

基于多方向滤波的强边缘深度图像补全方法

吕 浩^{1,2} 陈世峰¹

¹(中国科学院深圳先进技术研究院 深圳 518055)

²(中国科学院大学深圳先进技术学院 深圳 518055)

摘 要 传统的深度相机能够获取像素级配准的深度和彩色图像,但所获取的深度图像存在明显的像素缺失。针对这一问题,文章提出了一种快速深度图像补全算法,能够有效地填充像素缺失区域并保持锐利的深度图像边缘特征。首先,设计出一种边缘蒙版,通过边缘蒙版对联合双边滤波器进行改进。其次,与传统滤波器算法不同,由于滤波器本身存在的方向特性,文章采用不同方向模拟了真实场景的遮挡情况。通过设定 4 个滤波方向,用改进后的联合双边滤波器对孔洞深度图进行修补填充,然后再通过马尔科夫随机场模型,将 4 个不同方向滤波器获得的深度填充图融合成一幅图像。最后,通过不同场景的深度图像进行实验。结果表明,所提出的深度图像补全算法显著优于传统方法。

关键词 图像补全; 深度图像增强; 边缘保持; 马尔科夫随机场

中图分类号 TG 156 **文献标志码** A

Strong Edge-Aware Depth Image Completion with Multi-Direction Filtering

LV Hao^{1,2} CHEN Shifeng¹

¹(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

²(Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract Conventional depth-camera can provide pixel-wise aligned depth and color images. However, the obtained depth image usually contains a lot of vacant image regions subject to the device resolution and reflectance property of target scene. To solve this problem, a novel depth image completion algorithm was investigated in this paper. To preserve sharp edges in the depth image, an edge mask was first designed. With reference to the edge mask, an improved joint bilateral filtering scheme was proposed. By filtering the depth image in four directions, a Markov random field model was used to combine the filtered depth images into one. Different from conventional filter-based image completion algorithms, the scene occlusion problem is also considered in the proposed algorithm. A variety of depth images are used in the experiment. Comparative results are presented to demonstrate the improvement over some classical methods.

Keywords depth completion; depth image enhancement; edge preserving; Markov random field

收稿日期: 2016-04-13 修回日期: 2016-05-24

基金项目: 深圳市科技计划基础科研项目(JCYJ20150401145529049)

作者简介: 吕浩, 硕士研究生, 研究方向为计算机视觉、图像处理; 陈世峰(通讯作者), 硕士生导师, 研究方向为计算机视觉、图像处理,
E-mail: shifeng.chen@siat.ac.cn.

1 引 言

近年来图像的采集技术迅猛发展, 深度图像的采集设备越来越多样化和平民化。目前, 商用的 TOF (Time of Flight) 摄像机^[1,2]能够以 100 帧每秒的速度采集 320×240 分辨率的深度图像, 而消费级的深度图像采集设备 Kinect V2 的深度图像分辨率则达到了 512×424 。随着消费级深度信息采集设备的出现, 深度图像的应用研究吸引了越来越多的研究人员, 尤其是在计算机视觉和图形学领域的应用。目前深度图像已经被应用在一些室内场景的任务中, 如三维场景重建^[3-5]、行为识别^[6]和三维物体识别^[7,8]。

虽然深度图像采集设备多种多样, 但采集到的深度图像质量却远未达到令人满意的程度。由于物体的相互遮挡、镜面反射或者暗表面、场景超出设备采集范围和设备自有的量化噪声等原因, 深度图像中存在着许多像素值缺失的区域和数值噪声。这些缺陷一定程度上限制了深度图像的应用。因此需要一个作为深度图像预处理补全算法。

在深度图像出现之前, 对彩色以及灰度图像的补全就已经有相应研究, 这个针对彩色或灰度数字图像的补全领域被称为图像补全 (Image Inpainting)。传统的图像补全是通过该幅图像的其他部分来填充缺失的像素。一般来说, 传统图像补全方法可以分成两种: 一种是基于图像纹理的填充算法^[9], 一种是基于图像几何结构的填充算法^[10,11]。

