

基于融合代价和分段优化的立体匹配算法

谢小鹏^{1,2*}, 欧永东^{2**}, 王银安², 黄泽琼²

¹ 东莞理工学院城市学院智能制造学院, 广东 东莞 523419;

² 华南理工大学机械与汽车工程学院, 广东 广州 510640

摘要 针对已有的立体匹配算法在噪声侵扰中匹配率低, 在视差不连续和弱纹理区域中视差精度不高的问题, 提出了一种基于 Census 变换与互信息(MI)融合代价和分段优化的立体匹配算法。算法主要分两步进行, 即初始视差图获取和视差图优化。第一步, 由 MI 和 Census 融合形成初始匹配代价, 之后由改进引导滤波进行代价聚合, 求得最优匹配代价; 由胜者为王(WTA)策略得到初始视差图。第二步, 先将参考图像分割成超像素, 并为每个超像素拟合视差平面; 其后通过马尔可夫随机场(MRF)来估算超像素的平均视差, 再利用平均视差对邻接系统进行遮挡区域处理和段内的视差精度优化; 最后经中值滤波等后处理得到最终视差图。实验结果表明, 所提算法处理得到的 15 组 Middlebury 测试数据集的视差图在非遮挡区域的平均误匹配率仅为 7.60%, 各阶段运行时间短, 平均每对图像处理耗时 6.8 s, 运行效率高。

关键词 机器视觉; 立体匹配; 互信息; 引导滤波; 马尔可夫随机场; 视差图

中图分类号 TN911.73

文献标志码 A

doi: 10.3788/LOP202158.1215004

Stereo Matching Algorithm Based on Fusion Cost and Segmentation Optimization

Xie Xiaopeng^{1,2*}, Ou Yongdong^{2**}, Wang Yin'an², Huang Zeqiong²

¹ School of Intelligent Manufacturing, City College of Dongguan University of Technology, Dongguan, Guangdong 523419, China;

² School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou, Guangdong 510640, China

Abstract Aiming at solving the problems that the existing stereo matching algorithms have, a low matching rate in noise intrusion and low disparity accuracy in discontinuous disparity and weak texture regions, a stereo matching algorithm based on fusion cost of Census transformation and mutual information (MI) and the segment optimization is proposed in this study. The proposed algorithm mainly involves two steps: initial disparity map acquisition and disparity map optimization. In the first step, the initial matching cost is formed by the fusion of MI and Census, and then, the cost is aggregated by improved guided filtering to obtain the optimal matching cost; the winner-take-all (WTA) strategy is used to obtain the initial disparity map. In the second step, the reference image is divided into superpixels, and a disparity plane is fitted to each superpixel; next, the average disparity of the superpixels is estimated using the Markov random field (MRF). Then, the average disparity is used to process the occlusion area in the adjacent system and optimize the disparity accuracy. Finally, the final disparity map is obtained by median filtering. The experimental results show that the average mismatch rate of the disparity maps of the 15 sets of Middlebury test datasets obtained using the proposed algorithm in a nonocclusion area is only 7.60%, running time of each stage is short, and average processing time for each pair of images is 6.8 s. Overall, the proposed algorithm runs efficiently.

收稿日期: 2020-09-03; 修回日期: 2020-10-10; 录用日期: 2020-10-21

基金项目: 东莞理工学院城市学院重大科技培育基金(2018YZD001Z)

* E-mail: xiexp@scut.edu.com; ** E-mail: ou_yongdong@163.com

Key words machine vision; stereo matching; mutual information; guided filtering; Markov random field; disparity map

OCIS codes 150.1135; 100.6890; 330.1400

1 引言

双目视觉是人工智能领域的重要组成部分,在 3D 重构、导航、缺陷检测等方面有着举足轻重的作用。而立体匹配技术是研究双目视觉的核心问题。根据 Scharstein 等^[1]的理论,立体匹配算法通常需要依次执行 4 个步骤:匹配代价计算、代价聚合、视差计算和视差优化。经上述 4 个步骤对左右图像进行匹配,得到视差图,最终可利用视差值,由几何关系计算出深度信息。

现有的立体匹配算法可分为全局和局部立体匹配算法。全局立体匹配算法的核心思想是在计算出匹配代价后建立一个全局能量函数,通过最小化能量函数来估计视差。常见的全局算法包括协同算法^[2]、置信度传播(BP)^[3-5]、图割(GC)^[6-9]、最小生成树(MST)^[10-11]、段树(ST)等。全局立体匹配算法具有较高的匹配精度,但运行速度较慢,算法复杂度较高。局部立体匹配算法利用了像素周围的局部信息,通过代价聚合估算出视差。常见的匹配代价计算方法包括基于像素的匹配代价,例如绝对差(AD)、灰度差和平方差(SD)等;基于窗口的匹配代价,例如绝对差之和(SAD)、归一化互相关(NCC)等;非参数转换匹配代价计算,例如 Census^[12-13]、Rank^[14]等。与全局立体匹配算法相比,局部立体匹配算法精度略低,但运行速度较快,算法复杂度较低。此外,近年来还出现了一些新颖的匹配代价计算方法,如卷积神经网络(CNN)^[15-16]和深度学习^[17-19],但样本训练时间长。

目前基于像素和窗口的代价计算方法对噪声和光照变化非常敏感。Census 变换对光照的变化不敏感,但是在重复的纹理区域中效果较差。文献[20]提出了一种基于自适应颜色相关性的初始匹配成本计算方法,与传统的颜色度量相结合,但计算复杂度高。Zhang 等^[21]阐释了一种基于跨区域匹配窗口的匹配算法,该算法利用颜色相似度和距离约束产生一个由多个相邻的水平或垂直分割线组成的匹配窗口,但抗干扰还不强。Hirschmuller^[22]将互信息熵作为初始匹配成本。Birchfield 等^[23]提出了一种相邻像素插值匹配方法,提高了初始匹配精度。马晴晴等^[24]在代价聚合时融合了边缘和颜色

