

基于跨尺度引导图像滤波的稠密立体匹配

刘杰, 张建勋, 代煜, 苏赫

南开大学机器人与信息自动化研究所, 天津 300071

摘要 针对现有局部立体匹配算法在弱纹理表面、深度不连续处等特定区域匹配精度低、实时性难以满足要求等问题, 提出了一种基于跨尺度引导图像滤波的稠密立体匹配算法。利用图像分割技术对立体图像进行预分割, 得到分割区域内像素的聚合半径; 以此半径为指导, 在立体图代价空间中以 3 种不同尺寸的核进行滤波, 引入正则化项确保聚合代价的一致性, 以得到更有效的聚合代价; 运用简单高效的贪心策略获取初步视差。基于 Middlebury 测试平台的实验结果表明所提算法兼具实时性和高效性。

关键词 机器视觉; 信号处理; 引导滤波; 立体匹配; 正则化; 跨尺度; 聚合代价

中图分类号 TN911.73

文献标识码 A

doi: 10.3788/AOS201838.0115004

Dense Stereo Matching Based on Cross-Scale Guided Image Filtering

Liu Jie, Zhang Jianxun, Dai Yu, Su He

Institute of Robotics & Automatic Information System, Nankai University Tianjin 300071, China

Abstract To solve problems of the difficulty to meet the real-time requirements and the low matching accuracy of existing local stereo matching algorithms at some special regions, such as weak textured surfaces and the discontinuity boundary of depth, a dense stereo matching algorithm based on cross-scale guided image filtering is proposed. An image segmentation technology is used to realize pre-segmentation of stereo images and the aggregation radius of pixels in the segmented region is obtained. This radius is used as a guide, and kernels with three different sizes are used to carry out filtering in the cost space of stereo image. The regularization term is introduced to ensure the consistency of the aggregated cost, so as to obtain a more efficient aggregate cost. A simple and efficient winner-take-all strategy is used to obtain the initial disparity. The experimental results based on Middlebury test bench show that the proposed algorithm has both real time capability and high efficiency.

Key words machine vision; signal processing; guided filtering; stereo matching; regularization; cross-scale; aggregate cost

OCIS codes 150.1135; 330.1400; 100.6890

1 引 言

在计算机视觉和图形学中,超像素、降噪等都可以视为一个病态的逆问题。在多目视觉立体匹配中同样面临着病态问题的挑战。噪声、抖动、角度、畸变等因素的影响,使得精确得到立体图像间像素级的匹配关系和深度信息成为一个极具挑战性的问题。Scharstein 等^[1]将立体匹配问题分为匹配代价计算、聚合代价、初步视差计算和后处理 4 个步骤,

将相关匹配算法分为全局和局部立体匹配算法。全局立体匹配算法的主要思想是将立体匹配问题形式化为能量函数,通过最小化能量函数得到匹配结果,其特点是精度较高,但是计算复杂度高、实时性差。常见的全局匹配算法有图割^[2]和动态规划^[3]等。局部立体匹配算法的主要思想是通过支持窗口进行聚合代价,进而计算视差,其匹配精度相对较差,但模型复杂度较低,易于实现,因此应用广泛。

近年来,基于机器学习的立体匹配算法^[4-5]获得

收稿日期: 2017-07-19; **收到修改稿日期:** 2017-08-29

基金项目: 国家自然科学基金(51375494,61403212)、国家重点研发计划(2017YFC0110402)

作者简介: 刘杰(1990—),男,博士研究生,主要从事计算机视觉、立体匹配、机器学习方面的研究。

E-mail: liu_j@mail.nankai.edu.cn

导师简介: 张建勋(1961—),男,博士,教授,博士生导师,主要从事医疗图像处理、智能机器人控制、计算机网络与信息系

统方面的研究。E-mail: zhangjx@nankai.edu.cn

了较高的匹配精度,该算法以深度神经网络算法为基础,自主学习模型中的参数。但由于模型可解释性差,在现阶段高精度立体匹配图像数据较少的情况下难以有效验证其泛化能力,限制了其在现实系统中的应用。