针对深度图像, 图像补全工作通常采用引导补全 (Guided Inpainting) 的方式。一般地, 深度图像采集系统在采集深度图像时会对应地采集相同场景的彩色图像。引导补全是利用高清晰度的彩色图像作为附加信息辅助深度图像补全的一类算法。这类算法是基于深度图像与彩色图像的结构一致性假设。本文目标是对深度图像的补全, 因

此采用的是引导补全的策略。

虽然深度图像的补全是一个相对新的领域, 已经有一些卓越的研究成果被发表。本节将介绍几种重要的深度图像补全算法。

在传统图像补全算法中, 快速行进算法 (Fast Marching Method)^[11,12]通过在图像梯度方向上传播局部的图像信息, 提高了运行速度。Bertalmio 等^[13]提出的同时补全图像结构和纹理的补全算法能够自动生成纹理, 从而使缺失区域更加合理、美观。基于样本的纹理生成补全算法^[9]通过合理利用物体边缘的梯度强度, 尝试解决 Bertalmio 等^[13]的算法不能很好地将强方向性的结构填充到缺失区域的缺点。连续的视频序列往往具有很强的时间连续性, Lai 等^[14]首先利用传统图像补全方法^[11]对单帧进行补全, 然后扩展了联合双边滤波器 (Joint Bilateral Filter), 加入了时间序列间连续的约束。

观察到深度图像及其对应的高质量彩色图像间有很强的一致性, 因此针对深度图像的补全算法通常利用彩色图像和原始深度图像作为输入, 来补全深度图像中的像素缺失区域。基于优化的算法通常基于彩色图像和深度图像边缘的一致性进行设计, 在数值效果上表现优异, 但很耗时, 而且在深度边缘会不同程度地产生模糊^[15-17]。由于深度图像的补全过程通常是作为更高级的视觉任务之前的预处理, 因此算法的时间效率很重要。基于滤波的算法在时间效率上取得了很好的性能, 但不能很好地处理较大的像素缺失区域, 其中有联合双边滤波器^[18-23]、拉普拉斯滤波器^[7]和模态滤波器^[24,25]。

除传统补全方法外, Yu 等^[26]也提出了基于三维估计的深度图像增强方法。Lu 等^[27]基于图像块组成的矩阵低秩的特点, 提出了数据驱动的深度图像补全方案。这些方法在补全效果上表现很好, 然而时间复杂度较高、内存需求大, 不适用于作为视觉任务的预处理算法。

本文主要考虑算法的时间效率,因此以速度较快且效果具有竞争力的联合双边滤波器^[18]为基础,利用 Canny 边缘检测子提取到的边缘图,提出了具有强边缘保持特性的边缘模板,并将其加入到滤波器中。观察到不同方向的填充会产生不同的效果,本文提出了一种基于马尔科夫随机场的快速融合方案,将 4 个不同方向的填充结果融合在一起,得到一个更好的填充结果。在本文中,原始深度图像记为 D_0 ,对应的彩色图像记为 C_0 。

算法在广泛使用的 Middlebury 数据集^[28]以及利用 Kinect V2 采集的数据上进行了验证,算法在时间效率上达到了目前最快,并且在数值和视觉效果上可以和目前最好的算法媲美,甚至在某些情况下超越目前最好的算法。

2 强边缘保持的深度补全算法

2.1 联合双边滤波器

原始的联合双边滤波器^[18]解决的是深度图像上采样的问题,本文将其用于深度图像的补全任务。标准的联合双边滤波器可以定义成如下形式:

$$\tilde{D}_p = \frac{1}{k_p} \sum_{q \in \Omega} D_q f(\|p-q\|) g(\|I_p - I_q\|) \quad (1)$$

其中, p 和 q 表示像素的坐标; \tilde{D} 是滤波后的深度图; D 是输入的深度图; I 是用作引导的图像,在这里即是与深度图对应的彩色图像; k_p 是归一化因子,即 $f \cdot g$ 滤波器的和; Ω 是滤波器的空域窗口; f 和 g 分别称作空间滤波器核 (Spatial Filter Kernel) 和范围滤波器核 (Range Filter Kernel), 这两者通常是高斯核。

