

基于三角剖分与超像素结构的 3D 标签优化方法

李萌昊^{1,2}, 葛宝臻^{1,2*}, 权佳宁^{1,2}, 陈其博^{1,2}

¹天津大学精密仪器与光电子工程学院, 天津 300072;

²光电信息技术教育部重点实验室, 天津 300072

摘要 在立体匹配问题中, 基于 3D 标签的算法可以取得更高精度的亚像素视差图。针对 3D 标签的随机初始化问题, 提出一种基于超像素结构与三角剖分的标签初始化方法。利用基于超像素边界提取到的特征点构建三角剖分, 生成初始 3D 标签; 针对图割算法迭代优化 3D 标签时效率的问题, 在超像素结构上利用图割算法进行全局优化 3D 标签, 迭代中加入对当前标签状态的假设扩展标签候选, 提升了标签搜索效率。在 Middlebury2014 数据集上对方法进行验证, 实验结果表明, 所提方法的平均误匹配率(8.31%) 低于 LocalExp 算法的平均误匹配率(8.39%), 并且处理单幅图像耗费的平均时间约为 LocalExp 算法的 70%。

关键词 机器视觉; 立体匹配; 3D 标签; 超像素; 三角剖分; 图割

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202259.0815003

3D Label Optimization Based on Triangulation and Superpixel Structures

Li Menghao^{1,2}, Ge Baozhen^{1,2*}, Quan Jianing^{1,2}, Chen Qibo^{1,2}

¹*School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China;*

²*Key Laboratory of Opto-Electronics Information Technology, Ministry of Education, Tianjin 300072, China*

Abstract The algorithm based on a 3D label can obtain a more accurate sub-pixel disparity map in the stereo matching problem. To overcome the random initialization of 3D labels, we proposed a label initialization approach based on superpixel structure and triangulation. An initial 3D label is generated by triangulating the feature points retrieved from the superpixel boundary. To increase the efficiency of the graph cut approach in iterative optimization of 3D labels, we conduct optimization on the superpixel structure, adding the hypothesis of current label state throughout the iteration to expand the label candidates, which improves label search efficiency. Experiments on the Middlebury2014 dataset demonstrate that the proposed approach has a lower average error rate (8.31%) than the LocalExp algorithm (8.39%), and the average processing time for single image is ~70% that of the LocalExp algorithm.

Key words machine vision; stereo matching; 3D label; superpixel; triangulation; graph cut

1 引言

立体匹配作为双目立体视觉技术的关键环节,

已发展了多种立体匹配算法, 概括起来可分为两类, 一类是基于特征匹配的方法^[1], 另一类是基于深度学习的匹配方法^[2]。基于特征匹配方法的算法通过代

收稿日期: 2021-03-15; 修回日期: 2021-04-14; 录用日期: 2021-04-21

基金项目: 国家自然科学基金重点项目(61535008)

通信作者: *gebz@tju.edu.cn

价计算、代价聚合、视差计算、视差优化 4 个步骤求解匹配图像每个像素点的整数视差值,是 1D 视差。近年来,提出了基于 3D 标签的双目立体匹配方法, Bleyer 等^[3]提出了典型的 3D 标签算法——PatchMatch Stereo(PMS)算法,该算法以三个参数表示像素点的视差值与切平面法线方向,据此计算倾斜块匹配代价,在视差计算阶段搜索像素点的最优 3D 标签,找到平滑的、分段线性的视差图。随后的研究主要集中于提高 3D 标签的搜索速率与准确率。

Veldandi 等^[4]和 Bleyer 等^[5]引入分割信息,在更新标签时,假设处于同一分割区域的像素点来自于同一三维表面,缩小 3D 标签的搜索范围。Patch Match Filter (PMF) 算法^[6]与 PatchMatch-Based Superpixel Cut (PMSC) 算法^[7]以超像素作为搜索 3D 标签的基本单位,提高了搜索 3D 标签的效率。PMF 算法采用超像素结构与引导滤波器^[8]减少代价聚合时的冗余计算,大幅加快了 3D 标签的求解速率,但是其能量函数不惩罚视差的不平滑。

Olsson 等^[9]将二阶视差平滑引入 PMS 算法,构建 Markov Random Field (MRF) 模型,成功提高了 3D 标签估计的精度。Besse 等^[10]提出了 PatchMatch Belief Propagation (PMBP) 算法,该算法将 BP 算法与 PMS 算法结合优化目标函数,然而 BP 算法在计算时容易导致目标函数陷入局部最小值,与之相比,图割算法^[11]具备全局特性,不容易陷入局部最小值。PMSC 算法在每个超像素内部采用图割算法优化像素点的 3D 标签,最终融合不同超像素尺寸下的 3D 标签结果,但 PMSC 算法优化的中间结果不会反馈到标签的生成过程。Local Expansion (LocalExp) 算法^[12]构建多尺寸的图像网格结构时,在网格结构中执行了 PMS 算法的空间传播过程,在网格及其邻域内部采用图割算法优化 3D 标签,最后融合不同图像网格的 3D 标签结果。与 PMSC 算法相比,其优化后的中间结果会反馈标签生成过程,取得了优于 PMSC 算法的效果。高雅昆等^[13]在 LocalExp 算法的基础上针对无纹理和弱纹理区域的误匹配问题,提出了一种基于像素类别信息的 3D 标签立体匹配算法。