现有的局部立体匹配算法一般基于自适应支持窗口^[6]和自适应支持权重^[7],利用局部支持区域和支持权重的方法来增强中心像素的可辨别性。Shi等^[8]基于分割来获取像素的支持区域,但随后的多特征融合策略更多依赖于经验设定,泛化能力较差;Yang^[9]基于最小生成树将聚合代价形式化为权值滤波,将支持区域扩展到全图,但是其计算复杂度较高;Rhemann等^[10]将聚合代价视为在代价空间中进行滤波,利用引导图像滤波器^[11]有效提升了聚合代价的效率;Zhang等^[12]对立体图进行下采样,利用多尺度信息得到了高效的聚合代价,提出了一个多尺度立体匹配框架。但是下采样图不可避免地丢失了很多细节信息,且理论上后期难以弥补框架本身的固有问题;祝世平等^[13]改进了 Census 变换策略,有效提升了单像素匹配代价的可靠性,并改进了动态规划算法,有效提升了匹配精度。

综上所述,基于图像滤波的立体匹配^[10,12]算法具有实时性好、算法匹配精度高的优点,但是现有算法难以有效解决支持窗口的选择问题,为了有效利用现有的快速求和技术(如积分图^[14]和盒滤波器^[15]等),图像滤波基本上采用以中心像素为核心的固定尺寸矩形窗口来进行聚合代价,很大程度上限制了支持区域的表达能力。事实上,中心像素本身的支持窗口可为任意形状和尺寸。为了解决上述问题,本文设计了一种基于跨尺度图像引导滤波的聚合代价策略,首先用 Mean-Shift^[16]图像分割技术对立体图像进行预分割,获取像素的分割区域半径;以此区域半径作为指导,在代价空间中以 3 种不同尺寸的滤波核进行滤波,得到对应的代价空间;引入正则化项确保聚合代价的一致性,以得到更有效的聚合代价;以贪心(WTA)策略进行视差选择以验证聚合代价的可靠性。

2 跨尺度引导图像滤波立体匹配

立体匹配是多目视觉的关键,旨在获取多目立体图像间像素级的对应关系。聚合代价可以视为在代价空间中进行滤波,本节从尺度空间方面探讨了现有局部立体匹配方案的特点,基于图像滤波器^[7,11]提出了一种跨尺度图像滤波立体匹配方案。

2.1 立体匹配

立体匹配可以分为匹配代价计算、聚合代价、初始视差计算和视差精调 4 个步骤^[1]。代价计算主要是为了计算立体图像单像素间的匹配代价,常用的方法有绝对亮度差、归一化互相关和普查变换等。采用目前先进匹配算法中常用的梯度和强度进行单像素的匹配代价计算^[10,12],则像素 i 在视差为 d 时的单像素匹配代价可表示为

$$C(i, d) = \alpha \cdot \min(\|I_1(i) - I_r(i_d)\|, T_1) + (1 - \alpha) \cdot \min(\|\nabla_x I_1(i) - \nabla_x I_r(i_d)\|, T_2), \quad (1)$$

式中 $I_1(i)$ 和 $I_r(i_d)$ 分别表示参考图上中心像素和视差为 d 时该像素在目标图上对应像素的强度, $\nabla_x I_1(i)$ 和 $\nabla_x I_r(i_d)$ 分别表示参考图上中心像素和视差为 d 时该像素在目标图上对应像素的梯度, α 用来平衡颜色强度和梯度项, T_1 、 T_2 为截断值。

聚合代价即通过一定规则将得到的局部区域中的匹配代价进行累加聚合,如基于中心像素邻域特性的自适应支持权重聚合^[7]。一定程度上,局部区域的聚合代价相当于在代价空间中进行滤波,即

$$\tilde{C}(i, d) = \frac{1}{Z_i} \sum_{j \in N_i} K(i, j) C(i, d), \quad (2)$$

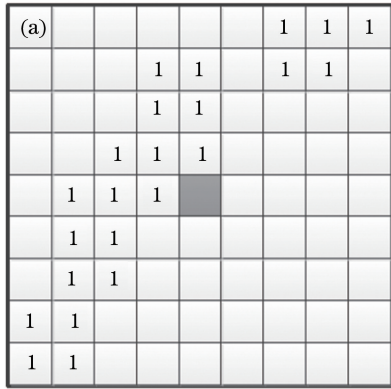
式中 $\tilde{C}(i, d)$ 为中心像素 i 在视差为 d 时的聚合代价, $K(i, j)$ 为其邻域像素 j 的支持权重, N_i 为中心像素 i 的支持窗口, $Z_i = \sum_{j \in N_i} K(i, j)$ 为归一化常数。

