[26],如图 2 所示。模型中残差单元的输出是由两个卷积层级联的输出 $F(X)$ 与输入 X 相加,再经过参数线性整流函数(PReLU)激活后得到。残差单元通过跳跃一个或多个层次的连接使得加深后模型所产生的训练误差不高于其所基于的较浅模型,避免了由于梯度消失而形成欠拟合缺陷,提高了重建深度图的质量。

3) 上采样。本文模型通过反卷积层实现上采样操作,实现了“端到端”的映射,而大部分算法采用网络外的双三次插值实现上采样。假设输入反卷积层的图像大小为 I ,反卷积层参数包括:核的大小 $R \times R$ 、步长 s 和填充大小 p ,那么经过反卷积之后的输出图像的大小为

$$o = s(I - 1) + R - 2p. \quad (3)$$

通过适当调整反卷积的参数便可以实现不同倍数的上采样。

4) 多尺度变换。高分辨率图像重建通常依赖于不同尺寸的特征,本文算法利用多尺度卷积核提取图像特征。多尺度重建部分由 10 层组成,开始的 4 个 64 维 3×3 的卷积层用于提取高维特征,同样用到跳跃连接。接下来利用 16 个 1×1 卷积核降低维度,多尺度卷积核大小分别为 $5 \times 5, 3 \times 3, 7 \times 7, 1 \times 1$,通过多尺度卷积层可以得到 16 个特征图,最后采用 1×1 的卷积核连接多尺度特征,输出重建结果。

2.2 浅层通道

浅层通道由 2 个卷积层和 1 个反卷积层组成,结构简单,用于实现对图像粗略特征的提取。浅层通道的设计源于 SRCNN,与 SRCNN 不同的是,本文利用反卷积在通道中实现上采样。模型通过将深、浅层两通道结合,实现深度图丰富特征的学习,

取得了优异的重建结果。

2.3 网络训练

本算法采用最小化欧氏距离来优化模型参数。对于一组由网络重建得到的深度图 $F^i(Y; \Theta)$, 和一组无损的高分辨率的深度图 X_i , 只要使得均值误差最小, 即可得到最优的参数 Θ , 均值误差为

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|F^i(Y; \Theta) - X_i\|^2. \quad (4)$$

训练过程中采用基于反向传播^[27]的随机梯度下降法来优化回归。本文模型卷积层中的所有权重采用均值为 0, 方差为 0.001 的高分布进行随机初始化。

2.4 数据集

本文从 MPI Sintel 和 Middlebury 数据集中收集了 61 幅深度图, 52 幅为原始训练样本, 9 幅为验

证样本。训练样本由网络的输入(低分辨率深度图 D_L)和标签(高分辨率深度图 D_H)两部分组成。首先对原始的 52 幅训练样本进行扩充, 采用步长 7 对降质图像进行切块, 得到 544388 个图像块作为训练集, 44956 个图像块作为测试集。网络训练好之后, 输入 9 幅验证样本, 评估不同放大因子时网络的重建深度图质量, 进一步验证网络的有效性。

3 实验结果与分析

模型由深、浅两通道组成, 各自实现不同的任务。图 3(a)为深层网络最终的输出结果, 该结果恢复了深度图的更多信息。图 3(b)为浅层网络最终的输出结果, 该结果恢复了深度图的轮廓信息。图 3(c)为深、浅两个通道组合后的输出结果, 实现了粗略与细节的融合, 完成了深度图的重建任务。