为提高 3D 标签的搜索速率与准确度,本文基于 LocalExp 算法提出了一种基于三角剖分与超像素结构的 3D 标签优化方法。为了验证所提方法的匹配精度与速度,在标准数据集上与 LocalExp 算法进行了对比,所提方法在保证匹配精度的前提下,

能够在更短时间内完成视差计算。

2 基本原理

根据 3D 标签的定义可知,某像素点 p 的视差由三个参数表示,形式^[14]为

$$d_p = a_p p_x + b_p p_y + c_p, \quad (1)$$

式中: (p_x, p_y) 为点 p 的图像坐标; d_p 为点 p 的视差; 点 p 的 3D 标签 l_p 为 (a_p, b_p, c_p) 。搜索 3D 标签 l_p 的过程也就是搜索该点的最优三维切平面的过程。

将倾斜窗口的匹配代价聚合与基于神经网络计算的匹配代价相结合^[7],得到

$$\Phi(\mathbf{p}, l_p) = \sum_{\mathbf{q} \in \mathbf{W}_p} \omega_{pq} S(\mathbf{q}, l_p), \quad (2)$$

式中: \mathbf{W}_p 为以点 p 为中心的 11×11 像素窗口内所有像素点; ω_{pq} 为自适应参数,根据颜色相似度描述 p 与 q 在同一平面的可能性; $S(\mathbf{q}, l_p)$ 为由神经网络计算的匹配代价。

$$S(\mathbf{q}, l_p) = \min [C_{\text{CNN}}(\mathbf{q}, \mathbf{q}'), \tau_{\text{CNN}}], \quad (3)$$

式中: $C_{\text{CNN}}(\mathbf{q}, \mathbf{q}')$ 表示以 \mathbf{q} 与 \mathbf{q}' 为中心的 11×11 像素大小的图像块之间由神经网络 (MC-CNN-acrt)^[15] 计算得到的匹配代价。采用引导滤波器^[8]计算自适应参数:

$$\omega_{pq} = \frac{1}{W^2} \sum_{k: (\mathbf{p}, \mathbf{q}) \in \mathbf{W}_k} \left[1 + (\mathbf{I}_p - \boldsymbol{\mu}_k)^\top (\boldsymbol{\Sigma}_k + \mathbf{e})^{-1} (\mathbf{I}_q - \boldsymbol{\mu}_k) \right], \quad (4)$$

式中: W 为窗口尺寸; \mathbf{I}_p 和 \mathbf{I}_q 为归一化的三通道颜色向量; $\boldsymbol{\mu}_k$ 与 $\boldsymbol{\Sigma}_k$ 表示在窗口 \mathbf{W}_k 中 \mathbf{I}_p 的均值与协方差矩阵; \mathbf{e} 是一个防止过拟合的正则化矩阵。

对于图像中的每个像素点,找到最优标签 l_{best} :

$$l_{\text{best}} = \arg \min \Phi(\mathbf{p}, l_p), \quad (5)$$

即可根据 (1) 式计算得到视差 d_p 。PMS 算法采用随机初始化、空间传播、平面细化等过程搜索像素点的最优 3D 标签。随机初始化过程为每个像素点指定一个随机的 3D 标签。空间传播是一个逐像素的迭代过程,在当前像素点与前一个像素点的标签中选择使 $\Phi(\mathbf{p}, l_p)$ 更低的标签,完成对所有像素点的遍历,即完成了整幅图的空间传播过程。平面细化过程采用二分法对当前标签进行幅度很小的调整,当新的 l_p 使得 $\Phi(\mathbf{p}, l_p)$ 更小时,则更新标签,最终得到更精细的 3D 标签。

LocalExp 算法仍采用了 PMS 算法随机初始化、空间传播、平面细化的框架,但 LocalExp 算法将

图割算法与空间传播过程结合,提出局部 α 扩展(alpha-expansion),采用图割优化的方法对3D标签进行全局优化。

基于LocalExp算法,提出了一种新的基于超像素结构与三角剖分初始化结果的3D标签计算方法。首先建立多尺寸超像素结构,利用超像素边界

采样得到特征点,对成功匹配的特征点进行三角剖分,采用三角形顶点的匹配信息代替随机初始化方案生成3D标签;使用图割算法迭代优化3D标签,通过对标签状态进行假设,得到更优的标签候选,提高求得像素点的最优3D标签的可能性。图1为所提方法的工作流程。