信息,使得最小生成树效率有所提升。文献[25]则在半全局立体匹配(SGM)中引入了鲁棒特征(SURF),且用改进的双边滤波优化视差,图像误匹配率得到降低。

基于此,本文提出了一种新颖的匹配算法来获取和改善视差图。所提算法可分为两步:利用 Census 变换和互信息(MI)的加权融合获得初始匹配代价,由改进引导滤波聚合得到聚合代价矩阵,之后由胜者为王(WTA)策略获得初始视差图;对原图像进行超像素分割,根据超像素之间的约束由马尔可夫随机场(MRF)优化视差平面,再对融合后的段内进行视差优化,最后经过中值滤波等处理获得较高精度的视差图。经过两步运算后,所得的视差图改善了原来弱纹理区域的视差精度,且运算速度快,算法实时性高。

2 算法步骤

所提算法的工作流程如图 1 所示。

2.1 初始视差图获取

MI^[22]对记录和照明变化不敏感。它能根据两个图像的熵(即它们的信息内容)和联合熵来定义:

$$S_{MI}(I_L, I_R) = H(I_L) + H(I_R) - H(I_L, I_R), \quad (1)$$

式中: $S_{MI}(I_L, I_R)$ 为左右图像的互信息; $H(I_L)$ 、 $H(I_R)$ 分别为左、右图像的熵; $H(I_L, I_R)$ 为左右图像的联合熵。那么基于互信息的相似测度为

$$C_{MI}(\mathbf{p}, \mathbf{d}) = -S_{MI}(I_L, I_R), \quad (2)$$

式中: $C_{MI}(\mathbf{p}, \mathbf{d})$ 为互信息的匹配代价; \mathbf{d} 为左右图像像素点的视差。

实际场景的光照条件并不理想,双目图像对不满足颜色一致性。而 Census 变换对光照变化的鲁棒性强,与互信息融合后能进一步提高匹配准确性。

Census 变换的表达式为

$$B_{cen}(\mathbf{p}) = \bigotimes_{q \in N_p} \xi[\bar{I}(\mathbf{p}), I(\mathbf{q})], \quad (3)$$

$$\xi[\bar{I}(\mathbf{p}), I(\mathbf{q})] = \begin{cases} 0, & \bar{I}(\mathbf{p}) < I(\mathbf{q}) \\ 1, & \bar{I}(\mathbf{p}) \geq I(\mathbf{q}) \end{cases}, \quad (4)$$

式中: $B_{cen}(\mathbf{p})$ 为点 \mathbf{p} 经 Census 变换后所得的二进

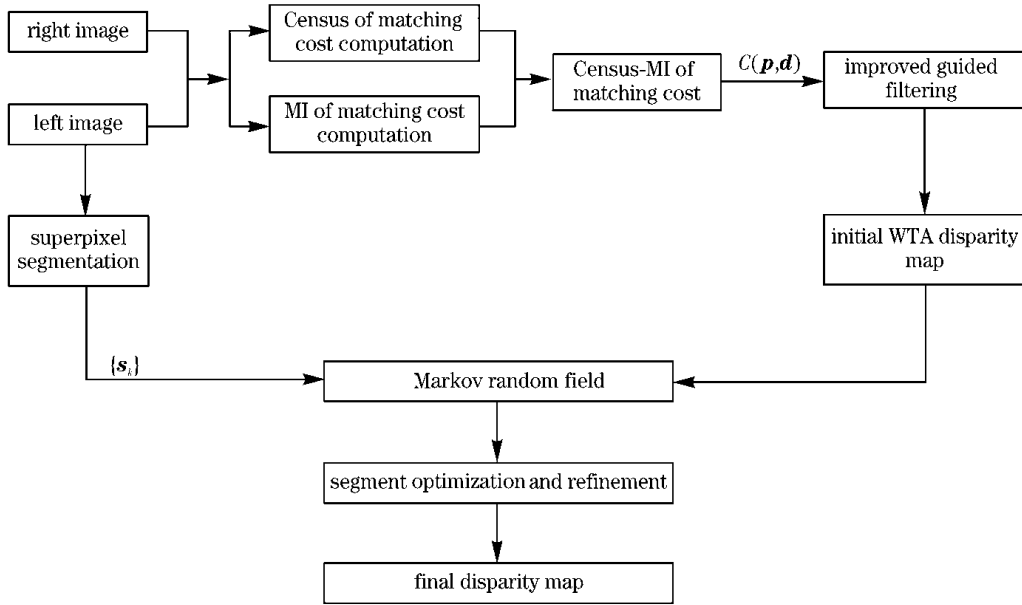


图 1 所提算法流程图

Fig. 1 Flowchart of the proposed algorithm

制比特串; N_p 为像素 p 的窗口变换域; q 为中心像素 p 的邻近像素; $I(q)$ 为像素 q 的灰度值; $\bar{I}(p)$ 为变换域中所有像素的灰度均值。那么 Census 变换的匹配代价为

$$C_{\text{cen}}(p, d) = 1 - \exp\left\{-\frac{\text{Hamming}[B_{\text{cen1}}(p), B_{\text{cen2}}(p+d)]}{\lambda_{\text{cen}}}\right\}, \quad (5)$$