初始视差计算环节利用聚合代价按照一定的方法构造目标函数,以得到初始视差图,经典算法有 WTA 算法和动态规划等。视差精调的作用主要包含降低初始视差阶段的误匹配率和按照一定规则计算亚像素级精度的视差两部分。

2.2 引导图像滤波

图像滤波器具备抑制噪声和抽取图像结构信息的特性。在解决立体匹配问题时,需要得到基于中心像素的支持窗口,支持窗口内的像素需要尽可能地与中心像素具备相近的视差(即平滑性),并摒弃与之距离相近且视差相差较大的相邻像素,这就使得聚合代价的计算与图像滤波具备一定的共通性,如文献^[10]中将聚合代价视为在代价空间中进行联合滤波。从滤波的角度来进行聚合代价,为立体视觉的研究指引了新的方向。由于图像滤波自身的高效性,用图像滤波器在代价空间中进行滤波以完成聚合代价受到了越来越多研究者的关注。但是,图像滤波的高效性使得滤波核形状局限在矩形上,甚

至是固定尺寸的矩形上。在矩形区域内进行权值分配,一定程度上意味着可以获取自适应形状支持区域(有最大尺度限定),如图 1 所示。其中深灰色像素为中心像素 i ,外层黑框表示支持区域。图 1(a)



中,1 表示该像素的支持像素权值,其余像素权值为 0;图 1(b)中,浅灰色像素为中心像素 i 的支持像素 p ,灰色像素组成了中心像素 i 的支持区域 R ,支持区域并不具备连通性。

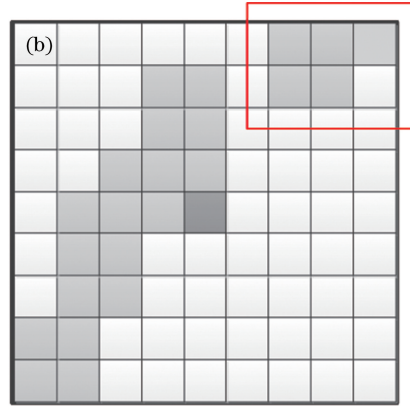


图 1 (a) 支持区域内有效像素的支持权值;(b)对应的自适应支持区域

Fig. 1 (a) Support weights of effective pixels in support regions; (b) corresponding adaptive support regions

需要说明的是,基于图像滤波的聚合代价^[10,12]采用了文献[7,11]中设置最大支持窗口尺寸的方案,而现实中难以预知像素的最大支持区域尺寸,因此在原始图上以固定尺寸的核进行滤波来完成聚合代价的思路,忽略了一部分支持像素,从而降低了聚合代价的有效性和稳定性。此外,支持区域不具备连通性,如图 1(b)右上角的红框区域,限制了其有效性和稳定性。因此,直接设置过大的核尺寸 r 可能会包含更多噪声,不利于代价的有效聚合。需要注意的是,这与完全自适应尺寸与形状的支持区域^[17]有着本质的区别。

2.3 跨尺度引导图像滤波

以优化的角度可以将聚合代价视为加权最小二乘问题^[18],即

$$\tilde{C}(i, d) = \operatorname{argmin}_z \frac{1}{Z_i} \sum_{j \in N_i} K(i, j) \| z - C(j, d) \|^2, \quad (3)$$

式中 z 为待计算的聚合代价。为了利用多尺度信息,文献[12]对原始立体图进行了下采样,并在下采样图中进行滤波(聚合代价),即

$$\tilde{C}(i^s, d^s) = \operatorname{argmin}_{\{z^s\}_{s=0}^S} \left(\sum_{s=0}^S \frac{1}{Z_i^s} \sum_{j^s \in N_i^s} K(i^s, j^s) \| z^s - C^s(j^s, d^s) \|^2 \right), \quad (4)$$

式中上标 s 为下采样尺度空间, S 为下采样的最大次数。某种程度上,这样的处理可以视为在原图尺寸上增大了支持窗口的最大尺寸,从而使得支持窗口包含尽可能多的有效像素;此外,增加的时间复杂