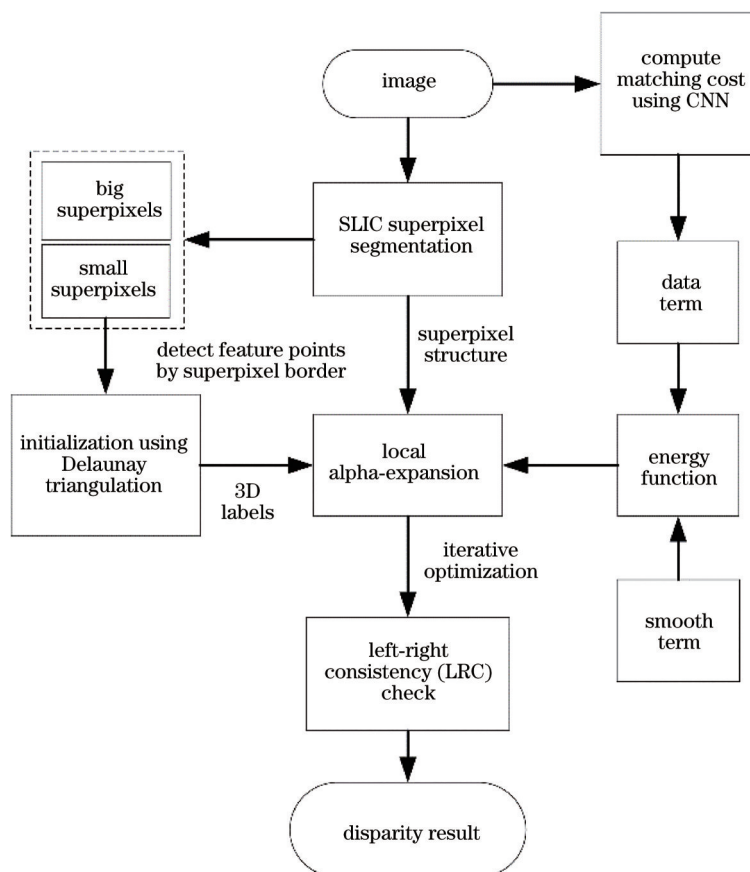


图1 所提方法的流程

Fig. 1 Flow chart of the proposed method

2.1 基于Delaunay三角剖分的3D标签初始化

根据像素点在图像中所处的位置决定其候选3D标签是一种缩小搜索范围的策略。Simple Linear Iterative Clustering (SLIC)算法^[16]采用线性迭代聚类的思想对像素点进行局部聚类,将图像中位置相邻且颜色、亮度、纹理等特征相似的像素点“聚合”为具有一定视觉意义的图像块——超像素。超像素可以作为3D标签求解的基本单位^[7]。

SLIC超像素分割会产生大量外观均匀的超像素 $S = \{s_1, s_2, s_3, \dots, s_n | \forall l \neq k, s_l \cap s_k = \emptyset\}$,其边界线对应着图像梯度的局部最大值。这些边界的交点是图像梯度或图像函数在多个方向上的局部极大值,这些交点即使在视点变换时也十分稳定,且可以被精确定位^[17]。将三个或者多个超像素边界的交点作

为候选点进行采样,通过遍历所有的超像素边界点,提取这些区域交点,只有在边界宽度为1 pixel时才可以获得唯一的区域交点。图2为基于超像素边界的特征点提取示例,图2(b)为超像素尺寸为8 pixel时的分割结果,图2(c)为图2(b)的局部放大图,这些特征点的八邻域包括至少三个超像素分割。

对所有提取到的候选点进行精确匹配。为确保匹配的效率,可以直接采用匹配代价作为特征点匹配的基准(亦可以采用其他基准),根据左右匹配一致性、最邻近距离比筛选匹配点。对保留下的匹配点进行Delaunay三角剖分,根据三角形三个顶点的匹配信息可以求得三角形区域内像素点的初始3D标签。三角剖分内的每一点 (p_x, p_y) 有

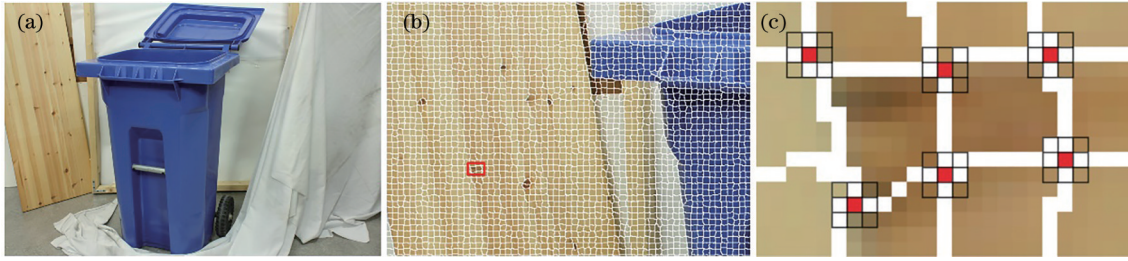


图 2 基于超像素边界的特征点提取示例。(a)原图;(b)SLIC超像素分割;(c)局部放大图

Fig. 2 Example of feature point extraction based on superpixel boundary. (a) Origin image; (b) SLIC superpixel segmentation; (c) local enlargement

$$d_i(\mathbf{p}) = a_i p_x + b_i p_y + c_i, \quad (6)$$

式中: i 是像素点 (p_x, p_y) 所属的三角形的序号。对于每个三角形,根据顶点 (p_x, p_y, d) 列出形如(6)式的线性方程组,求解方程组得到 (a_i, b_i, c_i) 。图3(a)、

(b)展示了最佳匹配点,图3(c)为三角剖分结果,每个三角形的顶点都在超像素的边界上,三角形内的像素点的初始3D标签为 $l_i(a_i, b_i, c_i)$ 。图4(d)为根据初始化的3D标签计算得到的视差图。

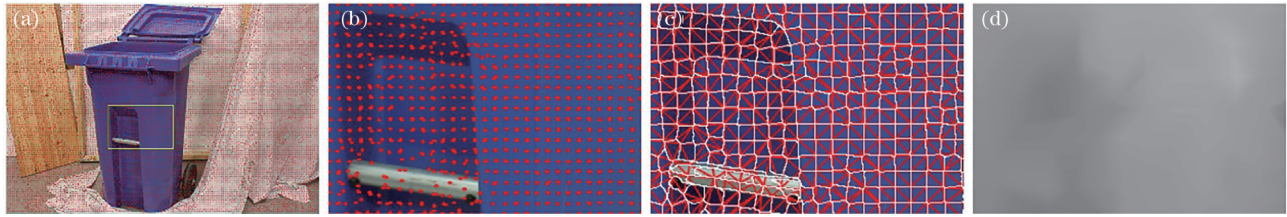


图 3 基于三角剖分的标签初始化。(a)特征点;(b)方框区域的放大图;(c)红线表示三角剖分,白线表示超像素边界;(d)方框区域的初始视差图

Fig. 3 Label initialization based on triangulation. (a) Feature points; (b) enlarged view of the box area; (c) the red line represents the triangulation, and the white line represents the superpixels' boundary; (d) initial disparity of the box area