标准的联合双边滤波器在处理大块深度像素缺失区域时效果很差,即便根据情况调整滤波器核的尺寸也不能得到令人满意的结果。因此本文扩展了标准的联合双边滤波器,使其支持原地像

素替换的滤波方式:

$$\tilde{D}_p^+ = \frac{1}{k_p} \sum_{q \in \Omega} \tilde{D}_q^- f(\|p-q\|) g(\|I_p - I_q\|) \quad (2)$$

其中, \tilde{D}_p^+ 是在填充了 p 点后的深度图; \tilde{D}_p^- 是在填充了 p 点前的深度图。初始状态时, \tilde{D}^- 被初始化成 D_0 。在获得一个新的像素值后, \tilde{D}^- 被更新成 \tilde{D}_p^+ , 再进行下一个像素的填充。扩展后的滤波器既能保持较好的边缘一致性,又能够在一定程度上处理较大的像素缺失区域。然而,此时的算法在处理边缘时仍然会产生模糊的情况,为解决这个问题,本文继而引入了 Canny 提取的边缘图像,并提出了边缘模板的概念,进一步改进滤波器。

2.2 引入边缘模板

好的深度图像补全结果应该在物体边缘边界锐利,然而联合双边滤波器在填充时会使像素溢出物体边界,特别是在物体表面与边界周围背景的颜色相近时。原因是联合双边滤波器中的范围项 (Range Term) 仅度量了像素间的颜色差异,这会在滤波时将边界两边的未缺失的像素都考虑在内。

考虑到联合双边滤波器的缺点,本文决定利用 Canny 边缘提取器首先提取彩色图像中的物体边界。在对原始深度图进行填充前,算法首先从 C_0 提取出边缘图像 E_0 。然后,算法基于 E_0 对每个待填充的像素定义一个边缘蒙版 (Edge Mask)。这些边缘蒙版能够在滤波时使越过边缘部分的像素失效,从而防止深度像素值越过物体边界溢到另一边,进而保持了锐利的物体边缘。

本文根据物体遮挡光线传播路径的情形来设计边缘蒙版。如图 1 所示,这是一个 7×7 像素的滤波器窗口,记中央像素为 D_p ,也就是将被填充的像素。假设根据边缘图像 E_0 ,红色圆点是滤波器的当前窗口中心,黄色、橙色和绿色圆点所在的像素是边缘像素。那么由于这三个像素的遮挡,它们后面的像素及其本身便不能够对当前

需填充的像素起作用。这些像素在图 1 中用遮挡住它们的像素的颜色标示出来了。这些像素应该在 p 点滤波时被排除在外。因此, 对于 p 点, 边缘蒙版可以定义成:

$$M_{q \in \Omega} = \begin{cases} 0, & \text{if } \tilde{D}_q^- \in \Lambda_p \\ 1, & \text{其他} \end{cases} \quad (3)$$

其中, Ω 是滤波器的空间窗口; Λ_p 是 p 点为中心的窗口中所有的边缘像素及被边缘像素遮挡的像素的集合。

利用这个边缘蒙版, 本文进一步扩展联合双边滤波器

$$\tilde{D}_p^+ = \frac{1}{k_p} \sum_{q \in \Omega} \tilde{D}_q^- M_q f(\|p-q\|) g(\|I_p - I_q\|) \quad (4)$$

