

基于像素类别优化的 PatchMatch 立体匹配算法

高雅昆, 刘涛, 李海滨*, 张文明

燕山大学工业计算机控制工程河北省重点实验室, 河北 秦皇岛 066004

摘要 基于 PatchMatch(同时估计像素点的视差和法向量的 3D 标签)的方案已经在立体匹配中取得高精度的亚像素视差,但该类方法无法有效解决图像无纹理区域的错误匹配。针对这一问题,对 LocalExp(local expansion move)算法进行了改进,并提出一种融合多维信息的自适应像素类别优化的立体匹配算法。该方法设计了一种交叉窗口,在窗口内基于颜色与颜色的自相关信息构建相关权重,并利用约束函数剔除匹配代价中的离群值;在 PatchMatch 的标签初始化阶段增加约束机制,改进视差标签的建议生成机制,并利用基于局部扩张运动的优化方法求解标签值;利用基于像素类别的填充策略进行视差优化。实验结果表明所提算法能够在 Middlebury 数据集上取得较低的匹配误差。

关键词 机器视觉; 立体匹配算法; PatchMatch; 像素分类; 视差优化

中图分类号 TP911.73

文献标识码 A

doi: 10.3788/AOS201939.0715006

Stereo Matching Algorithm Based on Pixel Category Optimized Patch Match

Gao Yakun, Liu Tao, Li Haibin*, Zhang Wenming

Key Laboratory of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao, Hebei 066004, China

Abstract PatchMatch-based algorithms that simultaneously estimate the disparities and normal unit of a disparity plane have achieved highly accurate sub-pixel disparities in the stereo matching problem; however, this kind of methods can not effectively deal with error matching in the non-texture regions of image. To solve this problem, we improve the LocalExp(local expansion move)algorithm and present a new stereo matching algorithm integrating multidimensional information for adaptive pixel category optimization. First, a crossover window is designed, the color and color self-correlation information in the window are used to establish the weight, and the restrained function is utilized to eliminate the outliers in the matching cost. Second, the constraint mechanism is added to the label initialization procedure, the proposal generation mechanism is modified, and the local expansion movement algorithm is used to optimize the label values. Finally, the pixel category information-based filling strategy is used to refine the disparity. The experimental results show that the proposed method can obtain a low matching error on the Middlebury dataset.

Key words machine vision; stereo matching algorithm; PatchMatch; pixel category; disparity refinement

OCIS codes 150.1135; 330.1400; 100.6890

1 引 言

双目立体匹配技术作为三维重建研究领域的核心技术之一,已广泛应用于无人驾驶、机器人导航定位、增强/虚拟现实、物体识别与跟踪、地理信息系统(GIS)等领域^[1]。在正向平行表面的视差一致性假设下(窗口内视差相等),对于在斜面及曲面区域的视差计算,传统匹配算法会产生阶梯效应。针对这

一问题,Bleyer 等^[2]提出了基于 3D 标签的倾斜窗口假设(窗口内视差值存在渐变),利用 PatchMatch(patch match)^[3]算法中图像修复的思想来解决立体匹配中视差问题,通过对像素点 3D 标签分别进行随机初始化、邻域传播、左右视图传播、随机搜索以及平面细化等来估计视差,并将该方法称为 PMS(patch match stereo)。3D 标签相对于单一视差值,增加了像素点的切平面法向量信息,更能表征像

收稿日期: 2019-01-03; 修回日期: 2019-03-01; 录用日期: 2019-04-01

* E-mail: hbli@ysu.edu.cn

素点所在平面的空间信息。但采用 PatchMatch 算法进行立体匹配时存在以下问题:1)随机对每个像素点进行参数初始化,初始化过程并不能保证每个像素都有一个几何意义下可行的视差平面参数;2)平面细化过程使用 Luus 和 Jaakola^[4]提出的优化方法来最小化代价函数,并不能保证找到代价的局部最小值;3)受代价函数或能量函数的限制,该方法不能有效解决弱或无纹理区域的错误匹配。