式中: $B_{\text{cen1}}(p)$ 为左图像经 Census 变换后 p 点的二进制比特串; $B_{\text{cen2}}(p+d)$ 为右图像经 Census 变换后 $(p+d)$ 点的二进制比特串; $\text{Hamming}[\cdot]$ 为汉明距离计算函数; λ_{cen} 为 Census 变换的控制参数。

MI 和 Census 变换融合后的总匹配代价为

$$C(p, d) = \lambda_1 C_{\text{MI}}(p, d) + \lambda_2 C_{\text{cen}}(p, d), \quad (6)$$

式中: λ_1 为 MI 匹配代价的权重值; λ_2 为 Census 变换匹配代价的权重值。

单个像素的匹配代价往往易受噪声影响并很容易导致误匹配, 因此需要周围像素的聚合限制来提高可靠性。在众多的聚合算法中, 引导滤波性能优越, 受到了广泛的关注。引导滤波的具体过程如下。设 I 为引导图像, 通常取左图像为引导图像; Y 为待滤波图像, 即上述构造的匹配代价 $C(p, d)$; U 为滤波后图像。假设在以各个像素为中心的窗口 $\omega_{k'}$ 内, U 由 I 线性转换得到:

$$U_i = a_{k'} I_i + b_{k'}, \quad i \in \omega_{k'}, \quad (7)$$

式中: k' 为窗口内的中心像素点; U_i 和 I_i 分别为窗

口内不同像素点在滤波后图像和引导图像中的对应点; $\omega_{k'}$ 为矩形窗口; $a_{k'}$ 和 $b_{k'}$ 为线性系数。得到 $a_{k'}$ 和 $b_{k'}$ 的公式为

$$E(a_{k'}, b_{k'}) = \sum_{i \in \omega_{k'}} [(a_{k'} I_i + b_{k'} - r_i)^2 + \epsilon a_{k'}^2], \quad (8)$$

式中: ϵ 为规整化参数, 通常 $\epsilon > 0$, 取值越大, 滤波后得到的图像越平滑, 取值越小, 滤波后图像的边缘性越好; r_i 为引导图像中与滤波后图像的对应点。传统的引导滤波算法忽视了不同窗口像素纹理的差异性, 因此本文将自适应权重的思想引入到引导滤波中, 并提出如下的改进策略。在窗口 $\omega_{k'}$ 内, 为使 ϵ 能根据像素纹理自适应改变, 将引入梯度信息并以此来定义权重因子。权重因子的计算公式为

$$\Lambda = \max\left[\frac{\|I(q_{\text{max}}) - I(p)\|}{\|I(q_{2\text{max}}) - I(p)\|}, \Theta\right] \times \frac{\|I(q_{\text{max}}) - I(p)\|}{\eta_{\Lambda}}, \quad (9)$$

式中: $I(q_{\text{max}})$ 为以点 p 为中心像素的周围像素的最大灰度值; $I(q_{2\text{max}})$ 为以点 p 为中心像素的周围像素的次最大灰度值; Θ 为比重阈值常数; η_{Λ} 为像素差缩放系数。由上得到 $\tilde{\epsilon} = \epsilon / \Lambda$, 用 $\tilde{\epsilon}$ 代替 (8) 式中的 ϵ 即可。

图 2 为自适应权重改进前后的 ArtL 图像视差图对比, 可以看出, 改进后视差图的物体边缘部分更加锐利突出, 且一些细节区域也被保留下来。

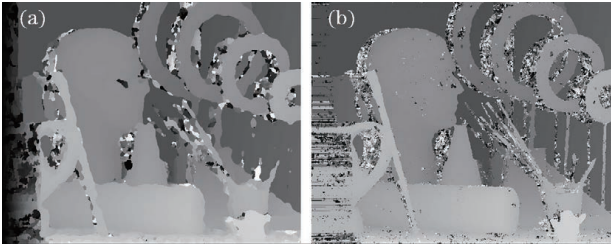


图 2 自适应权重改进前后的初始视差图。(a)改进前；
(b)改进后

Fig. 2 Initial disparity map before and after the adaptive weighting improvement. (a) Before adaptive weight improvement; (b) after adaptive weight improvement

在匹配代价聚合之后,使用 WTA 算法得到初始视差图,即将每个像素匹配代价最小的视差值选择为最终视差,计算公式为

$$d(\mathbf{p}) = \arg \min_{d \in [0, d_{\max}]} \tilde{C}(\mathbf{p}, d), \quad (10)$$

式中: $d(\mathbf{p})$ 为点 \mathbf{p} 的最终视差值; d_{\max} 为最大视差值; $\tilde{C}(\mathbf{p}, d)$ 为经过聚合后的代价值。

2.2 视差图优化

左参考图像通过图割^[26]分割为超像素 $\{s_k\}$,且把每个超像素的平均视差 $\{\mu_s\}$ 用作平面的约束拟合,提高了可靠性。将超像素视为图节点,通过 MRF 来最小化能量函数,能量函数表达式为

$$E(\mu) = \sum_{s \in \Omega} \varphi_s(\mu_s) + \rho \sum_{(s,t) \in \mathcal{D}} \chi_{st}(\mu_s, \mu_t), \quad (11)$$

式中: Ω 为超像素 s 的集合; \mathcal{D} 为相邻超像素的集合; $\varphi_s(\mu_s)$ 为数据项; $\chi_{st}(\mu_s, \mu_t)$ 为平滑项; ρ 为用于平衡平滑项的影响参数。

为了测量视差中心的置信度,将超像素的视差分布分为直方图块,计算超像素 s 中 WTA 视差 $d_s(x)$ 落入宽度为 L 的 $o(\mu_s)$ 的总数。 s 的数据项定义为

$$\varphi_s(\mu_s) = N_s - \sum_{i'=1}^{N_s} \mathcal{R}[d_s(x_{i'}) \in o(\mu_s)], \quad (12)$$