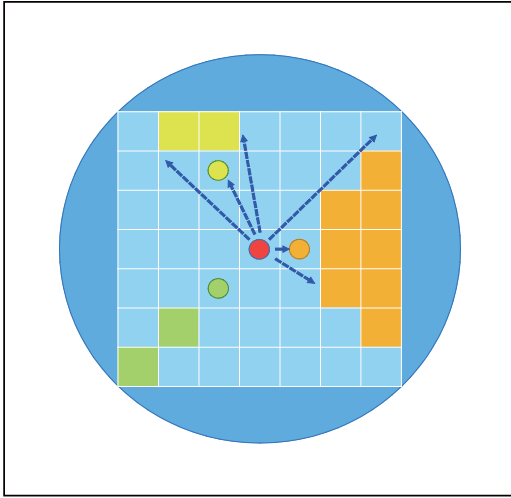


图 1 边缘模板示意图

Fig. 1 Edge mask map

2.3 基于马尔科夫随机场的深度图像融合

2.3.1 生成备选深度图

如前所述, 在不同方向滤波生成不同的结果, 本文沿着 4 个方向分别对原始深度图进行填充: 左上到右下、左下到右上、右上到左下和右下到左上, 生成四幅深度图像。这 4 幅深度图像质量上各有优劣, 下一步算法将通过马尔科夫随机场模型把这 4 幅深度图像融合成一幅图像, 即

最终结果。

2.3.2 融合备选深度图

本文将融合问题形式化为图像的标注问题 (Image Labeling Problem)。每个像素将会被赋予一个标号, 用于标识该像素是属于 4 幅备选图像中的哪一幅。对于马尔科夫随机场, 能量函数通常包含两项: 数据项和平滑项

$$E = \underbrace{\sum_p d_p(l_p)}_{E_d} + \lambda \underbrace{\sum_{\{p,q\} \in N} V_{pq}(l_p, l_q)}_{E_s} \quad (5)$$

其中, l_p 是赋给 p 点的标签值; d_p 是将标签 l_p 赋值给 p 点的代价函数; N 是 8 邻域像素对的集合; V_{pq} 是将标签 l_p 赋值给 p 点并且将 l_q 赋值给 q 点的代价函数; λ 是权重参数。

为了使融合后的深度图像更加合理, 本文定义数据项

$$d_p(l_p) = \frac{1}{c_p + 1} e^{-\frac{d_{l_p}}{\sigma^2}} \sum_{q \in \Omega} T_q \quad (6)$$

其中, c_p 是 T 中非零元素的个数; d_l 是像素 p 在滤波方向上到该行开始的距离; σ 是尺度因子 (在本文中为 10)。本文使用

$$T = 1 - G(E_0) \quad (7)$$

对 D_0 中像素的置信度进行建模, 用于模拟像素值传播过程中可信度的变化。这里 G 是高斯模糊算子。

本文使用空域可变的平滑项

$$V_{pq}(l_p, l_q) = w_{pq} v_{pq}(l_p, l_q) \quad (8)$$

其中,

$$v_{pq}(l_p, l_q) = \begin{cases} 0, & \text{if } l_p = l_q \\ C, & \text{其他} \end{cases} \quad (9)$$

$$w_{pq} = \begin{cases} H_p, & \text{如果 } p, q \text{ 是水平相邻} \\ V_p, & \text{如果 } p, q \text{ 是垂直相邻} \end{cases} \quad (10)$$

$$H_p = \begin{cases} a, & \text{if } h_p < t_1 \\ a - \beta h_p, & \text{if } t_1 \leq h_p \leq t_2 \\ b, & \text{if } h_p > t_2 \end{cases} \quad (11)$$

这里 C 是常数; h 是 C_0 在水平方向的归一化梯度图; t_1 和 t_2 是两个阈值; β 是尺度因子。相似地, V 是定义在垂直方向上。具体的参数设置将会在第三节给出。

最终的深度图是根据每个像素所属的备选深度图, 将对应的像素拷贝到结果图像中。马尔科夫随机场模型从备选中选择了最合适的结果, 保证了补全区域的平滑性, 充分利用了备选图像的优点。

3 实验

3.1 数据集

本文在两个数据集上进行了测试, 分别是 Middlebury 数据集^[28]和在实验室采集的 Kinect V2 数据集。图 2 中左边两幅是 Kinect V2 数据

集中的彩色图像和对应深度图像, 右边两幅是来自 Middlebury 数据集中的彩色图像和对应深度图像。下面将介绍这两个数据集。

3.1.1 Middlebury 数据集

Middlebury 数据集是一个通过立体视觉算法建立的数据集, 提供了相同尺寸的彩色图像和深度图像对, 用于验证立体视觉算法的有效性。该数据集也常被用于深度图像上采样及补全的研究中, 为了和其他算法比较, 本文基于此数据库对比测试了算法的时间效率和视觉效果, 并用均方根误差 (Root-Mean-Square Error, RMSE) 指标对比了几种目前最好的算法的补全效果。