不同学者针对这些问题提出了不同的立体匹配解决方案。Besse 等^[5]将 PMS 算法和粒子置信传播算法整合到统一的结构下,提出了一种基于置信传播 BP(belief propagation)的加速全局立体匹配算法(PMBP)。该方法在 BP 框架下将粒子采样标签更新和 PatchMatch 算法的邻域传播相融合解决了连续性标签处理问题。同时,该方法在 PMS 算法的代价函数基础上,考虑邻域平面的平滑项约束构造能量函数,得到了比 PMS 算法更准确的视差,但 BP 作为优化算法会陷入局部极小值。Lu 等^[6]提出的 PMF(patch match filter)算法采用超像素作为匹配单元代替 PMS 算法中的固定尺寸窗口,在代价函数中引入由粗到精的尺度一致性约束关系以缓解无纹理区域中的错误匹配,并通过在 PMS 标签搜索策略中增加一种有效的超像素相似性计算方案来搜索候选标签。该方法比 PMS 算法的运行速度快,但在能量函数中无正则化约束视差平滑变化。Heise 等^[7]提出的 PM-huber 算法将 PatchMatch 随机搜索环节同 Huber 算子约束的显式变分平滑方法相结合,构建能量函数的视差约束平滑项,并在 PMS 视差标签初始化时,根据经验限定标签取值范围,该方案相对于 PMF 能获得较好的精度。Li 等^[8]提出的 PMSC(PatchMatch-based superpixel cut)算法步骤为:1)利用超像素单元构建顶层匹配代价,以 Žbontar 等^[9]提出的卷积神经网络(CNN)方法计算得到的像素块相似度作为底层代价,并将两层代价进行组合以构建匹配能量函数,同时将基于曲率的二阶平滑项用于视差平滑;2)在视差优化过程中,在超像素间使用 PatchMatch 算法更新当前超像素区域标签,在超像素内利用 GC(graph cut)^[10]优化每一像素点的最优标签值;3)融合超像素块标签和像素点标签以获得最终视差图。由于利用 GC 算法在节点间的优化可同步改善连接节点状态,因此该方案可有效地减少图像在弱/无纹理区域的误匹配。不同于 PMSC 算法, Li 等^[11]提出的 3DMST(3D minimum spanning tree)匹配方案的步

骤为:1)采用分割树分割^[12]代替超像素分割,将图像分成多个最小生成树图结构;2)使用 Yang^[13]提出的代价聚合方法计算每个 MST 上像素点的匹配代价值;3)基于 PatchMatch 算法优化每个 MST 标签状态,并求解像素点标签。该方案进一步降低了图像在无纹理区域的匹配错误率。Tanai 等^[14]提出的局部扩展移动优化算法 LocalExp 与 PMSC、3DMST 算法的整体思想相似,但存在几点区别:1)构造由小到大的多尺寸规则窗口金字塔模型,在不同尺寸的窗口内计算匹配代价,进行标签传播;2)针对每一个窗口单元构造扩散区域用于 PatchMatch 的空间传播,构造滤波区域用于基于快速代价滤波的代价聚合方法中,求取粗略视差值;3)利用 GC 优化更新细化局部窗口内像素点的标签。目前该方法的结果在 Middlebury 网站上排名前列。

虽然上述匹配算法针对 PatchMatch 算法的局限性给出了不同的解决方法,但这些方法仍不能较好地解决图像中无纹理和弱纹理区域的误匹配问题。针对这一问题,本文基于 LocalExp 算法提出了一种新的基于像素分类的立体匹配解决方案。

2 基本原理

LocalExp 算法的流程如下:1)随机初始化图像的 3D 标签图,为每个像素点 $p(p_x, p_y)$ 随机分配单位表面法向量 $\mathbf{n}_p(n_x, n_y, n_z)$ 和随机视差 z , 用于像素点所在参数平面 $L_p(a_p, b_p, c_p)$ 的初始化,其转换关系为 $a_p = -n_x/n_z, b_p = -n_y/n_z, c_p = -(n_x p_x + n_y p_y + n_z z)/n_z$, 其中 p_x, p_y 为像素点 p 的坐标分量, n_x, n_y, n_z 为像素点 p 的 3D 单位法向量分量, a_p, b_p, c_p 为点 p 所在视差参数平面系数;2)使用 3 种不同尺寸的窗口构建局部窗口,并将其作为优化单元,利用颜色和梯度信息计算各窗口中像素点在当前标签下的匹配代价;3)利用 PatchMatch 算法在不同窗口之间进行 3D 视差标签传播以搜寻最优标签;4)基于 GC 算法优化窗口内像素点的标签。