式中: N_s 为超像素 s 中的像素个数; $\mu_s = 0, L, 2L, \dots$, 较低的数据项由于负号而具有较高的置信度。 $\mathcal{R}(\cdot)$ 是条件的函数,定义为

$$\mathcal{R}(\cdot) = \begin{cases} 1, & \cdot \text{ is true} \\ 0, & \cdot \text{ is false} \end{cases} \quad (13)$$

数据项的数值是基于投票策略的,因此落入同一仓中的结果越多,置信度越高。但遮挡区域的视差很容易受到干扰,难以保持一致,故遮挡区域的函数值将由平滑项主导。平滑项增强了相邻超像素之

间视差分布中心的相似性,定义为

$$\chi_{st}(\mu_s, \mu_t) = \max(\omega_{st}, \delta) \Gamma(s, t) P(\mu_s, \mu_t), \quad (14)$$

式中: $\Gamma(s, t)$ 为相邻超像素 s 和 t 之间的共享边界长度; δ 为下界截断值。 ω_{st} 为色相相似权重,定义为

$$\omega_{st} = \exp\left(\frac{-\|I(s) - I(t)\|^2}{\gamma}\right), \quad (15)$$

式中: γ 为控制颜色权重影响的参数; $I(s)$ 为超像素 s 的平均色相灰度值。 P 的定义为

$$P(\mu_s, \mu_t) = \begin{cases} 0, & |\mu_s - \mu_t| = 0 \\ 1, & |\mu_s - \mu_t| = a \\ \kappa, & |\mu_s - \mu_t| > a \end{cases}, \quad (16)$$

式中: a 为共平面阈值; κ 为非共面惩罚值。对于一对超像素 $(s, t) \in \mathcal{D}$,如果它们的平均视差不相似,即 $|\mu_s - \mu_t| > a$,则该对超像素不属于同一个平面。因此,这切断了可能的前景与背景的连接,提高了遮挡区域的视差精度。其中 $\kappa > 2$ 。

经过 MRF 优化后,全局图像的视差得到了优化。对于具有相同标识符的超像素,将它们融合成一个新的超像素,再考虑该超像素内的视差。每个新的超像素内的视差近似为线性,且具有适当的边界,那么其视差分布应均匀分布。考虑到超像素的不规则边界形状,将超像素内的视差分布建模为高斯分布:

$$F_{\text{Gauss}}(d) = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{(d - \mu_s)^2}{2\sigma_s^2}\right], \quad (17)$$

式中: σ_s 为新超像素 s 的视差方差。

随后,找到超像素内视差值符合 $d \notin [\mu_s - 3\sigma_s, \mu_s + 3\sigma_s]$ 的值,并将其替换为相邻像素视差值的均值,最后再经过中值滤波获得最终的视差图。

图 3 为各个阶段处理的视差图,可以看出,视差图经所提算法处理后有了较为明显的改善。图 3(b)中的视差图有不少的误匹配点,视差图精度低。经 MRF 优化后,图 3(c)显示出了较为明显的视差改善。进一步地,经过分段优化等后续处理后,局部区域的精度也进一步提高了,如图 3(d)所示。但是可以看到,相比图 3(c),图 3(d)圆圈标志中的直线存在断裂情况。这是由于图 3(c)圆圈标志中有线条过细的部分,因此在后续的中值滤波处理中,细小的线条被当成噪声点过滤掉了。而图像中其他绝大部分物体的边缘部分则更明显,噪声点总体上减少了,因此所提算法提高了整幅图像的视差精度。

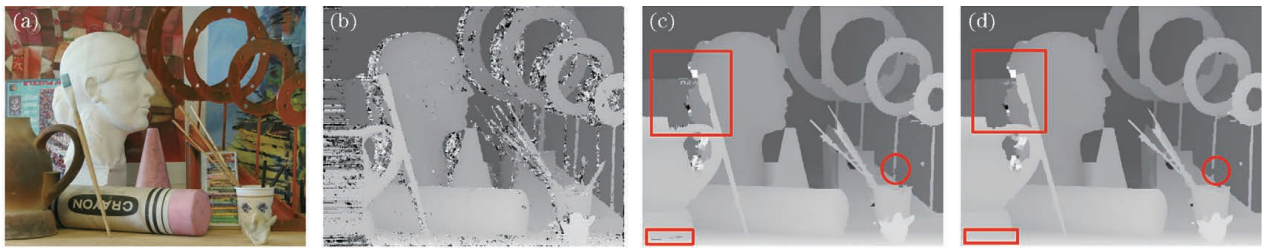


图 3 各个阶段的视差图。(a) ArtL 原图像;(b)经 Census 变换和 MI 的融合代价及改进引导滤波聚合后得到的初始视差图;(c)经 MRF 优化得到的视差图;(d)经分段优化和中值滤波后得到的最终视差图

Fig. 3 Disparity maps of each stage. (a) ArtL original image; (b) initial disparity map obtained by fusion cost of Census transformation and MI and improved guided filter aggregation; (c) disparity map optimized by MRF; (d) final disparity map obtained after segmentation optimization and median filtering

3 实验结果

为验证所提算法的各方面性能,用 C++ 语言编程实现算法,硬件平台中央处理器(CPU)为 Intel (R) Core(TM) i5-9300H,主频为 2.40 GHz。在包括各种静态室内场景的 15 个测试图像对的 Middlebury 数据集上进行实验测试,并对所提算法