度较小,使系统仍然可以满足实时性的要求。但是,下采样图丢失了很多细节信息,使聚合代价的可靠性有所下降。在下采样的立体图像(即粗糙尺度)上进行视差计算虽然有效减小了视差搜索空间,但是容易导致有效信息的丢失,且理论上无法靠后续步骤来弥补,因此下采样层数只能设定为常数,不能由立体图的特性自适应地决定。上述问题都限制了聚合代价的有效性,最终体现为匹配精度的损失。

为了避免上述问题,采用 Mean-Shift 算法^[16]对立体图像进行预分割,得到一系列分割区域 R_s , 由此得到给定像素 i_s 的分割区域 R_s 的半径 r_s 。根据指导半径 r_s 计算立体图像内每个像素对应的两个聚合半径 r_{\min} 和 r_{\max} , 即

$$r_{\min} = \begin{cases} r_c^{\min}, & r_s \cdot \operatorname{floor}(r_s / r_c^{\min}) < r_c^{\min} \\ r_s \cdot \operatorname{floor}(r_s / r_c^{\min}), & \text{else} \end{cases}, \quad (5)$$

$$r_{\max} = \begin{cases} r_c^{\max}, & r_s \cdot \operatorname{ceil}(r_s / r_c^{\min}) > r_c^{\max} \\ r_s \cdot \operatorname{ceil}(r_s / r_c^{\min}), & \text{else} \end{cases}, \quad (6)$$

式中 $\operatorname{floor}(\cdot)$ 和 $\operatorname{ceil}(\cdot)$ 分别表示向下和向上取整; r_c^{\min} 和 r_c^{\max} 分别表示预设滤波核的最小、最大尺寸,后者是前者的整数倍。设置滤波核的尺寸范围,一方面是为了使图像滤波器保持一定尺寸,具备足够的支持信息;另一方面是为了防止因核尺寸过大导致更多噪声。引导滤波器(GF)每次只能采用一个固定尺寸的核在代价空间中进行滤波,以完成聚合代价,而以上计算得到的两个聚合半径是动态变

化的。如果精确计算每个像素的聚合代价,则需要大量重复的图像滤波进行聚合代价,这增加了算法复杂度,也计算了大量无用的聚合代价。为了利用图像滤波器每次只能用一个特定尺寸滤波核进行聚合代价的特性,尽量降低算法的时间复杂度,同时得到可靠的聚合代价,统计图像内像素的 3 个半径 r_c^{\min} 、 r_{\min} 和 r_{\max} 各自在不同量级上的频率,以出现频率最高的半径作为滤波核的最终尺寸 \tilde{r}_c^{\min} 、 \tilde{r}_{\min} 和 \tilde{r}_{\max} 。对核尺寸为 \tilde{r}_c^{\min} 、 \tilde{r}_{\min} 和 \tilde{r}_{\max} 的图像滤波器在原尺寸立体图的代价空间中进行滤波,得到 3 个不同的聚合代价空间。用 WTA 策略在 Middlebury^[19] 测试平台内的立体图像 Teddy 上计

算初始视差图。图 2 为不同尺寸滤波核得到的初始视差图,由图 2(a)~(c)可见,这 3 种不同尺寸的滤波器得到的聚合代价已经具备相当程度的可靠性,说明以出现频率最高的聚合半径代替每个像素的聚合半径是可行的。仔细分析视差图可知,无论是整体区域还是特殊区域(物体边界、弱纹理区域等),都得到了较好的匹配精度。另外,经过观察可以得出:核尺寸越大,视差图越平滑。但是随之而来的是部分细节的丢失。

为了更好地利用小尺寸下的细节信息和较大核尺寸下的平滑度,引入了正则化项,则加入正则项后的聚合代价为

$$\hat{C}(i, d) = \operatorname{argmin}_{\{z^l\}_{l=\tilde{r}_c^{\min}}^{\tilde{r}_{\max}}} \left[\sum_l \frac{1}{Z^l} \sum_{j \in N^l} K^l(i, j) \|z^l - C^l(j, d)\|^2 + \lambda \sum_l (\|z^{l_2} - z^{l_1}\|^2 + \|z^{l_3} - z^{l_2}\|^2) \right], \quad (7)$$