图 1 给出 LocalExp 算法和本文改进的算法内容,其中黑色字体为 LocalExp 算法的结构图,而黑色加粗字体为本文改进的环节,图中 K_s, K_b 为窗口分组数量。具体改进内容如下。

1)提出一种能够提取更多局部窗口用于计算像素点匹配代价和共享标签的交叉窗口模型,将颜色内相关性信息作为权重构建因素,在局部窗口内对匹配代价进行多维权重滤波运算,构建能量函数;同时,用 Geman-McClure 函数^[15]代替传统的截断

函数处理异常值,以保证初始代价值的稳定与平滑。

2) 采用具有约束信息的视差标签初始化方法 (IPMS),使初始化视差标签值更合理有效。

3) 根据像素分类信息改进 LocalExp 算法标签建议机制,即在迭代过程中通过多种机制生成候选标签集合用于迭代更新视差标签。

4) 在像素点视差值优化处理方面,设计了一种基于像素分类的视差优化方法,利用交叉线搜索同标签下可靠的视差标签对图像弱纹理和无纹理区域中的视差误匹配点进行错误填充;采用最近邻域填充方法对边缘、纹理区域以及遮挡区域中的错误点进行填充。

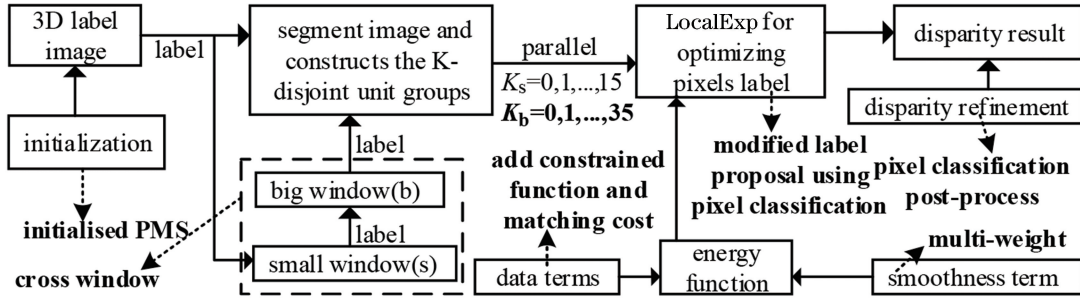


图 1 算法内容细节对比

Fig. 1 Detail comparison of algorithm content

2.1 多维权重能量函数

利用在 PairwiseMRF 模型下定义的最小化能量函数来估计像素点的 3D 视差标签。能量函数表达式为

$$E(L) = \sum_{p \in \Omega} \phi_p(L_p) + \lambda \sum_{(p,q) \in N_4} \varphi_{pq}(L_p, L_q), \quad (1)$$

式中: Ω 为像素点集合; λ 为平滑系数; N_4 为四邻域关系; L 为 3D 标签量集合; $\phi_p(L_p)$ 为像素点 p 的数据项; $\varphi_{pq}(L_p, L_q)$ 为像素点 p, q 间的平滑项,其中 L_p 为像素点 p 的 3D 标签值, L_q 为像素点 q 的 3D 标签值, p 为窗口中心像素点, q 代表窗口中的其他像素点。 $\phi_p(L_p)$ 用于计算匹配像素间相似度的数据项,可表示为

$$\phi_p(L_p) = \frac{1}{Z_p} \sum_{p,q \in W_p} \omega_{pq} \Gamma(q, L_p), \quad (2)$$

式中: $\Gamma(q, L_p)$ 为像素点 q 的标签值为 $L_p(a_p, b_p, c_p)$ 时,以 p 为中心的窗口内的像素点 q 与待匹配图对应点 q^* 的匹配代价值; W_p 为以点 p 为中心的滤波窗口; Z_p 为权重归一化参数; ω_{pq} 为权重,可表示为

$$\omega_{pq} = \frac{1}{|W^2|} \cdot \sum_{(p,q) \in W_k} [\mathbf{I} + (\mathbf{I}_p - \boldsymbol{\mu}_k)^T \cdot (\boldsymbol{\Sigma}_k + \mathbf{e})^{-1} (\mathbf{I}_q - \boldsymbol{\mu}_k)]. \quad (3)$$

其中 $\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{e}$ 分别为局部窗口 W_k 内的颜色均值、协方差和避免过拟合的矩阵^[16]; W 为窗口尺寸; \mathbf{I}_p 为 p 的颜色向量; \mathbf{I}_q 为 q 的颜色向量。