2.2 基于MRF模型的能量函数

基于MRF模型构建全局能量函数,能量函数中的数据项衡量匹配点的相似性,平滑项衡量相邻像素之间3D标签的连续性。全局能量函数的表达式^[12]为

$$E(l) = E_{\text{data}}(l) + \lambda E_{\text{smooth}}(l), \quad (7)$$

式中:常数 λ 用于平衡数据项与平滑项的权重。数据项定义为

$$E_{\text{data}}(l) = \sum_{\mathbf{p} \in \Omega} \Phi(\mathbf{p}, l_{\mathbf{p}}), \quad (8)$$

式中: Ω 表示整个图像区域。平滑项定义为

$$E_{\text{smooth}}(l) = \sum_{\mathbf{p} \in \Omega} \sum_{\mathbf{q} \in N_{\mathbf{p}}} \psi_{pq}(l_{\mathbf{p}}, l_{\mathbf{q}}), \quad (9)$$

式中: $N_{\mathbf{p}}$ 表示点 \mathbf{p} 的邻域像素。基于曲率的二阶平滑正则化表达式^[9, 12]为

$$\psi_{pq}(l_{\mathbf{p}}, l_{\mathbf{q}}) = \max(\omega'_{pq}, \epsilon) \min[\psi'_{pq}(l_{\mathbf{p}}, l_{\mathbf{q}}), \tau_{\text{dis}}], \quad (10)$$

式中:平滑项的权重 $\omega'_{pq} = \exp\left(\frac{\|I_{\mathbf{p}} - I_{\mathbf{q}}\|_1}{\gamma}\right)$, γ 为常数; ϵ 为权重的下限值; τ_{dis} 是一个截断值,允许在视差边界存在跳变。

$$\psi'_{pq}(l_{\mathbf{p}}, l_{\mathbf{q}}) = |d_{\mathbf{p}}(l_{\mathbf{p}}) - d_{\mathbf{p}}(l_{\mathbf{q}})| + |d_{\mathbf{q}}(l_{\mathbf{q}}) - d_{\mathbf{q}}(l_{\mathbf{p}})|, \quad (11)$$

式中:第一项,通过像素点 \mathbf{p} 处 $l_{\mathbf{p}}$ 和 $l_{\mathbf{q}}$ 所得视差值的不同,衡量 $l_{\mathbf{p}}$ 和 $l_{\mathbf{q}}$ 的不连续性,第二项在点 \mathbf{q} 处进行相同计算。(11)式满足优化条件^[12]:

$$\psi(\alpha, \alpha) + \psi(\beta, \gamma) \leq \psi(\beta, \alpha) + \psi(\alpha, \gamma). \quad (12)$$

2.3 基于局部超像素标签传播的图割优化方法

常见3D标签的图割优化都在图像块而非全图上进行,局部 α 扩展计算该图像块中所有像素点的能量函数,选择能量最小对应的3D标签组合,迭代地解决图像块上的二元标签 $\{l'_p, \alpha\}$ 的标记更新问题,迭代公式为

$$l^{(t+1)} = \arg \min E(l' | l'_p \in \{l'_p, \alpha\}). \quad (13)$$

基于图像的超像素结构,在当前超像素和相邻的超像素组成的区域内执行局部 α 扩展,即可更新该区域的3D标签。由(13)式可知,执行迭代公式,需要不断采样新标签 α 。对当前标签添加扰动 Δ ,生成新标签 α :

$$\alpha = (a_p, b_p, c_p) + \Delta. \quad (14)$$

将点 p 的 3D 标签转化为点 (p_x, p_y, d) 与法向量 $\mathbf{n} = (n_x, n_y, n_z)$ 表示, 即可对标签代表的平面进行更平滑的采样, 其中 $a_p = -\frac{n_x}{n_z}$, $b_p = -\frac{n_y}{n_z}$, $c_p = \frac{n_x p_x + n_y p_y + n_z d}{n_z}$ 。

Δ_d^{\max} 为对 d 随机搜索时的最大步长, Δ_n^{\max} 为对 \mathbf{n} 随机搜索时的最大步长, Δ_d 与 Δ_n 在 $[-\Delta, \Delta]$ 范围内随机取值。文献 [3, 7, 12] 使用 $d + \Delta_d$ 与 $u(n_x + \Delta_n, n_y + \Delta_n, n_z + \Delta_n)$ 进行 3D 标签求精, 其中 $u(\cdot)$ 表示求单位向量, 迭代中采用二分法缩小搜索范围, 随机性较大。

生成新标签 α 时, 当前两个参数 (d, \mathbf{n}) 可能具备以下状态 [18]: 两个都是最优解、两个都不是最优解、两个中的某一个是最优解。将随机扰动 Δ_d 与当前可能为理想法线向量的 \mathbf{n} 组合作为新的 3D 标签 (Δ_n 与当前理想视差 d 的组合亦然), 增大采样理想