式中上标 l 表示不同尺寸的滤波核, l_1 、 l_2 、 l_3 分别表示核尺寸为 \tilde{r}_c^{\min} 、 \tilde{r}_{\min} 、 \tilde{r}_{\max} 。该算法加入正则化项,可以强制不同尺寸滤波器核得到的聚合代价保持较强的一致性,使不同代价空间的差距降低,确保了滤波器在代价空间中进行聚合代价的有效性;也

使得跨尺度图像引导滤波得到的聚合代价最大可能地利用了矩形核内像素的支持信息,从而得到更有效的聚合代价信息,进一步提升匹配精度。由图 2(d)可以看出, $\hat{C}(i, d)$ 的可靠性相对 $\tilde{C}(i, d)$ 要更高,保留了较多细节和视差的层次感,同时边缘足够平滑。

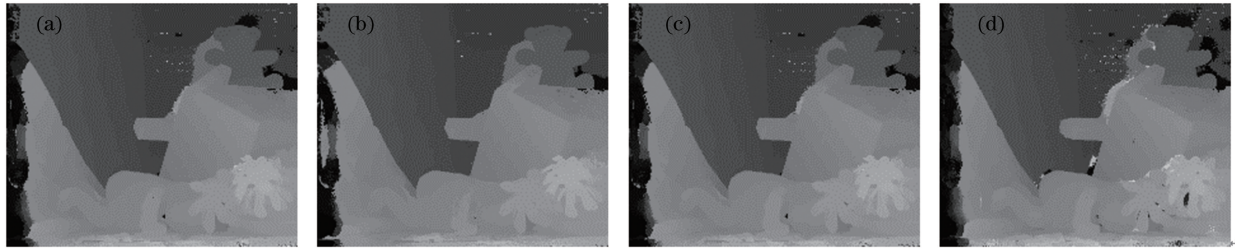


图 2 不同尺寸滤波核得到的初始视差图。(a) $l = \tilde{r}_c^{\min}$; (b) $l = \tilde{r}_{\min}$; (c) $l = \tilde{r}_{\max}$; (d) 加入正则化项后得到的初始视差图

Fig. 2 Initial disparity maps with different sizes of filter kernels. (a) $l = \tilde{r}_c^{\min}$; (b) $l = \tilde{r}_{\min}$;

(c) $l = \tilde{r}_{\max}$; (d) initial disparity map with regularization term

2.4 视差选择

计算得到聚合代价后,为了简单高效地验证在代价空间中进行图像滤波所得到聚合代价的有效性,采用 WTA 策略计算初步视差,可表示为

$$d_i = \operatorname{argmin}_{d \in S_d} \hat{C}(i, d), \quad (8)$$

所提算法的主要贡献在聚合代价环节,因此在视差计算环节只进行简单处理。用不同尺寸的滤波核在代价空间中进行聚合,并加入正则化项以确保不同代价空间的一致性,进而得到更可靠的聚合代价。之后采用 WTA 策略完成视差选择,得到初步视差后,可以用视差精度对跨尺度引导滤波(S-GF)

得到的聚合代价进行可靠性验证。图 3 为整个算法的流程图。

3 实验结果与分析

为了验证所提算法的匹配精度,在 Visual studio 集成开发环境中配合图像处理开源库 Opencv 实现了所提算法,并在立体匹配算法测试平台 Middlebury^[19]上进行了评测。基于图像滤波的立体匹配算法视差图如图 4 所示,原始彩色图像如图 4(a)所示,从上到下依次为 Tsukuba、Venus、Teddy 和 Cones。将不同算法得到的视差图