与其他视差优化算法进行了比较。全分辨率数据集的分辨率约为 3000×2000 ,本文使用半分辨率数据集进行实验和评估。

3.1 参数设置及影响

经过多次的算法实验比对,选取了实验效果最为优异的参数作为最终的参数值。最终实验所用的参数如表 1 所示。

表 1 实验相关参数设置

Table 1 Experimental related parameter settings

Parameter	λ_{cen}	λ_1	λ_2	Θ	η_A	ϵ	γ	a	κ	δ
Value	25	0.3	0.7	2	3	10^{-5}	20	2	16	0.01

在 MRF 优化中,关键参数是(11)式中的影响因子 ρ 。此参数影响着平滑度约束。当 ρ 设置为 0 时,平滑约束将失去作用,因此会退化为局部约束。随着 ρ 增加到无穷大,平滑性约束将主导估计,导致像素点之间的标识符差异不明显。因此,为了获得良好的性能,且验证参数 ρ 对实验结果的影响,选择合适的 ρ ,将 Adirondack、ArtL、Jadeplant、Playroom、

Vintage 作为测试图,并用视差图全部和非遮挡区域的误匹配率作为评价标准,通过分析不同 ρ 值所得视差图的误匹配率来确定 ρ 。 ρ 相当于一个权重值,表示平滑项在能量函数中的占比,而传统的全局匹配算法能量函数均取值为 1。为了探索 ρ 对整个能量函数的优化作用,将在 $[0, 1]$ 区间范围搜寻最佳 ρ 值。结果如图 4 所示。

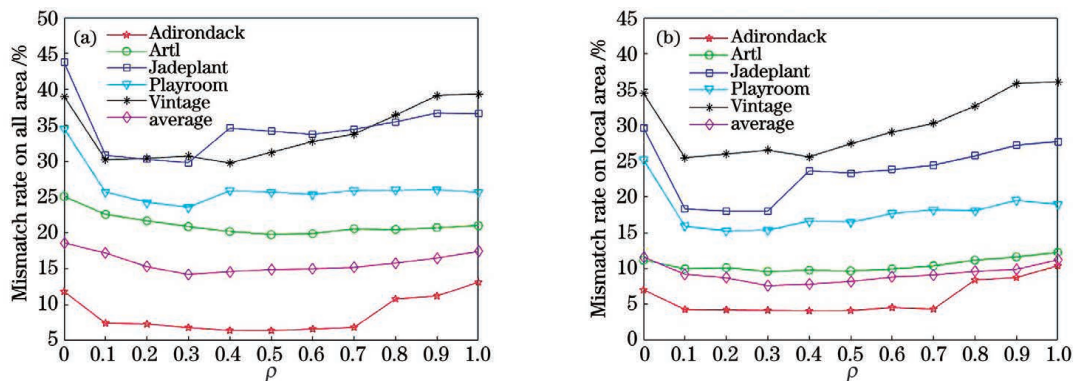


图 4 ρ 对视差图的影响。(a) ρ 对视差图全部区域的影响;(b)参数 ρ 对视差图非遮挡区域的影响

Fig. 4 Influence of ρ on the disparity map. (a) Influence of ρ on all area of the disparity map; (b) influence of ρ on the non-occlusion area of the disparity map

图 4 绘制了 5 组图像对的误匹配率曲线和全部 15 幅图像对的平均误匹配率曲线。当 ρ 在 0.1 到

0.7 之间取值时,5 条曲线所反映的误匹配率较低,且对参数 ρ 不敏感。在 Middlebury 训练集中,当 ρ

设置为 0.3 时,可取得最佳的匹配效果。

k 为控制图像所分割出的超像素大小的物理量。同理,参数 k 会影响超像素的大小,而较大的 k 会导致产生较少的超像素数量。为了验证 k 对实

验效果的影响并确定参数 k 值,同样选取以上 5 组图像对,分析并对比参数 k 的取值对视差图误匹配率的影响,确定最终的 k 值。实验的数据结果如图 5 所示。

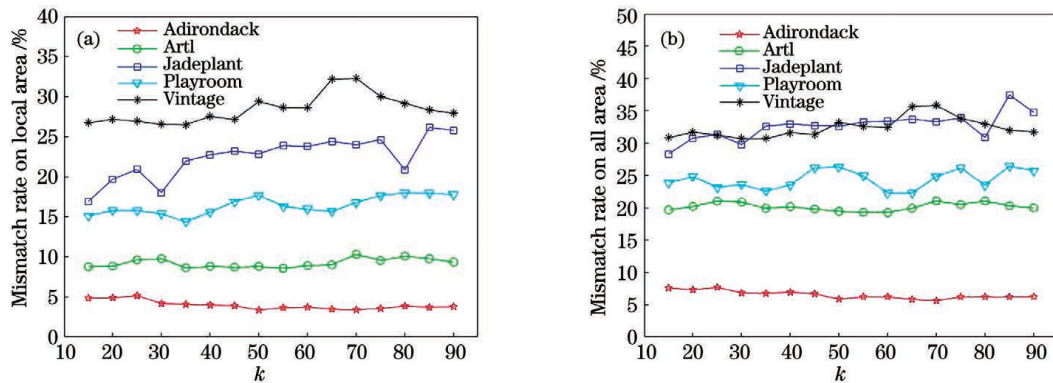


图 5 k 对视差图的影响。(a) k 对视差图非遮挡区域的影响;(b) k 对视差图全部区域的影响

Fig. 5 Influence of k on the disparity map. (a) Influence of k on the non-occlusion area of the disparity map; (b) influence of k on all area of the disparity map