标签的可能性。每次平面细化的最优标签从以下组合选出:

$$L_{\text{candidate}} = \{(d, \mathbf{n}), (d_{\text{prt}}, \mathbf{n}), (d, \mathbf{n}_{\text{prt}}), (d_{\text{prt}}, \mathbf{n}_{\text{prt}})\}, \quad (15)$$

式中: $d_{\text{prt}} = d + \Delta_d$; $\mathbf{n}_{\text{prt}} = u(n_x + \Delta_n, n_y + \Delta_n, n_z + \Delta_n)$ 。

当前超像素为 s_{center} , 则 $S_{\text{neighborhood}} = s_{\text{center}} \cup N(s_{\text{center}})$, $N(s_{\text{center}})$ 表示与 s_{center} 相邻的超像素。根据当前的 s_{center} 及其 $L_{\text{candidate}}$ 生成一个新的标签 α , 在 $S_{\text{neighborhood}}$ 区域内迭代使用 (13) 式更新标签, 这个过程将 s_{center} 的 α 标签向其余超像素传播。局部 α 扩展过程可以使用图 4 表示, 该 $S_{\text{neighborhood}}$ 中共包含 9 个超像素。对整个图像区域中不同的超像素并行地进行该过程, 即可完成一次对整幅图的标签更新。

为了并行过程满足 (12) 式, 局部 α 扩展的过程不能同时处理有重叠的区域, 所以需要将每个超像素的 $S_{\text{neighborhood}}$ 注册为一个新的超像素, 根据原有的边界关系找到 $S_{\text{neighborhood}}$ 的边界关系, 对这些更大的超像素 $S_{\text{neighborhood}}$ 采用四色定理分组, 保证同时处理的 $S_{\text{neighborhood}}$ 区域不相互重叠。

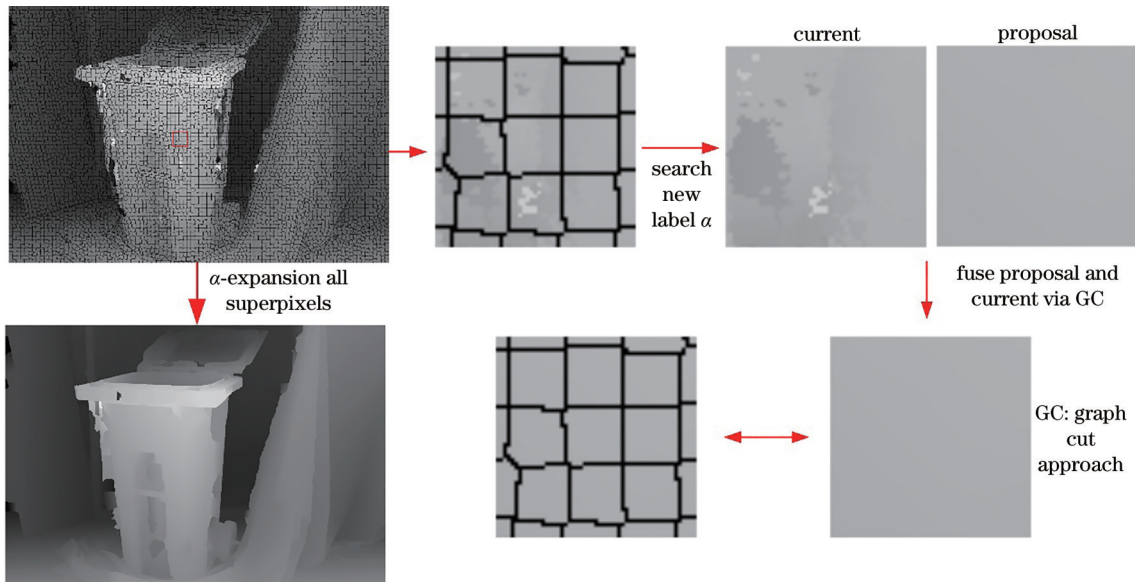


图 4 超像素结构下的 α 扩展

Fig. 4 α -expansion in superpixel structure

3 实验结果及分析

计算机实验平台配置为 Intel Core i7-8700 处理器, 主频 3.2 GHz, Windows10 家庭版 64 位操作系统, 内存 8 GB。超像素尺寸设置为 $\left\{ \frac{\min(W, H)}{100}, \frac{\min(W, H)}{50}, \frac{\min(W, H)}{25} \right\}$ pixel, 其中 W, H 为图像的水平像素数与垂直像素数。根据文

献 [12] 设置参数, $\lambda = 0.5$, $W = 20$, $\tau_{\text{CNN}} = 0.5$, $\gamma = 10$, $\tau_{\text{dis}} = 1.0$, $\epsilon = 0.01$, N_p 为八邻域, (13) 式的迭代次数为 5。

为了验证所提算法有效性, 选取了 Middlebury2014 数据集中的 12 幅提供视差真值的立体图像对作为实验对象。所提算法与 LocalExp 算法在每轮迭代结束时记录目标函数值与误匹配率 (error rate)。采用归一化后的目标函数值绘制收

敛曲线;采用默认标准判定误匹配,且仅统计非遮挡区域的像素点,与真值间误差大于 2 个像素即被认为是误匹配。为了减小随机性的影响,每对图像独立运行 10 次取平均值。