利用由颜色分量、图像水平和垂直梯度分量组合的相似度函数 $\Gamma(q, L_p)$ 计算像素点间的匹配代

价。多代价值融合的策略对一些变化区域的匹配精度有小幅提升,并且能够在边缘处得到更加准确和平滑的视差图。

由颜色分量、图像水平和垂直梯度分量组合的相似度函数 $\Gamma(q, L_p)$ 可表示为

$$\Gamma(q, L_p) = (1 - \alpha) G_{\text{col}}(\|\mathbf{I}_q - \mathbf{I}_{q^*}\|_1, \tau_{\text{col}}) + 0.5\alpha [G_{\text{grad}_x}(\|\nabla_x I_q - \nabla_x I_{q^*}\|_1, \tau_{\text{grad}_x}) + G_{\text{grad}_y}(\|\nabla_y I_q - \nabla_y I_{q^*}\|_1, \tau_{\text{grad}_y})], \quad (4)$$

式中: $\nabla_x I$ 表示灰度图像在梯度算子滤波处理后水平方向灰度梯度; $\nabla_y I$ 为垂直方向灰度梯度; G_{col} 为像素点颜色信息的约束函数; G_{grad_x} 为水平方向梯度约束函数; G_{grad_y} 为垂直方向梯度约束函数; $\tau_{\text{col}}, \tau_{\text{grad}_x}, \tau_{\text{grad}_y}$ 分别表示匹配代价的不同分量的调节参数; α 为比例因子。在(4)式中,可以通过预定义参数 τ 调整异常值控制范围,如图 2 所示,原始代价 c 经过变化后趋于稳定、平滑。LocalExp 在匹配代价

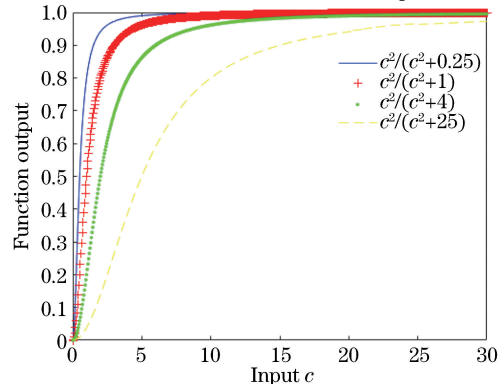


图 2 Geman-McClure 函数不同参数对比曲线

Fig. 2 Comparison curves for different parameters of Geman-McClure function

计算中用线性截断处理异常点代价值,而本文用

$$G(c, \tau) = \frac{c^2}{c^2 + \tau^2} \quad (5)$$

表示的约束函数替代线性截断函数,当代价值 c 超过某值时,函数的输出将平滑下降。

能量函数中用于约束视差变化的平滑项 $\varphi_{pq}(L_p, L_q)$ 由 Olsson 等^[17] 提出的二阶曲率正则化约束关系表示为

$$\varphi_{pq}(L_p, L_q) = \max(\omega_{pq}^*, \epsilon) \min[\psi(L_p, L_q), \epsilon_{\text{dis}}], \quad (6)$$

式中: ω_{pq}^* 代表平滑项的权重; $\psi(L_p, L_q)$ 为视差惩罚函数; ϵ 为下限值,能够减小权重受噪声的影响; ϵ_{dis} 为最大视差跳变量。

相对于 LocalExp, 本文将颜色内相关性分量 I_{Cn} 和颜色分量作为平滑项的权重 ω_{pq}^* 组合, 组合后的权重关系反映了色彩内部的一些特征, 更能表征出相邻像素颜色的一致性和差异性。 I_{Cn} 和 ω_{pq}^* 可分别表示为

$$I_{\text{Cn}}(p) = [I_n^r(p) - I_n^g(p), I_n^g(p) - I_n^b(p), I_n^b(p) - I_n^r(p)], \quad (7)$$

$$\omega_{pq}^* = \exp(-\{\beta[\|I_{\text{Cn}}(p) - I_{\text{Cn}}(q)\|_1/\gamma_n]\} - (1-\beta)[\|I_{\text{Cn}}(p) - I_{\text{Cn}}(q)\|_1/\gamma_{\text{cn}}]), \quad (8)$$