经过实验数据验证, k 较小时,所提算法处理的图像细节效果比较理想;而 k 较大时,则可以处理较大的无纹理区域。在给定的范围(15~90)内, k 的取值对结果影响较小。数据显示,把 k 取为 35 时,所提算法对整个 Middlebury 训练集的处理效果最为理想,即可获得最低的平均误匹配率。

3.2 各个阶段的测试性能

表 2 为 k 值对全部图像的运行时间的影响。可以看出,随着 k 值的增大,平均运行时间在逐步下降,且当 k 值大于 60 时,平均运行时间保持在 4.0 s 左右。所提算法所取的 k 值为 35,平均运行时间约为 6.8 s,今后还可进一步地研究如何利用图形运行处理器(GPU)进行并行运算以缩短处理时间。

表 2 参数 k 对图像的平均处理时间的影响

Table 2 Influence of k on the average processing time of the images

k	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90
Time / s	19.8	8.9	7.4	6.8	6.4	5.8	5.5	5.0	5.3	4.5	4.7	4.2	4.3	4.3	4.2

图 6 为所提算法各个阶段在处理 Vintage 图像时的运行时间,显示了 k 值对实验各个阶段的运行时间的影响。结果表明: k 值小于 30 时,随着 k 值

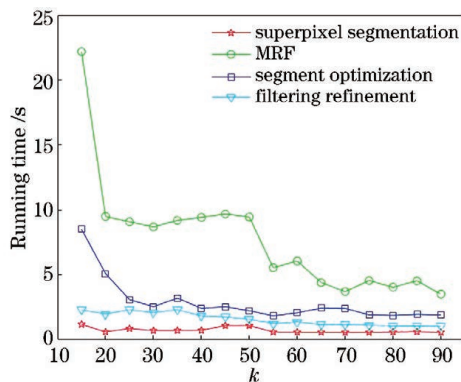


图 6 k 对各个优化阶段的运行时间的影响

Fig. 6 Influence of k on the running time of each optimization stage

的增大,超像素分割和 MRF 优化的运行时间显著下降; k 值在 30 与 50 之间时,各阶段的运行时间对 k 值不敏感; k 值大于 50 时,超像素分割的运行时间进一步下降,且在之后的取值范围内, k 值对整体的运行时间影响不显著。

实验结果显示,所提算法各个阶段的运行速度均较快。

3.3 15 组标准 Middlebury 图像测试结果

图 7(a)和(b)为第一步处理获取的初始视差图,整体较为粗糙;图 7(c)和(d)为经第二步优化后得到的视差图。

为了能够更加客观地评价所提算法的性能,计算了 15 组标准 Middlebury 图像对在非遮挡区域的误匹配率,并与 SED^[27]、r200high^[28]、ICSG^[29]、LS-ELAS^[30]、AD_Cens^[31]、ADSM^[32] 算法进行了对照比较,结果如表 3 所示。