表 1 为所提算法、LocalExp 算法、PMSC 算法在所有 12 组实验图像下的误匹配率结果,后处理过程采用标准的左右一致性检测方法(LRC)。所提算法的平均误匹配率(8.31%)与 LocalExp 算法的平均误匹配率(8.39%)相比更低,且算法消耗的时间更少。表 1 中粗体表示算法在该组图像中的误匹配率最低,所提算法在 Piano、Recycle、Vintage 图像上优势更为明显,其中所提算法对 Vintage 图像的误匹配率(9.7%)与 LocalExp 算法(12.54%)相差最大。图 5 第 4 行为两算法对 Vintage 图像的结果对比,在细节放大区域与图像左上部分的墙面区域所提算法表现更优,因为这些区域较其他区域纹理更少且平面更多。而 Piano 与 Recycle 图像也大量存在类似区域。PMSC 算法同样构建了超像素结构,而与 PMSC 算法的平均误匹配率相比,所提算法的平均误匹配率更低;由于 PMSC 算法没有开源代码,无法在相同的实验平台下复现其结果,只能根据 Middlebury 网站上的结果进行比较。PMSC 算法仅在 Teddy 图像表现更优,所提算法在该图像上的表现优于 LocalExp 算法。

表 1 不同算法在 12 组实验图像下的误匹配率结果
Table 1 Error rates of different algorithms in 12 experimental images unit: %

Image	Proposed algorithm	LocalExp	PMSC
Adirondack	1.18	1.22	1.46
ArtL	4.82	4.65	3.69
Jadeplant	9.85	8.95	11.2
Motorcycle	3.49	3.54	3.68
Piano	8.76	9.82	11.9
Pipes	4.29	4.11	5.25
Playroom	9.27	9.35	12.6
Playtable	7.28	6.02	8.03
Recycle	6.32	6.66	7.58
Shelves	31.8	30.78	31.6
Teddy	2.89	3.13	2.67
Vintage	9.7	12.54	17.9
Average	8.31	8.39	9.80

图 5 为所提算法与 LocalExp 算法在 Adirondack、Recycle、Piano、Vintage 的视差图结果对比。误匹配图中错误的视差值为黑色,正确的视差值为白色,灰色表示遮挡区域的像素点,不纳入统计范围。图 5 中的插图部分为细节对比区域,在一部分纹理较低的区域,所提算法的视差图结果更好。

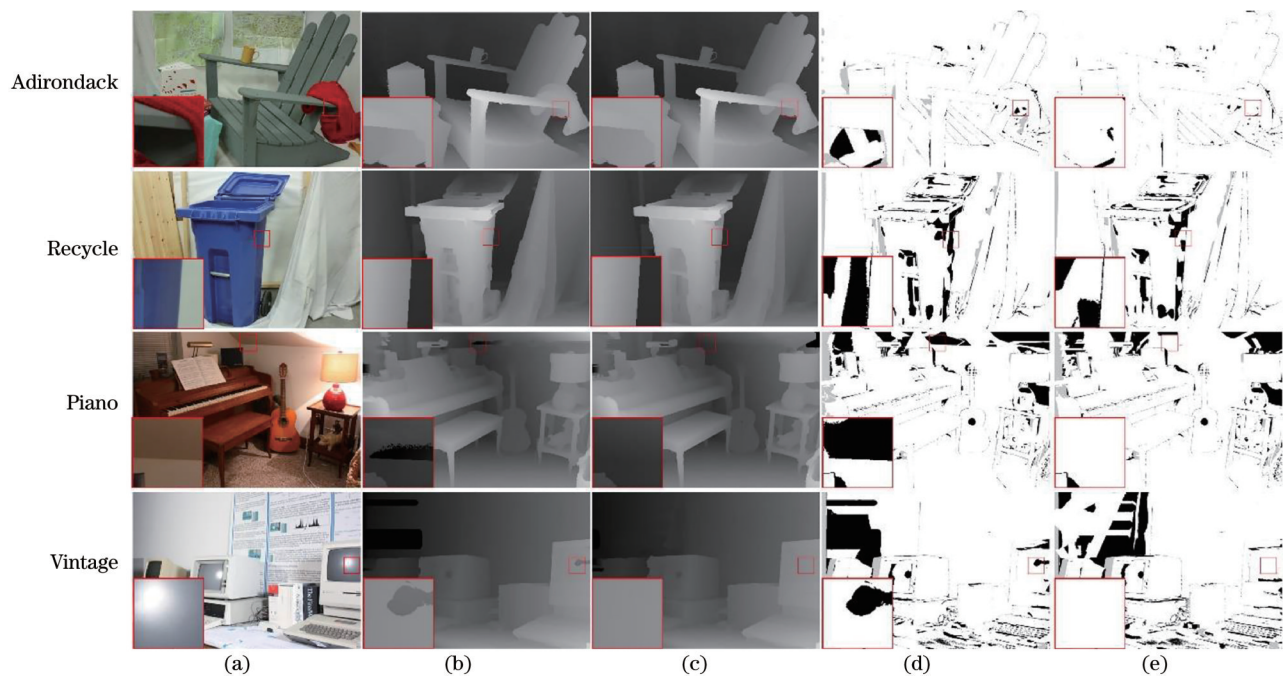


图 5 不同算法的视差图和误匹配图。(a)左图;(b) LocalExp 算法的视差图;(c)所提算法的视差图;(d) LocalExp 算法的误匹配图;(e)所提算法的误匹配图

Fig. 5 Disparity maps and error maps of different algorithms. (a) Left image; (b) disparity map of LocalExp; (c) disparity map of proposed algorithm; (d) error map of LocalExp; (e) error map of proposed algorithm