3.1.2 Kinect V2 数据集

随着硬件设备的发展, 深度图像的质量有了显著的提升。为了测试算法在最新的深度图像采集设备上的效果, 本文在实验室中搭建了较为复杂的环境, 并利用最新的 Kinect V2 采集了 100 对彩色-深度图像。图 3 展示了其中部分数据。由于缺乏可靠的真值, Kinect V2 数据集仅用作视觉效果比较。



图 2 彩色图像及对应的深度图像

Fig. 2 Color image and its corresponding depth map

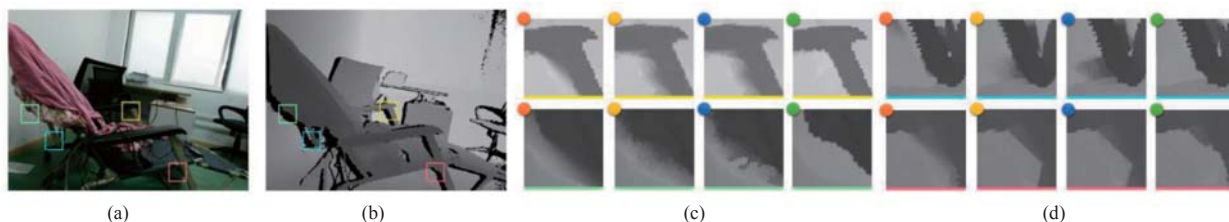


图 3 在 Kinect V2 数据集上的结果 (橙色、黄色、蓝色、绿色圆标分别表示 JBF、Levin、Park 和本文算法的结果)

Fig. 3 Results on Kinect V2 dataset (Orange, yellow, blue and green round points represent the result of JBF, Levin, Park and our method)

3.2 数据配准

Middlebury 数据集提供了天然配准的图像对。对于 Kinect V2 的数据,本文利用 Kinect SDK V2 中提供的相应函数将深度图转换到彩色图的坐标系下,完成配准工作。

3.3 实验过程

本文分别在数值上和视觉效果上与目前最好的深度图像补全算法进行比较。

在数值上,本文以广泛应用的均方根误差(RMSE)为标准,在 Middlebury 数据集上进行比较。由于缺乏标准的真实值图像(Ground Truth),本文通过随机去掉深度图像物体边缘的一些像素来模拟 Kinect 数据的真实缺陷。在计算 RMSE 时,原始图像中的像素缺失部分不计算在内(因为缺少真实值)。本文在具有大面积像素缺失和小面积像素缺失的数据样本上都进行了实验。

在视觉上,本文在实验室采集的 Kinect V2 数据集上进行了实验,对比了 JBF^[18]、Levin^[29]、Park^[17]和本文算法的效果。为便于比较,本文特别地截取了几处关键的被填充区域,在实验结果一节将会看到不同的填充效果。

所有的实验都在同一台拥有 4G 内存和英特尔 I7 CPU 的笔记本上进行。

4 实验结果

4.1 量化结果比较

表 1 为运行时间比较结果,表 2 为 RMSE 指标下不同算法的性能比较结果。从表 2 可以看出,在大部分情形下,本文算法达到了最低的 RMSE。但从表 1 的时间效率比较可以看出,在 RMSE 指标下,本文的算法能够取得与其他算法相似或更好的效果,而且速度上比其他算法快一个数量级。

4.2 视觉效果比较

图 3 和图 4 分别比较了算法在 Kinect V2 数

表 1 运行时间比较

Table 1 Algorithm runtime measured

图片名称	算法运行时间 (s)			
	JBF	Levin	Park	本文算法
Aloe	4.455	9.251	15.974	0.075
Baby3	4.501	9.077	15.797	0.172
Bowling1	4.282	8.619	14.977	0.077
Cones	4.713	9.496	16.317	0.072
Lampshade1	4.450	9.096	15.686	0.118
Reindeer	4.644	9.988	16.770	0.032
Rocks1	4.403	8.875	15.341	0.118
Teddy	4.767	9.391	16.308	0.057