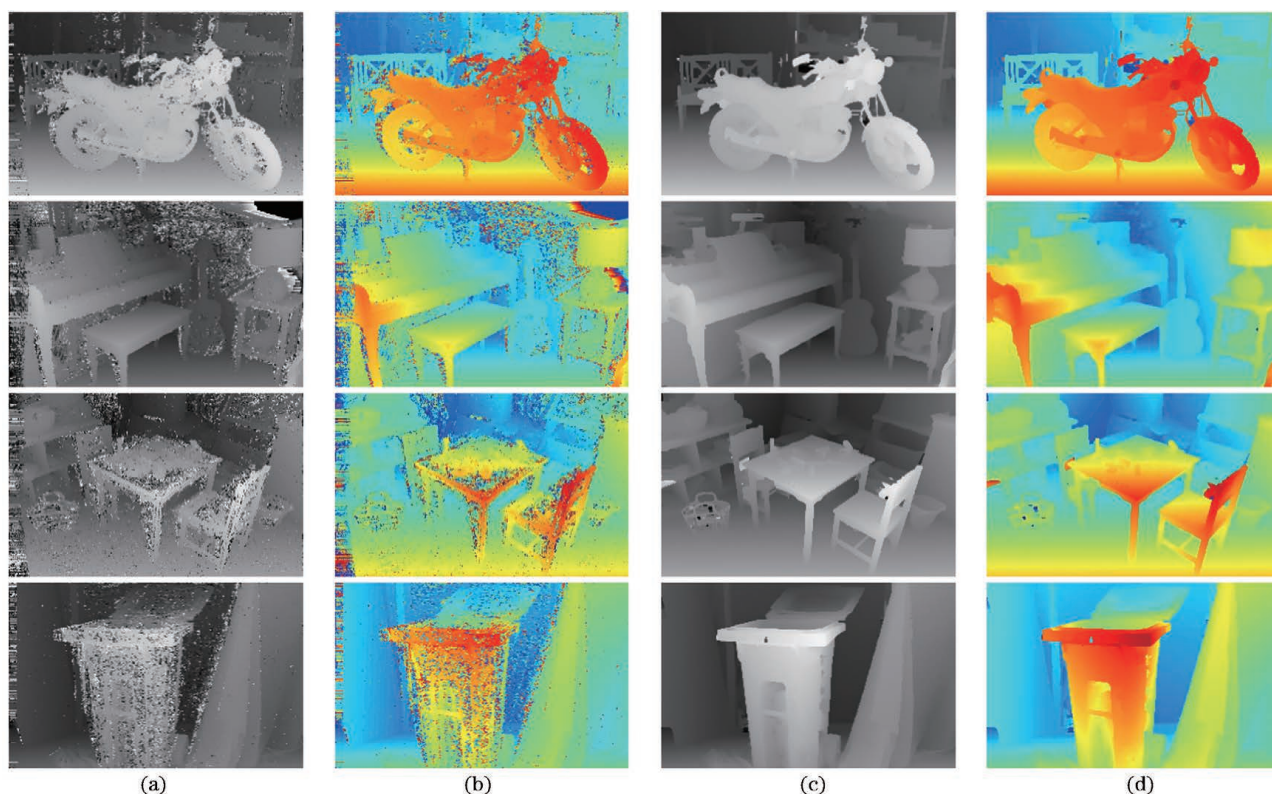


图 7 初始视差图和优化后的视差图。(a)初始视差图;(b)彩色初始视差图;(c)所提算法优化后的视差图;(d)所提算法优化后的彩色视差图

Fig. 7 Initial disparity map and optimized disparity map. (a) Initial disparity map; (b) color initial disparity map; (c) disparity map optimized by the proposed algorithm; (d) color disparity map optimized by the proposed algorithm

表 3 不同算法对非遮挡像素点的视差平均误差

Table 3 Disparity average error of non-occluded pixel points obtained by different algorithms unit: %

Image	SED ^[27]	r200high ^[28]	ICSG ^[29]	LS_ELAS ^[30]	AD_Cens ^[31]	ADSM ^[32]	Proposed algorithm
Adirondack	23.70	22.60	24.00	8.46	8.81	13.30	4.17
ArtL	15.60	12.80	6.93	3.83	5.27	6.10	3.57
Jadeplant	10.60	45.90	54.20	41.10	21.30	15.00	11.16
Motorcycle	18.30	12.00	12.00	5.12	4.10	3.67	2.45
MotorcycleE	17.70	12.00	10.40	5.80	5.58	5.67	2.36
Piano	17.70	19.20	15.60	5.54	6.57	7.08	8.82
PianoL	29.70	31.80	29.30	8.97	31.40	20.60	17.76
Pipes	28.50	15.30	18.40	7.44	8.50	6.57	3.92
Playroom	21.30	28.30	24.70	8.76	7.44	13.20	6.00
Playable	18.20	36.40	26.70	22.40	24.80	23.10	5.87
PlayableP	15.90	18.00	10.70	3.47	5.48	3.55	3.00
Recycle	16.20	18.30	17.70	6.93	3.32	5.76	4.11
Shelves	14.40	27.80	23.60	8.26	12.20	17.20	17.21
Teddy	6.65	10.70	7.73	2.29	3.11	3.05	1.95
Vintage	31.60	78.90	72.90	13.10	10.20	10.10	21.75
Average error	25.90	23.10	21.30	9.66	9.21	8.95	7.60

从表 3 可知:所提算法的平均误匹配率仅为 7.60%, 低于 SED 等对比算法, 且相对于 SED、r200high、ICSG 算法, 所提算法具有较大的优势, 大

幅降低了误匹配率; 但所提算法处理“Vintage”、“Shelves”、“PianoL”图像对时的平均误差较大, 这是因为 Vintage 图像对的视差达到了 380, 是 15 幅

测试图像对中最大的,且图像尺寸大,因此处理起来耗时长,在大范围视差内的误匹配点多,在全部测试图像中误匹配率最高;“Shelves”为架子图像对,图像中木板物体的重复纹理区域多,容易干扰初始匹配代价的计算;而“PianoL”为“Piano”图像对的加强曝光后的图像对,曝光强度大,造成了较高的误匹配率。所提算法对其余图像对的处理表现较为良好,在弱纹理、视差不连续区域能够提高视差精度,故而提高了整体的视差精度。因此,所提算法在视差范围较小、重复纹理区域少及弱曝光的条件下能达到较为可观的效果。

4 结 论

提出了一种基于 Census 变换和 MI 融合的匹配代价和分段优化的立体匹配算法。该算法将 Census、MI 独立的代价值加权融合形成初始匹配代价,经过改进的引导滤波聚合得到聚合代价矩阵,有效地提高了图像匹配的稳健性和抗光照干扰能力。另外,在得到初始视差图之后,又经过了超像素分割和 MRF 全局的能量优化处理,使得平面视差精度得到了优化,且在段内进行了视差精度提升,进一步改善了图像在遮挡区域的视差精度。通过比对其余近年来的算法,结果表明,所提算法的视差精度优于其他比对算法,具有良好的抗干扰性和运行效率,在视差不连续和弱纹理区域的视差均有所提高。但是,所提算法在某些细节区域的视差还存在精度不足的问题,一些交界点、线的视差处理仍有待提高和精细化,且 MRF 的处理时间也需要进一步降低,今后将继续往降低运行时间和提高细节区域及交界点、线视差精度方向进行研究。