式中: $I_n^r(p)$ 、 $I_n^g(p)$ 、 $I_n^b(p)$ 为像素点 p 的 RGB (red-green-blue) 强度值; $(p, q) \in N_4$, N_4 为 4 邻域; β 为组合因子; $\gamma_n, \gamma_{\text{cn}}$ 为颜色相似计算的参数。平滑项中的视差惩罚函数表达式为

$$\psi(L_p, L_q) = |d_p(L_p) - d_q(L_q)| - |d_q(L_q) - d_q(L_p)|, \quad (9)$$

式中: $d_q(L_p) = a_p q_x + b_p q_y + c_p$ 为像素点 q 在 L_p 标签下的视差值。函数 $\psi(L_p, L_q)$ 用于惩罚邻域像素点 $(p, q) \in N_4$ 分别在标签 L_p, L_q 下视差的不连续性。由于平滑项函数满足 GC 算法中子模块能量优化条件, 即

$$\psi(\alpha^1, \alpha^1) + \psi(\beta^1, \gamma^1) \leq \psi(\alpha^1, \beta^1) + \psi(\alpha^1, \gamma^1), \quad (10)$$

式中: $\alpha^1, \beta^1, \gamma^1$ 为 3D 视差平面参数, 因此可根据 GC 算法最小化能量函数, 获取标签值, (10) 式的详细证明可参考文献[14]。

2.2 局部优化窗口模型

2.2.1 交叉窗口

在 LocalExp 中设定网格大小将图像分割成网格区域 $C_{ij} \in \Omega$, 下标 $(i, j) \in \mathbf{Z}^2$ 代表其网格单元坐标, \mathbf{Z} 为整数集。LocalExp 中采用 3 种不同尺寸 ($5 \times 5, 10 \times 10, 25 \times 25$) 的网格大小构建图像金字塔模型来平衡 PM 算法中的随机搜索和局部传播的互

相影响。为了实现算法加速, 对多匹配单元进行并行优化处理, 规定了一种互不相交的局部区域窗口分组方法, 图像的网格分组数 $K = 6 \times (j \bmod 6) + (i \bmod 6)$, 如图 3(a) 所示为 $K = 36$ 分组后的图像细胞索引图。针对每个网格单元定义三个区域用于视差计算, 图 3(b) 中标号 12 的白色网格单元为中心区域 C_{ij} , 相连的灰色单元为扩张区域 $R_{ij} = C_{ij} \cup_{(m,n) \in N_4(i,j)} C_{mn}$, 其中 m, n 表示细胞单元 C 的横轴坐标。

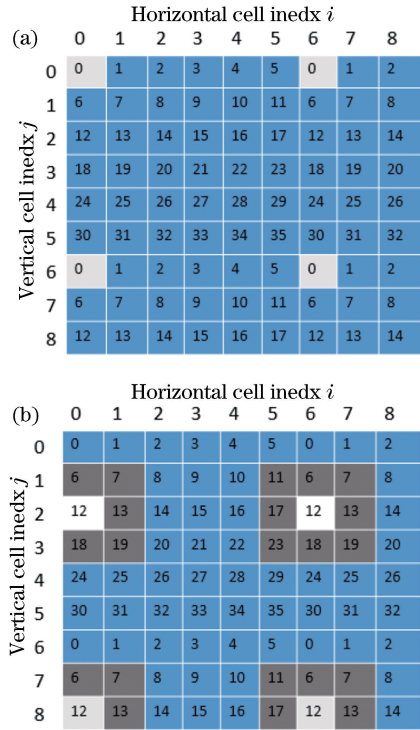


图 3 分组索引。(a) 图像细胞索引分组; (b) 局部扩张区域

Fig. 3 Grouping index. (a) Image cell index groups;

(b) locally extended region

受局部立体匹配中多窗口和自适应窗口思想的启发, 提出一种交叉窗口模型以获取更多局部特性窗口。相对于 LocalExp 的窗口模型, 本文模型中心窗口的滑动距离小于细胞单元尺寸大小, 即保留相连细胞单元的部分区域作为新细胞的中心区域。新中心区域与原细胞区域存在局部信息交叉, 故称交叉窗口。图 4(a) 中给出了同分组下的两种窗口模型(交叉窗口为半个细胞单元移动长度)单元标号结果。交叉窗口结构能够产生更多的窗口去捕捉局部信息, 每一个像素能够被分配更多的标签求取代价, 因此更有利于获取最优标签。图 4(b) 中 $M_{ij} = \bigcup_{p \in R_{ij}} W_p$, 表示为中心区域 C_{ij} 的滤波区域。图 5 给出两种窗口模型在 5 次实验后的平均错误率结果对