为了与 LocalExp 算法比较,设定图割算法的迭代次数为 5,结果如图 6 所示。由图 6 的收敛曲线可以看出,所提算法能更快完成 3D 标签初始化,以 Recycle 的结果为例,LocalExp 算法在约 50 s 时完成初始化,所提算法在 30 s 时完成初始化;在 5 次迭代结束后,所提算法的误匹配率低于 LocalExp 算法,此时所消耗时间约为 LocalExp 算法的 70%。

图 7 中,设定算法运行时间为 430 s 左右。在算法运行时间相近的情况下,所提算法的图割优化次数更多。采用(15)式改进二分法生成标签后,单次迭代耗费的时间更少,且最终优化效果更好。

图 8 为实际远距离拍摄的混凝土模型的立体图像对的视差图对比。根据右图与视差关系重构得到的左图与原左图应该十分相近,光度一致性损失^[19]可以量化相近的程度:

$$L_p = \frac{1}{N} \sum \delta \|I - I'\|_1, \quad (16)$$

式中: N 为像素点个数; δ 用于避免非感兴趣区域的异常值; I 与 I' 为原图和重构图。光度一致性损失越小,说明两图越相近。如表 2 所示,两种算法所得重构图像的光度一致性损失十分接近,但是耗费时间上所提算法更少。

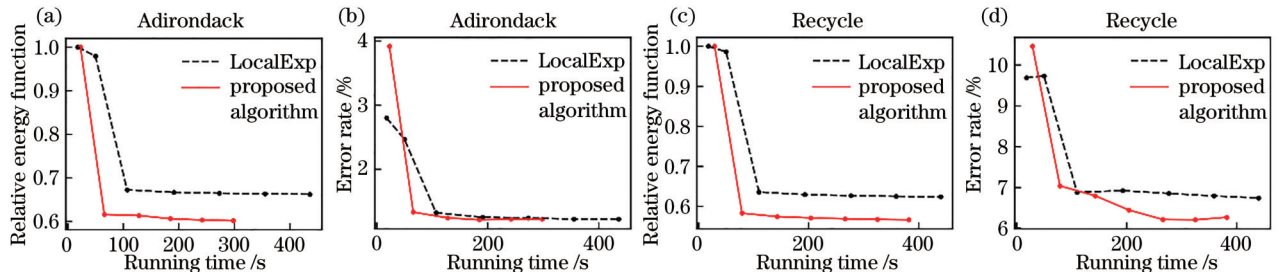


图 6 在 Adirondack 与 Recycle 图像上的实验结果。(a)(c)相对能量函数值;(b)(d)误匹配率

Fig. 6 Results of experiments on Adirondack and Recycle images. (a)(c) Relative energy function value; (b)(d) error rate

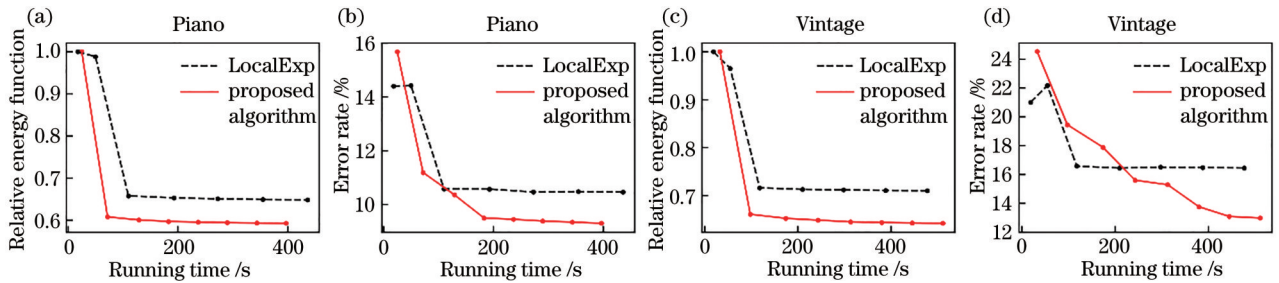


图 7 在 Piano 与 Vintage 图像上的实验结果。(a)(c)相对能量函数值;(b)(d)误匹配率

Fig. 7 Results of experiments on Piano and Vintage images. (a)(c) Relative energy function value; (b)(d) error rate

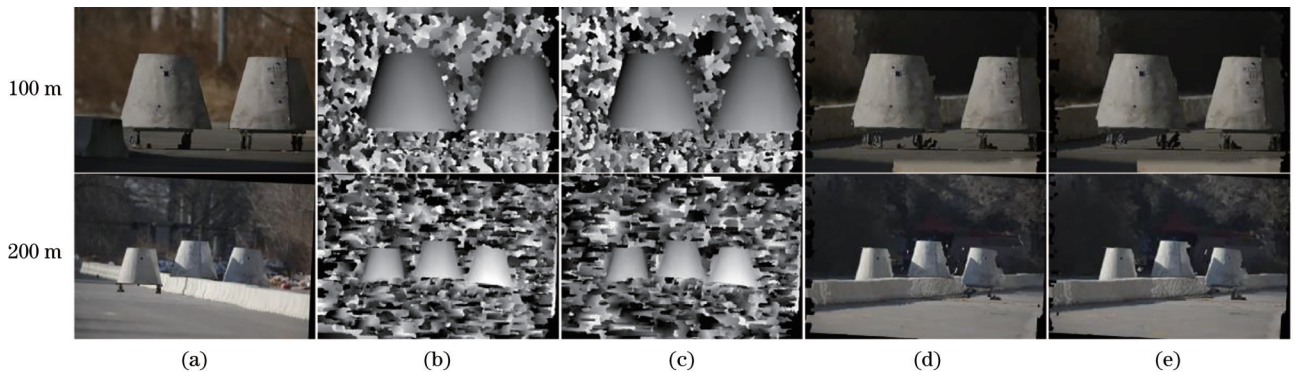


图 8 实际拍摄图像。(a)左图;(b)LocalExp 算法的视差图;(c)所提算法的视差图;(d) LocalExp 算法重构的左图;(e)所提算法重构的左图