参 考 文 献

- [1] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *International Journal of Computer Vision*, 2002, 47(1/2/3): 7-42.
- [2] Zitnick C L, Kanade T. A cooperative algorithm for stereo matching and occlusion detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(7): 675-684.
- [3] Yang Q X, Wang L, Yang R G, et al. Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(3): 492-504.
- [4] Sun J, Zheng N N, Shum H Y, et al. Stereo matching using belief propagation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(7): 787-800.
- [5] Besse F, Rother C, Fitzgibbon A, et al. PMBP: PatchMatch belief propagation for correspondence field estimation [J]. *International Journal of Computer Vision*, 2014, 110(1): 2-13.
- [6] Gong W B, Gu G H, Qian W X, et al. Stereo matching algorithm based on image segmentation and adaptive support weight [J]. *Acta Optica Sinica*, 2015, 35(z2): s210002.
龚文彪, 顾国华, 钱惟贤, 等. 基于图像分割和自适应支撑权重的立体匹配算法 [J]. *光学学报*, 2015, 35(z2): s210002.
- [7] Worbly J, Maclean W J. Establishing visual correspondence from multi-resolution graph cuts for stereo-motion [C] // *Fourth Canadian Conference on Computer and Robot Vision (CRV '07)*, May 28-30, 2007, Montreal, QC, Canada. New York: IEEE Press, 2007: 313-320.
- [8] Hong L, Chen G. Segment-based stereo matching using graph cuts [J]. *Proceedings of the IEEE*, 2004, 1: 74-81.
- [9] Boykov Y, Veksler O, Zabih R, et al. Fast approximate energy minimization via graph cuts [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(11): 1222-1239.
- [10] Yang Q X. Stereo matching using tree filtering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(4): 834-846.
- [11] Li L C, Yu X, Zhang S L, et al. 3D cost aggregation with multiple minimum spanning trees for stereo matching [J]. *Applied Optics*, 2017, 56(12): 3411-3420.
- [12] Liu J G, Yu L, Liu S J, et al. Stereo matching algorithm based on improved Census transform and multi-scale space [J]. *Journal of South China University of Technology (Natural Science Edition)*, 2017, 45(12): 43-49.
刘建国, 俞力, 柳思健, 等. 基于改进 Census 变换和多尺度空间的立体匹配算法 [J]. *华南理工大学学报 (自然科学版)*, 2017, 45(12): 43-49.
- [13] Ding M, Cao Y F, Wu Q X, et al. Crater region detection based on Census transform and Boosting [J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2009, 41(5): 682-687.
丁萌, 曹云峰, 吴庆宪, 等. 基于 Census 变换和 Boosting 方法的陨石坑区域检测 [J]. *南京航空航天大学学报*, 2009, 41(5): 682-687.
- [14] Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence [M] // *Eklundh J*

- O. Computer vision-ECCV '94. Lecture notes in computer science. Berlin: Springer, 1994, 801: 151-158.
- [15] Žbontar J, Lecun Y. Computing the stereo matching cost with a convolutional neural network[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1592-1599.
- [16] Mayer N, Ilg E, Häusser P, et al. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4040-4048.
- [17] Cheng M Y, Gai S Y, Da F P, et al. A stereo-matching neural network based on attention mechanism[J]. Acta Optica Sinica, 2020, 40(14): 1415001.
程鸣洋, 盖绍彦, 达飞鹏, 等. 基于注意力机制的立体匹配网络研究 [J]. 光学学报, 2020, 40(14): 1415001.
- [18] Wang Y F, Wang H W, Yu G, et al. Stereo matching algorithm based on three-dimensional convolutional neural network[J]. Acta Optica Sinica, 2019, 39(11): 1115001.
王玉锋, 王宏伟, 于光, 等. 基于三维卷积神经网络的立体匹配算法 [J]. 光学学报, 2019, 39(11): 1115001.
- [19] Tulyakov S, Ivanov A, Fleuret F, et al. Weakly supervised learning of deep metrics for stereo reconstruction[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 1348-1357.
- [20] Heo Y S, Lee K M, Lee S U, et al. Robust stereo matching using adaptive normalized cross-correlation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(4): 807-822.
- [21] Zhang K, Lu J B, Lafruit G, et al. Cross-based local stereo matching using orthogonal integral images[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(7): 1073-1079.
- [22] Hirschmuller H. Stereo processing by semiglobal matching and mutual information [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 328-341.
- [23] Birchfield S, Tomasi C. Depth discontinuities by pixel-to-pixel stereo[C]//Sixth International Conference on Computer Vision, January 7, 1998, Bombay, India. New York: IEEE Press, 1998: 1073-1080.
- [24] Ma Q Q, Wang C F. Non-local stereo matching algorithm based on color and edge information[J]. Laser & Optoelectronics Progress, 2020, 57(10): 101020.
马晴晴, 王彩芳. 基于颜色和边缘信息的非局部立体匹配算法 [J]. 激光与光电子学进展, 2020, 57(10): 101020.
- [25] Wang Y P, Qin A N, Hao Q, et al. Semi-global stereo matching of remote sensing images combined with speeded up robust features [J]. Acta Optica Sinica, 2020, 40(16): 1628003.
王阳萍, 秦安娜, 郝旗, 等. 结合加速鲁棒特征的遥感影像半全局立体匹配 [J]. 光学学报, 2020, 40(16): 1628003.
- [26] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167-181.
- [27] Peña D, Sutherland A. Disparity estimation by simultaneous edge drawing[M]//Chen C S, Lu J W, Ma K K. Computer vision-ACCV 2016 workshops. Lecture notes in computer science. Cham: Springer, 2017, 10117: 124-135.
- [28] Keselman L, Woodfill J I, Jepsen A G, et al. Intel(R) RealSense(TM) stereoscopic depth cameras [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1267-1276.
- [29] Shahbazi M, Sohn G, Théau J, et al. Revisiting intrinsic curves for efficient dense stereo matching [J]. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016, 3: 123-130.
- [30] Jellal R A, Lange M, Wassermann B, et al. LS-ELAS: line segment based efficient large scale stereo matching[C]//2017 IEEE International Conference on Robotics and Automation (ICRA), May 29-June 3, 2017, Singapore. New York: IEEE Press, 2017: 146-152.
- [31] Mei X, Sun X, Zhou M C, et al. On building an accurate stereo matching system on graphics hardware[C]//2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), November 6-13, 2011, Barcelona, Spain. New York: IEEE Press, 2011: 467-474.
- [32] Ma N, Men Y B, Men C G, et al. Accurate dense stereo matching based on image segmentation using an adaptive multi-cost approach [J]. Symmetry, 2016, 8(12): 159.