表 2 RMSE 指标下的算法结果比较

Table 2 Quantity compare in RMSE

图片名称	RMSE			
	JBF	Levin	Park	本文算法
Aloe	14.772 0	13.994 0	13.877 7	13.800 9
Baby3	27.409 5	27.660 5	27.265 0	27.477 9
Bowling1	8.598 5	14.211 3	15.858 4	16.551 9
Cones	18.547 2	18.357 7	18.107 8	18.100 9
Lampshade1	11.773 6	16.797 0	16.538 4	10.739 4
Reindeer	10.019 6	9.237 2	8.978 8	8.593 1
Rocks1	18.868 7	20.159 3	20.264 6	20.888 4
Teddy	19.031 5	19.010 8	18.876 5	18.734 8

据集部分数据、Middlebury 数据集和 Kinect V2 数据集部分数据上的效果。综合图 3 和图 4 可看出, JBF 能够保持锐利的边缘,但有时会产生粗糙的区域; Levin 和 Park 在待填充区域较小时给出了比 JBF 更好的结果,但在处理较大待填充区域时这两者都会产生模糊;而本文算法不仅生成了非常锐利的边缘(图 4(j)),而且处理在较大待填充区域时,依然能够在马尔科夫随机场模型的作用下保持原有物体的结构(图 4(t))。由此可知,边缘蒙版能够有效地保持锐利边缘,融合处理则能够有效地从备选深度图像中选择出更好的结果。