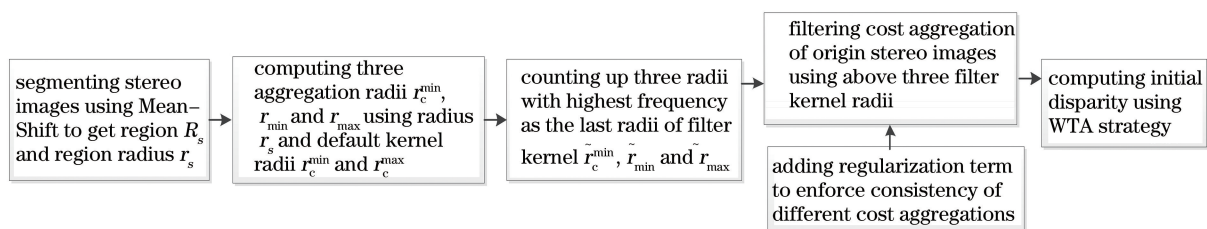


图 3 所提算法的流程图

Fig. 3 Flowchart of proposed algorithm

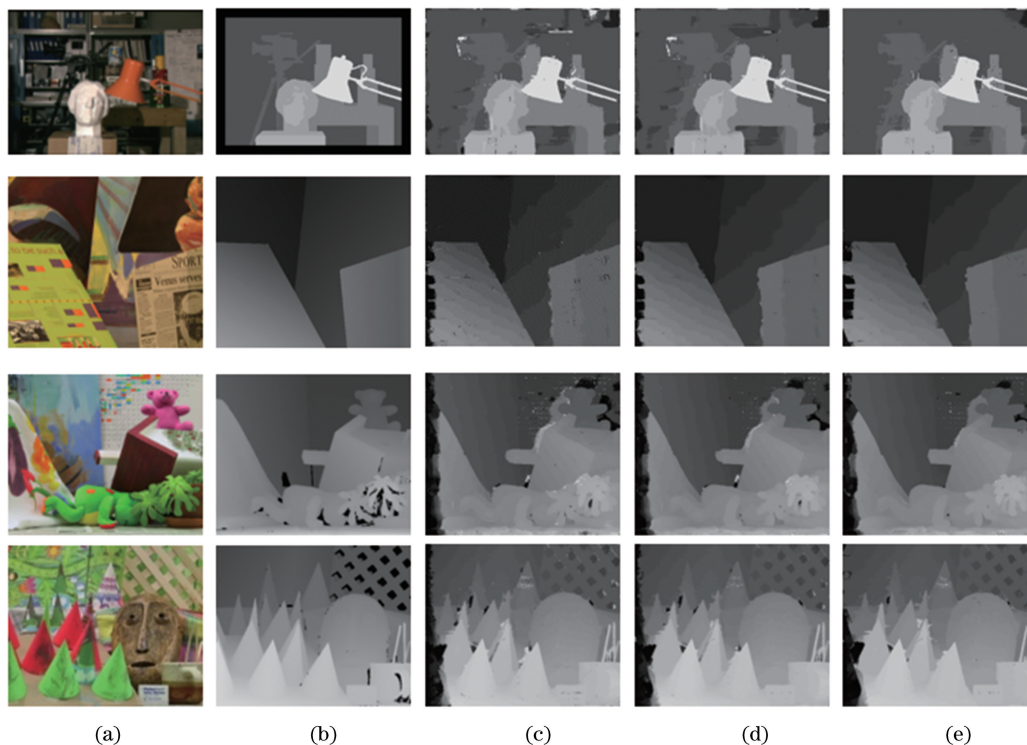


图 4 基于图像滤波的立体匹配算法视差图。(a)原始彩色图像;(b)真实视差图;

(c)引导滤波聚合代价得到的视差图;(d)多尺度引导滤波聚合代价得到的视差图;(e)所提算法得到的视差图

Fig. 4 Disparity maps of stereo matching algorithms based on image filtering. (a) Original color images;

(b) real disparity maps; (c) disparity maps obtained by guided filtering cost aggregation; (d) disparity maps obtained by multiscale guided filtering cost aggregation; (e) disparity maps obtained by proposed algorithm

[图 4(c)~(e)]与真实视差图[图 4(b)]进行比较,进而有效评价各个算法的匹配情况。

值得注意的是,所提算法的主要目标是在确保相当匹配精度的同时进一步提升实时性,并不单纯以提升匹配精度为目标。同时,为了更好地观察聚合代价算法本身对匹配精度的影响,对采用视差精调处理的算法进行了标注,无标注的即为聚合代价后简单采用 WTA 算法得到的初始视差。虽然仅使用了简单的 WTA 策略,所提算法依然得到了较高的匹配精度。与文献[10]中基于 GF 和文献[7]中基于双边滤波器(BF)的两个经典方案相比,跨尺度信息的加入使匹配精度得到明显提升。可以很清楚地看出,所提算法得到的视差图在视差层次感、细节