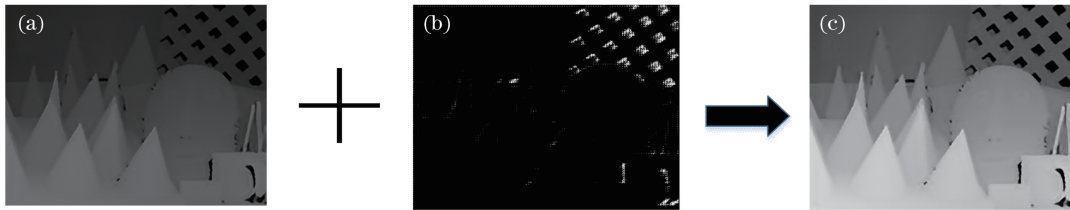


图 3 (a)深层网络、(b)浅层网络和(c)融合后网络输出结果

Fig. 3 (a) Deep network, (b) shallow network and (c) converged network output

大部分算法在深度图重建中会面临图像模糊问题, 图 4 所示为本文模型所得深度图“Teddy”与双三次插值算法和 SRCNN 算法^[24]在相同采样因子

下的重构结果对比。由局部放大图可以看出, 本文算法在一定程度上解决了图像模糊问题, 更接近于真实深度图。

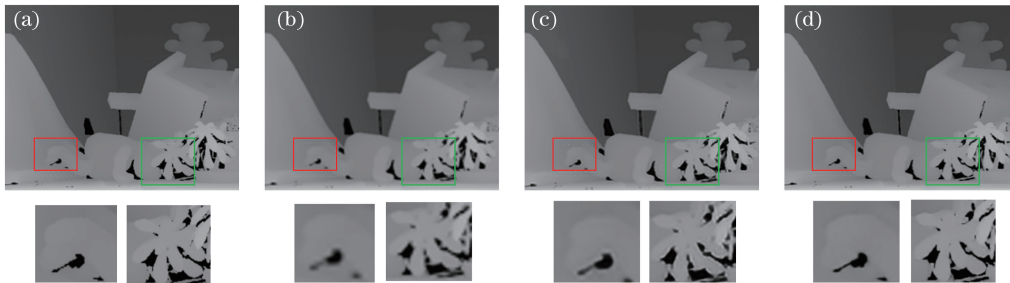


图 4 不同方法获得的重建结果对比。(a)原始图像;(b)双三次插值算法;(c) SRCNN 算法;(d)本文算法

Fig. 4 Comparison of reconstruction results obtained by different methods. (a) Original image; (b) bicubic algorithm; (c) SRCNN algorithm; (d) proposed algorithm

通过对 Middlebury RGBD 数据集中的 9 幅测试图进行分析, 证明模型的超分结果, 由于所对比算法并没有都进行 9 幅图的测试, 为了客观公正, 本文按对比文献进行了 3 种测试深度图的比较, 如表 2~4 所示, 表中用粗体表示最优的重建结果, 用下划线表示次优的重建结果。从表中可以看出, 本文算法对于不同的测试深度图均能取得较好的重建结

果。本文模型与传统算法相比, 无需人工提取特征; 与现有的基于稀疏的算法相比, 无须明确学习字典, 没有复杂的预处理过程。所对比的 SRCNN 模型是针对深度图重新训练得到的, 表 4 中 SRCNN2 是 SRCNN 改进后的网络, 原始 SRCNN 采用 ReLU 作为激活函数, 改进后的 SRCNN 采用 PReLU 作为激活函数。调整了激活函数之后, RMSE 平均降

低了 0.093,充分验证了 PReLU 作为网络激活函数的有效性。本文模型在采样因子为 8 时取得了十分优异的结果,在 4 倍采样因子时部分图片上的验证结果低于文献[21]和[25],由于这两篇文献都利用

了稀疏编码的思想,因此可知稀疏理论在重建任务中仍具有一定的研究价值,以后工作中可以考虑将稀疏理论与卷积神经网络相结合搭建更优异的重建模型。

表 2 不同算法在 Art、Books、Moebius 上的定量比较(RMSE)

Table 2 Quantitative comparison of different algorithms on Art,Books and Moebius (RMSE)

Algorithm	Art		Books		Moebius	
	4×	8×	4×	8×	4×	8×
Bilinear	4.147	5.995	1.673	2.394	1.499	2.198
Ref.[23]	3.794	5.503	1.546	2.209	1.439	2.054
Ref.[13]	4.056	4.712	1.701	1.949	1.386	1.820
Ref.[15]	3.498	4.165	1.530	1.994	1.349	1.804
Ref.[12]	3.788	4.974	1.572	2.097	1.434	1.878
Ref.[18]	3.785	4.787	1.603	1.992	1.458	1.914
Ref.[22]	—	5.798	—	2.728	—	2.422
Ref.[25]	2.525	3.957	1.098	1.646	0.979	1.459
Proposed	1.613	2.185	1.195	1.448	1.198	1.457

表 3 不同算法在 Dolls、Laundry、Reindeer 上的定量比较(RMSE)

Table 3 Quantitative comparison of different algorithms on Dolls,Laundry and Reindeer (RMSE)

Algorithm	Dolls		Laundry		Reindeer	
	4×	8×	4×	8×	4×	8×
Bicubic	1.305	1.855	2.408	3.452	2.809	3.986
Ref.[15]	1.301	1.745	2.132	2.770	2.407	2.987
Ref.[9]	1.977	—	2.969	—	3.178	—
CLMF0 ^[28]	1.271	1.878	2.312	3.084	2.690	3.417
CLMF1 ^[28]	1.267	1.707	2.512	2.892	2.699	3.331
Ref.[18]	1.355	1.859	2.511	3.757	2.712	3.789
Ref.[21]	0.921	1.259	1.212	2.077	1.559	2.583
Ref.[25]	0.989	1.445	1.630	2.466	1.914	2.878
Proposed	0.842	1.462	1.285	1.547	1.237	1.600

表 4 不同算法的在 Venus、Teddy、Cones 上的定量比较(RMSE)