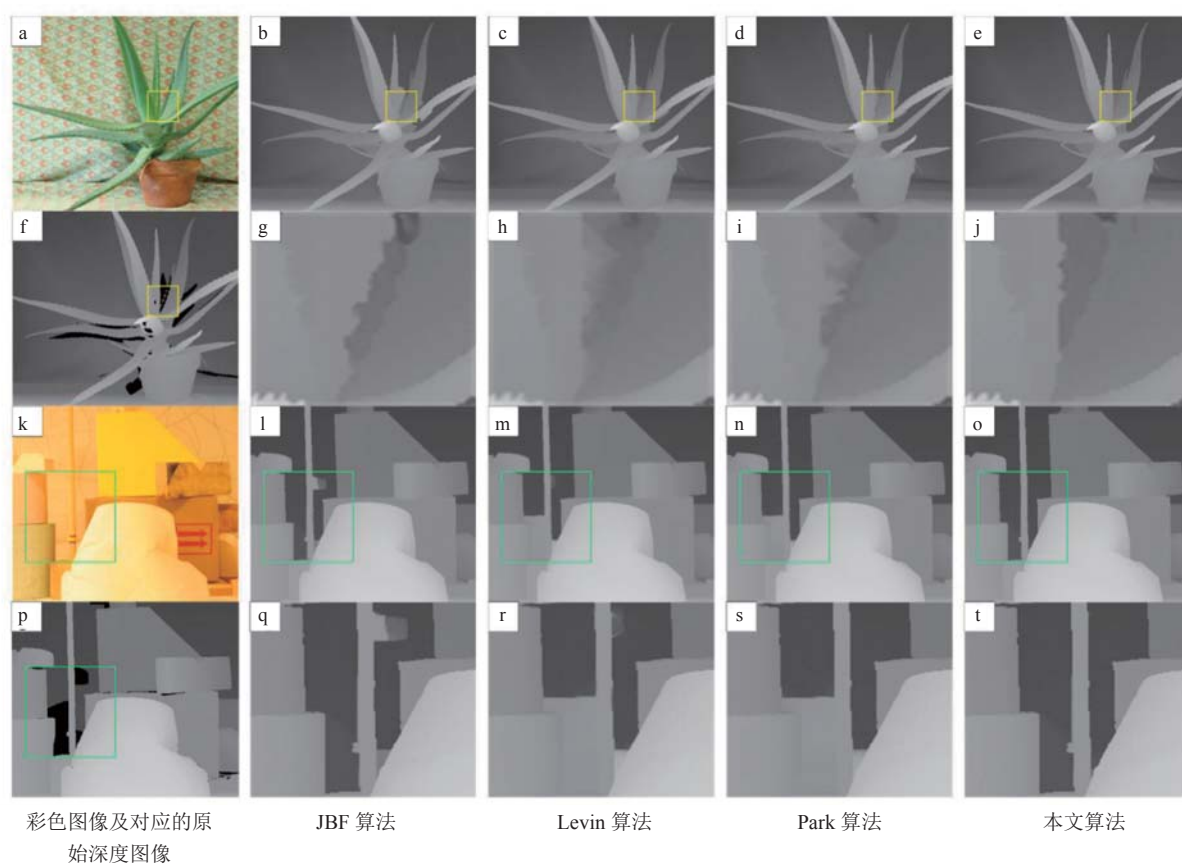


图 4 在 Middlebury 数据集上的结果

Fig. 4 Results on Middlebury dataset

5 讨论

深度图像的补全任务通常作为更高级的视觉算法的预处理，因此时间效率和空间效率是本文算法的主要考虑方面，与此同时又要在数值和视觉效果上能够与当前最好的算法媲美。

基于全局优化的补全算法^[15-17]往往能够得到较好的结果，但计算复杂度较高、运行时间较长。注意到深度图像中实际上只有部分的像素能够对补全区域有影响，因此本文仅通过利用待填充区域邻域像素来做填充，大幅降低了算法处理的数据量，速度提升在图像质量较好时尤为明显。算法主要采用计算复杂度较低的联合双边滤波器，因此整体的计算复杂度低。传统的图像补全算法注重像素填充顺序，而这一过程较为耗

时。本文的算法预设了 4 个固定的方向，避免了确定每个像素填充顺序的问题，从而大幅提升了速度。

本文假设彩色图像与深度图像在结构上具有一致性，在填充和融合两个步骤中都充分利用了此性质，因此实验结果在边缘处特别锐利。

正如所有其他引导补全的算法一样，在彩色图像与深度图像的一致性假设不成立的局部区域，算法会产生不好的结果。然而，在真实情景下这种情况比较少见，因此不作为算法设计的重点考虑。

6 结论

本文提出了一种快速的深度图像补全算法，通过利用高质量的彩色图像作为引导图像，填充

深度图像中像素缺失的区域, 算法在取得具有竞争力的视觉效果同时, 具有非常高的时间效率和空间效率。

算法在 4 个预定的方向上对原始深度图分别进行填充, 然后通过马尔科夫随机场模型将 4 幅深度备选图融合成最终的结果。算法充分利用了引导图像的边缘信息, 因此在填充区域保持了锐利的深度边缘。算法适合作为原始深度图的后期处理。我们将进一步利用 GP-GPU 加速本文的算法, 使之用于实时的深度图像补全任务。