保留程度、边缘平滑度及弱纹理区域上的匹配精度均有提升,视差图整体上也显得更加精确。相较于文献[12],所提算法得到的视差图匹配精度更高、误匹配区域更少,但整体匹配精度的提升并不明显。所提算法采用 3 种不同的尺寸在原始立体图上进行聚合代价,得到的代价空间保持了一致性,因此可在正则化环节计算插值时节省开销,使算法的整体时间复杂度降低,实际运行时间减少了大约 15%。几种基于引导图像滤波立体匹配算法的运行时间与匹配精度的对比如图 5 所示。

需要说明的是,与 GF 相比,BF 的滤波核尺寸与运行时间是呈非线性的,对时间复杂度而言,其并不适用于 S-GF 算法(实际运行时间为引导滤波的

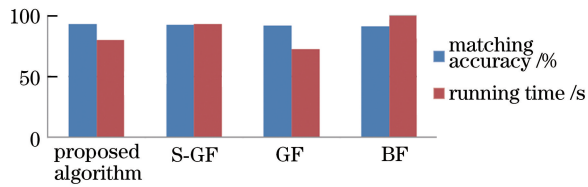


图 5 基于引导图像滤波的几种立体匹配算法的运行时间与匹配精度对比

Fig. 5 Comparison of running time and matching accuracy of several stereo matching algorithms based on guided image filtering

30 倍左右),因此图 5 所示 BF 的运行时间(纵坐标最大值 100 s)并不是实际运行时间。由图 5 可知,所提算法在视差匹配精度和时间复杂度上具备整体优越性。考虑到当前立体图像的分辨率越来越高,以及立体匹配技术在移动平台上的广泛应用,所提算法的实时性无疑具备极大的现实意义。视差匹配的具体情况如表 1 所示,表中 n-occ 为非遮挡区域的错误匹配率, disc 为深度不连续区域的错误匹配率, All 为全部区域的错误匹配率, AvgPBM 为平均错误匹配率,带 * 算法具备视差后处理环节。

表 1 不同算法误匹配像素百分比

Table 1 Percentage of mismatching pixels of different algorithms

Algorithm	Tsukuba			Venus			Teddy			Cones			AvgPBM
	n-occ	All	disc	n-occ	All	disc	n-occ	All	disc	n-occ	All	disc	
Proposed	2.38	2.85	8.4	1.13	1.98	9.26	7.05	14.9	16.8	3.32	11.0	7.99	7.09
BF ^[7]	2.93	3.37	9.40	1.89	2.67	9.43	9.39	17.2	19.6	5.29	13.7	12.3	8.94
GF ^[10]	2.62	3.38	8.16	1.72	2.81	13.2	8.15	16.5	17.5	3.34	12.2	8.69	8.20
S-GF ^[12]	2.30	2.94	8.12	1.09	1.99	9.67	7.04	15.1	16.7	2.99	11.4	8.17	7.29
GlobalGCP* ^[20]	0.87	2.54	4.69	0.46	0.53	2.22	6.44	11.5	16.2	3.59	9.49	8.9	5.60
VarCross* ^[21]	1.99	2.65	6.77	0.62	0.96	3.20	9.75	15.1	18.2	6.28	12.7	12.9	7.60

4 结 论

提出了一种基于跨尺度引导图像滤波的立体匹配算法,该算法对立体图进行预分割,以得到分割区域内像素的半径,并以该半径为指导,用 3 种不同尺寸的滤波核在原尺寸立体图代价空间中进行滤波,以完成聚合代价。通过加入正则化项,确保 3 个聚合代价空间的一致性,以得到更有效的聚合代价。在 Middlebury 测试平台上的实验结果表明所提算法兼备实时性和精确性,在特殊区域的精度高于现有基于图像滤波的方案。现阶段仅在计算机上实现了所提算法,今后可以在多个图形处理器平台上进一步提升算法的实时性。