Table 4 Quantitative comparison of different algorithms on Venus,Teddy and Cones (RMSE)

Algorithm	Venus		Teddy		Cones	
	4×	8×	4×	8×	4×	8×
Ref.[9]	2.597	—	4.030	—	5.740	—
Ref.[10]	2.331	—	3.718	—	5.490	—
Ref.[21]	0.819	1.169	1.822	2.370	2.974	4.516
Ref.[18]	1.742	—	2.595	—	3.498	—
Ref.[22]	1.734	2.134	2.723	3.468	3.985	5.344
SRCNN ^[24]	0.789	1.706	1.985	3.252	3.585	5.180
SRCNN2	0.718	1.593	1.891	3.136	3.439	5.171
Ref.[25]	1.191	1.786	2.026	3.015	3.078	4.865
Proposed	0.531	0.749	1.375	1.780	1.448	1.898

表 5 所示为重构的高分辨率深度图在 4 倍采样因子时,在 PSNR 和 SSIM 两类客观评价指标上与 SRCNN 和 Bicubic 算法的客观比较结果,本文算法在 9 幅测试图像上都取得了最高的 PSNR 值,对比

SRCNN 算法平均提高约 0.5 dB。在测试深度图“cones”上,本文模型得到的 SSIM 值略低于 SRCNN 算法,但总体上,本文模型得到的 SSIM 值十分优异。

表 5 4 倍采样因子下不同算法的定量比较

Table 5 Quantitative comparison of different algorithms by a scaling factor of 4

Dataset	Bicubic algorithm		SRCNN algorithm ^[24]		Proposed algorithm	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Art	29.81	0.8799	31.20	0.9260	31.82	0.9562
Books	30.22	0.9283	31.33	0.9425	31.88	0.9639
Dolls	31.67	0.9179	32.44	0.9300	32.61	0.9438
Laundry	32.64	0.9189	34.24	0.9330	34.37	0.9442
Moebius	30.92	0.9258	31.89	0.9403	32.52	0.9536
Reindeer	31.39	0.9290	33.30	0.9519	33.86	0.9647
Cones	28.61	0.8993	29.81	0.9190	30.57	0.9048
Teddy	27.61	0.9167	28.52	0.9343	29.70	0.9474
Venus	42.45	0.9827	45.52	0.9898	46.19	0.9923

为了验证本文模型的时效性,表 6 比较了本文算法与 Bicubic 和 SRCNN 算法在 9 幅深度图上的测试时间复杂度,单位为 s。所有算法运行在 8 G 内存 2.6 GHz 主频的处理器,Win7 64 位系统,使用 MATLAB 2014a 版本软件。从表 6 可以看出,Bicubic 算法的时间复杂度最低,处理一幅深度图平均只需要 0.0079 s,但从图 4 可以看出,该算法重构的图像边缘模糊、超分辨率能力有限。本文算法除了在 Art 上时间复杂度略高于 SRCNN 算

法外,在其他深度图上所需要的测试时间均低于 SRCNN 算法,平均测试时间高于 SRCNN 算法 0.0502 s。SRCNN 算法需要双三次插值的预处理,在一定程度上限制了其测试的速度,而本文模型通过引入反卷积实现了端到端的上采样,避免了模型之外的预处理,大幅提高了模型的时效性。对于一幅 450 pixel×370 pixel 的深度图,本文模型需要的测试时间为 0.2014 s,能够满足实际应用的需要。

表 6 不同算法时间复杂度比较

Table 6 Comparison of time complexity of different algorithms

Algorithm	Art	Books	Dolls	Laundry	Moebius	Reindeer	Cones	Teddy	Venus
Bicubic	0.0398	0.0045	0.0034	0.0048	0.0035	0.0036	0.0042	0.0040	0.0036
SRCNN	0.2613	0.2580	0.2550	0.2459	0.2519	0.2456	0.2492	0.2479	0.2498
Proposed	0.2747	0.2008	0.1690	0.1624	0.2261	0.2107	0.1670	0.2324	0.1697

4 结 论

针对 SRCNN 存在的不足进行了改进,搭建了基于双通道卷积神经网络的深度图超分辨率重建模型。模型由深、浅两个通道组成,分别侧重学习深度图的细节与轮廓特征,两者相互补充实现了深度图的重建。实验结果表明本文算法在测试集上较其他算法取得了更好的实验结果,尤其是在采样因子为 8 时仍能取得较好的重建效果。而且,本文算法在一定程度上克服了传统算法手工提取特征不准确的缺陷,本文算法实现了端到端的重建降低了测试的时间复杂度,使得模型更适合实际应用。后续应进一步研究深度图与一般自然图像的特点以搭建能同时适用于两类图像超分任务的模型。

参 考 文 献

[1] Kolb A, Barth E, Koch R, *et al.* Time-of-flight cameras in computer graphics [J]. Computer

Graphics Forum, 2010, 29(1): 141-159.