参 考 文 献

- [1] Hansard M, Lee S, Choi O, et al. Time-of-Flight Cameras: Principles, Methods and Applications [M]. Springer, 2012.
- [2] Heptagon. SwissRanger [OL]. [2016-04-06]. <http://hptg.com/industrial/>.
- [3] Izadi S, Kim D, Hilliges O, et al. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera [C] // Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, 2011: 559-568.
- [4] Ren XF, Bo LF, Fox D. RGB-(D) scene labeling: features and algorithms [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012: 2759-2766.
- [5] Henry P, Krainin M, Herbst E, et al. RGB-D mapping: using Kinect-style depth cameras for dense 3D modeling of indoor environments [J]. The International Journal of Robotics Research, 2012, 31(5): 647-663.
- [6] Lei JN, Ren XF, Fox D. Fine-grained kitchen activity recognition using RGB-D [C] // The Proceedings of the 2012 ACM Conference on Ubiquitous Computing, 2012: 208-211.
- [7] Janoch J, Karayev S, Jia YQ, et al. A category-level 3D object dataset: putting the Kinect to work [C] // 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 2011: 1168-1174.
- [8] Wang NB, Gong XJ, Liu JL. A new depth descriptor for pedestrian detection in RGB-D images [C] // 2012 21st International Conference on Pattern Recognition (ICPR), 2012: 3688-3691.
- [9] Arias P, Facciolo G, Caselles V, et al. A variational framework for exemplar-based image inpainting [J]. International Journal of Computer Vision, 2011, 93(3): 319-347.
- [10] Masnou S. Disocclusion: a variational approach using level lines [J]. IEEE Transactions on Image Processing, 2002, 11(2): 68-76.
- [11] Telea A. An image inpainting technique based on the fast marching method [J]. Journal of Graphics Tools, 2004, 9(1): 23-34.
- [12] Gong XJ, Liu JY, Zhou WH, et al. Guided depth enhancement via a fast marching method [J]. Image and Vision Computing, 2013, 31(10): 695-703.
- [13] Bertalmio M, Vese L, Sapiro G, et al. Simultaneous structure and texture image inpainting [J]. IEEE Transactions on Image Processing, 2003, 12(8): 882-889.
- [14] Lai K, Bo LF, Ren XF, et al. A large-scale hierarchical multi-view RGB-D object dataset [C] // 2011 IEEE International Conference on Robotics and Automation (ICRA), 2011: 1817-1824.
- [15] Diebel J, Thrun S. An application of Markov random fields to range sensing [C] // Advances in Neural Information Processing Systems, 2005: 291-298.
- [16] Zhu JJ, Wang L, Gao JZ, et al. Spatial-temporal fusion for high accuracy depth maps using dynamic MRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(5): 899-909.
- [17] Park J, Kim H, Tai YW, et al. High-quality depth map upsampling and completion for RGB-D cameras [J]. IEEE Transactions on Image Processing, 2014, 23(12): 5559-5572.
- [18] Kopf J, Cohen MF, Lischinski D, et al. Joint bilateral upsampling [J]. ACM Transactions on Graphics, 2007, 26(3): 96.
- [19] Tomasi C, Manduchi R. Bilateral filtering for gray and color images [C] // Proceedings of the Sixth International Conference on Computer Vision, 1998: 839-846.

- [20] Wasza J, Bauer S, Hornegger J. Real-time pre-processing for dense 3D range imaging on the GPU: defect interpolation, bilateral temporal averaging and guided filtering [C] // 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 2011: 1221-1227.
- [21] Kim SY, Cho JH, Koschan A, et al. Spatial and temporal enhancement of depth images captured by a time-of-flight depth sensor [C] // 2010 20th International Conference on Pattern Recognition (ICPR), 2010: 2358-2361.
- [22] Richardt C, Stoll C, Dodgson NA, et al. Coherent spatiotemporal filtering, upsampling and rendering of RGBZ videos [J]. *Computer Graphics Forum*, 2012, 31(2pt1): 247-256.
- [23] Lazcano V, Arias P, Facciolo G, et al. A gradient based neighborhood filter for disparity interpolation [C] // 2012 19th IEEE International Conference on Image Processing (ICIP), 2012: 873-876.
- [24] Min D, Lu J, Do MN. Depth video enhancement based on weighted mode filtering [J]. *IEEE Transactions on Image Processing*, 2012, 21(3): 1176-1190.
- [25] Parker R. Kinect depth inpainting and filtering [OL]. [2016-04-06]. <http://www.radfordparker.com/papers/kinectinpainting.pdf>.
- [26] Yu LF, Yeung SK, Tai YW, et al. Shading-based shape refinement of RGB-D images [C] // 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013: 1415-1422.
- [27] Lu S, Ren XF, Liu F. Depth Enhancement via low-rank matrix completion [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 3390-3397.
- [28] Daniel Scharstein and Richard Szeliski. Middlebury stereo datasets [OL]. [2016-04-06]. <http://vision.middlebury.edu/stereo/data/>.
- [29] Levin A, Lischinski D, Weiss Y. Colorization using optimization [J]. *ACM Transaction on Graphics*, 2004, 23(4): 689-694.