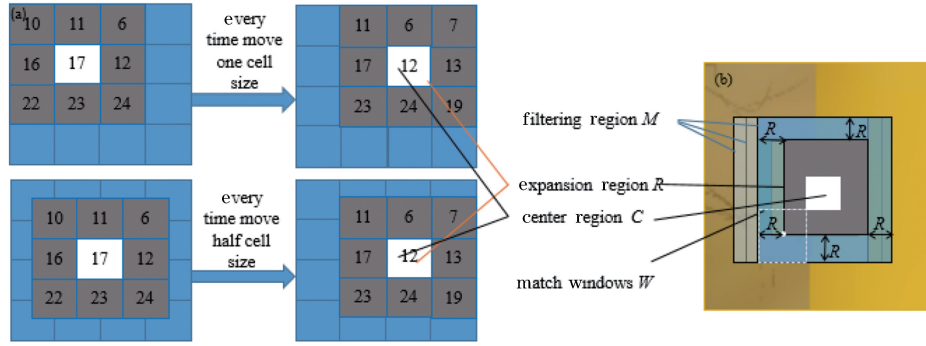


图 4 交叉窗口。(a)不同分组的扩展区域;(b)局部图像区域

Fig. 4 Cross windows. (a) Expansion regions for different groups; (b) local image region

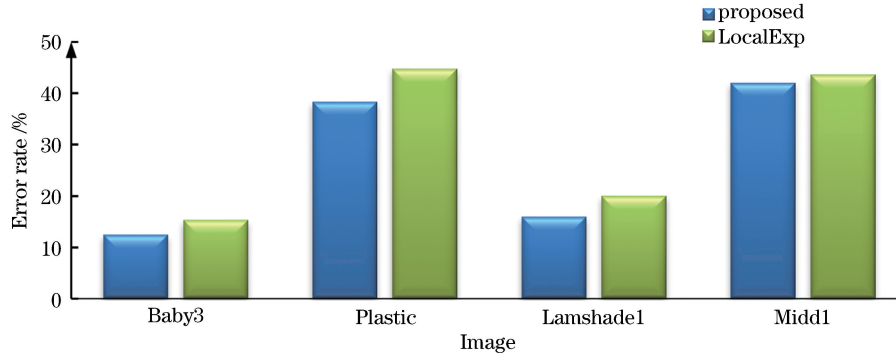


图 5 交叉窗口与原窗口方案在测试图像上的错误匹配率

Fig. 5 Error matching rates of cross window and original method on test images

比,相比于 LocalExp,本文方法的平均匹配准确率明显提高。

2.2.2 窗口标签初始化

基于 PatchMatch 算法的匹配算法流程如下:

1)初始化窗口内像素点标签;2)在每次迭代过程中,在窗口间对标签进行空间传播和随机搜索,通过判断窗口在当前标签下的能量函数增减性来改变窗口内像素的视差分布。算法迭代收敛速度取决于初始化过程的随机 3D 标签,随机数值越可靠,算法收敛速度越快。所以,初始化阶段是重要组成部分。本文替换 LocalExp 中视差标签初始化方式,采用 Ahmed 等^[18]提出的增加约束原则的 PatchMatch 推理方法 IPMS,该方法建立窗口内的像素点所在视差平面法向量、视差范围以及窗口内视差分布关系,给出可靠的初始化标签以加速收敛速度。

IPMS 中给出了两种初始化约束手段:

1)左右视图下视差可见性约束原则。根据射影变换关系

$$\mathbf{n}' = \frac{\mathbf{M}^{-\text{T}} \mathbf{n}}{\|\mathbf{M}^{-\text{T}} \mathbf{n}\|}, \quad (11)$$

建立左视图 I_0 法向量 $\mathbf{n} = (n_x, n_y, n_z)$ 和右视图 I_1

法向量 $\mathbf{n}' = (n'_x, n'_y, n'_z)$ 的转换关系 $\mathbf{M}, \mathbf{M} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,得到任意像素点 p 的 $\mathbf{n} = (n_x, n_y, n_z)$ 中变量的限制关系为 $n_z > 0, n_x < n_z$ 。