- [2] Liu J L, Li S M, Li Y D, *et al.* High-resolution depth maps based on TOF-Stereo fusion [J]. Information Technology, 2016, 40(12): 190-193.
刘娇丽, 李素梅, 李永达, 等. 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, 40(12): 190-193.
- [3] Schuon S, Theobalt C, Davis J, *et al.* High-quality scanning using time-of-flight depth superresolution [C] // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008: 1-7.
- [4] Wang Q L, Li J Y, Shen H K. Target tracking system of binocular vision and laser range sensor[J]. Acta Optica Sinica, 2016, 36(9): 0912002.
王琪龙, 李建勇, 沈海阔. 双目视觉-激光测距传感器目标跟踪系统[J]. 光学学报, 2016, 36(9): 0912002.
- [5] Li X, Tang R F, Li Z L, *et al.* Laser ranging data processing based on the analysis of the binary image [J]. Chinese Journal of Lasers, 2014, 41(12):

1208005.
李熙, 汤儒峰, 李祝莲, 等. 基于二值图像的卫星激光测距数据处理[J]. 中国激光, 2014, 41(12): 1208005.
- [6] Schuon S, Theobalt C, Davis J, *et al.* LidarBoost: Depth superresolution for ToF 3D shape scanning[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2009: 343-350.
- [7] Gevrekci M, Pakin K. Depth map super resolution [C] // Proceedings of 18th IEEE International Conference on Image Processing, 2011: 3449-3452.
- [8] Rajagopalan A N, Bhavsar A, Wallhoff F, *et al.* Resolution enhancement of PMD range maps[C] // Proceedings of the 30th DAGM Symposium on Pattern Recognition, 2008: 304-313.
- [9] Aodha O M, Campbell N D F, Nair A, *et al.* Patch based synthesis for single depth image super-resolution[C]// Proceedings of European Conference on Computer Vision, 2012: 71-84.
- [10] Timofte R, de Smet V, Gool L V. A+: Adjusted anchored neighborhood regression for fast super-resolution[C]// Proceedings of Asian Conference on Computer Vision, 2014: 111-126.
- [11] Kopf J, Cohen M F, Lischinski D, *et al.* Joint bilateral upsampling [J]. ACM Transactions on Graphics, 2007, 26(3): 96-101.
- [12] He K M, Sun J, Tang X O. Guided image filtering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6): 1397-1409.
- [13] Yang Q X, Yang R G, Davis J, *et al.* Spatial-depth super resolution for range images[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2007: 1-8.
- [14] Liu M Y, Tuzel O, Taguchi Y. Joint geodesic upsampling of depth images [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2013: 169-176.
- [15] Park J, Kim H, Tai Y W, *et al.* High quality depth map upsampling for 3D-TOF cameras [C] // Proceedings of International Conference on Computer Vision, 2011: 1623-1630.
- [16] Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one[J]. Multiscale Modeling & Simulation, 2005, 4(2): 490-530.
- [17] Buades A, Coll B, Morel J M. Imaged enoising methods. A new nonlocal principle [J]. SIAM Review, 2010, 52(1): 113-147.
- [18] Ferstl D, Reinbacher C, Ranftl R, *et al.* Image guided depth upsampling using anisotropic total generalized variation [C] // Proceedings of IEEE International Conference on Computer Vision, 2013: 993-1000.
- [19] Li Y J, Xue T F, Sun L F, *et al.* Joint example-based depth map super-resolution [C] // Proceedings of IEEE International Conference on Multimedia and Expo, 2012: 152-157.
- [20] Kwon H, Tai Y W, Lin S. Data-driven depth map refinement via multi-scale sparse representation[C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015: 159-167.
- [21] Kiechle M, Hawe S, Kleinstaub M. A joint intensity and depth co-sparse analysis model for depth map super-resolution [C] // Proceedings of IEEE International Conference on Computer Vision, 2013: 1545-1552.
- [22] Lu J J, Forsyth D. Sparse depth super resolution[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2245-2253.
- [23] Diebel J, Thrun S. An application of Markov random fields to range sensing [J]. Advances in Neural Information Processing Systems, 2005: 291-298.
- [24] Dong C, Loy C C, He K M, *et al.* Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [25] Wang Z W, Liu D, Yang J C, *et al.* Deep networks for image super-resolution with sparse prior [C] // Proceedings of IEEE International Conference on Computer Vision, 2015: 370-378.
- [26] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [27] 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.
- [28] Lu J B, Shi K Y, Min D B, *et al.* Cross-based local multipoint filtering [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2012: 430-437.