由表 1 可见,与其他基于图像滤波的方案相比^[7, 10, 12],即使未进行视差精调,基于图像滤波的立体匹配方案在整体匹配精度上也有一定程度的提高,在时间复杂度上更有较大程度的降低。与文献^[12]对原始立体图进行下采样、而后采用特定尺寸滤波核的引导图像滤波器进行聚合代价相比,所提算法在原始立体图代价空间中以不同尺寸核进行滤波得到的聚合代价更具可靠性。进一步分析可知,所提算法在深度不连续处具备较大优势,尤其适用于处理自然场景下边界比较复杂的情况。与进行视差后处理的局部匹配方案(如 VarCross^[20]得到 7.60%的整体匹配精度)相比,所提算法在整体匹配精度上依然具备优势,这反映了所提算法在代价空间中利用图像滤波得到的聚合代价的有效性。虽然与全局匹配方案 GlobalGCP^[20] 5.60%的整体匹配精度相差较大,但所提算法的实时性是全局匹配方案难以达到的。综合来看,所提算法在匹配精度和时间复杂度上达到了较好的平衡。在保证相当匹配精度的前提下,在深度边缘区域的性能依然优良,同时具备高实时性。

参 考 文 献

- [1] Scharstein D, Szeliski R, Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[C]. IEEE Stereo and Multi-Baseline Vision, 2002, 47(1): 7-42.
- [2] Bleyer M, Gelautz M. Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions[J]. Signal Processing: Image Communication, 2007, 22(2): 127-143.
- [3] Kim J C, Lee K M, Choi B T. A dense stereo matching using two-pass dynamic programming with generalized ground control points[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 2(2): 1075-1082.
- [4] Luo W J, Schwing A G, Urtasun R. Efficient deep

- learning for stereo matching[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 5695-5703.
- [5] Žbontar J, LeCun Y. Stereo matching by training a convolutional neural network to compare image patches[J]. The Journal of Machine Learning Research, 2016, 17(1): 2287-2318.
- [6] Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: Theory and experiment [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(9): 920-932.
- [7] Yoon K J, Kweon S. Adaptive support-weight approach for correspondence search[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4): 650-656.
- [8] Shi H, Zhu H, Wang J, *et al.* Segment-based adaptive window and multi-feature fusion for stereo matching[J]. Journal of Algorithms and Computational Technology, 2016, 10(1): 184-200.
- [9] Yang Q X. Stereo matching using tree filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(4): 834-846.
- [10] Rhemann C, Hosni A, Bleyer M, *et al.* Fast cost-volume filtering for visual correspondence and beyond [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2011: 3017-3024.
- [11] 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.
- [12] Zhang K, Fang Y Q, Min D B, *et al.* Cross-scale cost aggregation for stereo matching[J]. IEEE Conference on Transactions on Circuits and Systems for Video Technology, 2017, 27(5): 965-976.
- [13] Zhu S P, Yan L N, Li Z. Stereo matching algorithm based on improved Census transform and dynamic programming [J]. Acta Optica Sinica, 2016, 36 (4): 0415001.
- 祝世平, 闫利那, 李政. 基于改进 Census 变换和动态规划的立体匹配算法 [J]. 光学学报, 2016, 36 (4): 0415001.
- [14] Crow F C. Summed-area tables for texture mapping [C]. ACM Conference on Computer Graphics and Interactive Techniques, 1984: 207-212.
- [15] Pires B R, Singh K, Moura J M F. Approximating image filters with box filters[C]. IEEE International Conference on Image Processing, 2011: 85-88.
- [16] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [17] Liu J, Zhang J X, Dai Y. Dense stereo matching based on region growing[J]. Robot, 2017, 39(2): 182-188.
- 刘杰, 张建勋, 代煜. 基于区域增长的稠密立体匹配 [J]. 机器人, 2017, 39(2): 182-188.
- [18] Milanfar P. A tour of modern image filtering: New insights and methods, both practical and theoretical [J]. IEEE Signal Processing Magazine, 2013, 30 (1): 106-128.
- [19] Scharstein D, Szeliski R, Hirschmüller H. The middlebury stereo vision page[EB/OL]. <http://vision.middlebury.edu/stereo/>.
- [20] Zhang K, Lu J B, Lafruit G. Cross-based local stereo matching using orthogonal integral images[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(7): 1073-1079.
- [21] Wang L, Yang R G. Global stereo matching leveraged by sparse ground control points[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2011: 3033-3040.