Fig. 8 Real images. (a) Left image; (b) disparity map of LocalExp; (c) disparity of proposed algorithm; (d) reconstructed left image of LocalExp; (e) reconstructed left image of proposed algorithm

表 2 不同算法在实际拍摄图像上的实验结果

Table 2 Experimental results of different algorithms on real images

Image	Resolution	Photometric consistency loss		Running time /s	
		LocalExp	Proposed algorithm	LocalExp	Proposed algorithm
100 m	2880×1920	1583.8	1591.5	3630	2300
200 m	2880×1920	2602.8	2591.3	2879	2088

4 结 论

提出一种基于超像素与三角剖分的 3D 标签全局立体匹配方法。首先采用超像素与三角剖分构建初始化的 3D 标签,对 3D 标签当前状态进行假设,改进二分法标签采样,最后通过对超像素及其邻域进行图割算法求解获得最优 3D 标签。在初始化与迭代过程中改进了 3D 标签的生成,加快了算法迭代的速度。在处理 Middlebury2014 数据集的训练集与实际拍摄图片的实验中,与 LocalExp 算法相比,相同迭代次数下,所提算法所消耗的时间更少,能量函数的收敛速度更快,在误匹配率方面两者十分接近。

参 考 文 献

- [1] Ma R H, Zhu F, Wu Q X, et al. Dense stereo matching algorithm based on image segmentation[J]. *Acta Optica Sinica*, 2019, 39(3): 0315001.
马瑞浩,朱枫,吴清潇,等.基于图像分割的稠密立体匹配算法[J].*光学学报*,2019,39(3):0315001.
- [2] Wang X C, Liu H H, Niu Y M. Binocular stereo matching by combining multiscale local and deep features[J]. *Acta Optica Sinica*, 2020, 40(2): 0215001.
王旭初,刘辉煌,牛彦敏.融合多尺度局部特征与深度特征的双目立体匹配[J].*光学学报*,2020,40(2):0215001.
- [3] Bleyer M, Rhemann C, Rother C. PatchMatch stereo-stereo matching with slanted support windows [C]//*Proceedings of the British Machine Vision Conference 2011*, August 29-September 2, 2011, Dundee. London: BMVA Press, 2011: 14.
- [4] Veldandi M, Ukil S, Govindarao K. Robust segment-based stereo using cost aggregation[C]//*Proceedings of the British Machine Vision Conference 2014*, September 1-5, Nottingham. London: BMVA Press, 2014: 1-11.
- [5] Bleyer M, Rother C, Kohli P. Surface stereo with soft segmentation[C]//*2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE Press, 2010: 1570-1577.
- [6] Lu J B, Yang H S, Min D B, et al. Patch match filter: efficient edge-aware filtering meets randomized search for fast correspondence field estimation[C]//*2013 IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2013, Portland, OR, USA. New York: IEEE Press, 2013: 1854-1861.
- [7] Li L C, Zhang S L, Yu X, et al. PMSC: PatchMatch-based superpixel cut for accurate stereo matching[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(3): 679-692.
- [8] 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.
- [9] Olsson C, Ulén J, Boykov Y. In defense of 3D-label stereo[C]//*2013 IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2013, Portland, OR, USA. New York: IEEE Press, 2013: 1730-1737.
- [10] 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.
- [11] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(11): 1222-1239.
- [12] Taniar T, Matsushita Y, Sato Y, et al. Continuous 3D label stereo matching using local expansion moves [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(11): 2725-2739.
- [13] Gao Y K, Liu T, Li H B, et al. Stereo matching algorithm based on pixel category optimized PatchMatch[J]. *Acta Optica Sinica*, 2019, 39(7): 0715006.
高雅昆,刘涛,李海滨,等.基于像素类别优化的 PatchMatch 立体匹配算法[J].*光学学报*,2019,39(7):0715006.
- [14] Zhou J L, Chen Y, Wu C, et al. Binocular stereo matching algorithm based on labeled matching region correction[J]. *Pattern Recognition and Artificial*

- Intelligence, 2020, 33(8): 681-691.
- 周佳立, 陈育, 吴超, 等. 基于标签化匹配区域校正的双目立体匹配算法[J]. 模式识别与人工智能, 2020, 33(8): 681-691.
- [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] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.
- [17] Mustafa A, Kim H, Imre E, et al. Segmentation based features for wide-baseline multi-view reconstruction[C]//2015 International Conference on 3D Vision, October 19-22, 2015, Lyon, France. New York: IEEE Press, 2015: 282-290.
- [18] Schönberger J L, Zheng E L, Frahm J M, et al. Pixelwise view selection for unstructured multi-view stereo[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9907: 501-518.
- [19] Yang G R, Zhao H S, Shi J P, et al. SegStereo: exploiting semantic information for disparity estimation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 660-676.