2)在窗口内的视差约束。基于中心点 $p(x_p, y_p)$ 的半径为 r 的窗口内 $q(x_q, y_q)$ 的视差 d_q 满足

$$d_q = \frac{n_x}{n_z}(x_p - x_q) + \frac{n_y}{n_z}(y_p - y_q) + d_p, \quad (12)$$

式中: d_p 为像素点 p 的视差值。窗口内任意像素点 q 的视差范围为 $[d_{\min}, d_{\max}]$ 及其与中心像素 p 的坐标差绝对值范围可分别表示为

$$\begin{cases} d_q \in \omega_p = \{d_{\min}, d_{\max}\} \\ |x_q - x_p| \leq r, |y_q - y_p| \leq r \end{cases}, \quad (13)$$

将(13)式代入(12)式可得

$$-n_z(d_p - d_{\min}) \leq n_x(x_p - x_q) - n_y(y_p - y_q) \leq n_z(d_{\max} - d_p), \quad (14)$$

式中: d_{\min} 为最小视差值; d_{\max} 为最大视差值。 d^* 为 d_p 与 d_{\min} 和 d_{\max} 差值中的最小值,满足

$$d^* = \min(d_p - d_{\min}, d_{\max} - d_p), \quad (15)$$

$$-n_z \cdot d^* \leq n_x \cdot r \pm n_y \cdot r \leq n_z \cdot d^*. \quad (16)$$

求解(12) ~ (16)式可得

$$-\frac{n_z}{r+d^*}d^* \leq n_x \leq \frac{n_z}{r}d^*, \text{ if } p \in I_0, \quad (17)$$

$$-\frac{n_z}{r}d^* \leq n_x \leq \frac{n_z}{r+d^*}d^*, \text{ if } p \in I_1, \quad (18)$$

$$n_y \in \frac{s}{r}[-1,1], \text{ where } s = \begin{cases} \min[|n_z d^* + n_x r|, |(n_z + n_x)d^* + n_x r|], & \text{if } p \in I_0 \\ \min[|n_z d^* + n_x r|, |(n_z - n_x)d^* + n_x r|], & \text{if } p \in I_1 \end{cases}, \quad (19)$$

式中: I_0 为参考图; I_1 为待匹配图; s 为约束因子。通过上述约束条件得到的单位法向量 (n_x, n_y, n_z) 取值范围可在初始化3D标签时控制异常3D标签的产生,只保留满足(17) ~ (19)式以及 $n_z > 0, n_x < n_z$ 的标签值。该方案对单位法向量进行多次采样,直到满足约束条件,增加了正确估计像素点3D平面参数的概率。

2.3 基于像素类型的迭代优化

2.3.1 像素分类器

像素分类器由 Meanshift 分类器^[19]、SNIC (simple non-iterative clustering)^[20]分类器、LRC (left right consistency)分类器以及纹理分类器4部分构成。其中 Meanshift 分类器利用颜色概率密度分布聚类颜色相似的像素点,有很好的全局特性,但颜色渐变小的区域分离效果不佳,如图6(b)方框中局部分割标签图所示,SNIC分类器对分割区域具有更好的局部特性,能够弥补 Meanshift 分割的单纯聚类方法缺点,如图6(c)中方框区域所示。纹理

即获取左视图像素点 n_x 的上下边界。最后利用法向量在窗口内部的视差平面关系求解不等式[(16)式],得到左视图 n_y 的取值范围为

分类器可表示为

$$T(p) = \begin{cases} 0, & n \geq 0.75 \times N \times N \\ 1, & \text{else} \end{cases}, \quad (20)$$

式中: n 为像素点个数; N 为窗口半径。其原理是依据权重计算公式[(8)式]计算中心像素点与在半径为 N 的窗口内其他点的相似度大小,后累计相似度值满足大于0.8的像素点个数 n ,将大于设定阈值个数的像素点记作无纹理区域点,反之则为弱纹理点。图6(d)给出输入图像的像素纹理性质分类结果,黑色代表有纹理点,白色代表无纹理点。LRC分类器可表示为

$$\text{LRC}(p) = \begin{cases} 1, & D_0(p) - D_1(p) < 1 \\ 0, & \text{else} \end{cases}, \quad (21)$$

式中: D_0 为参考图视差图; D_1 为待匹配图的视差图。(21)式用于分离视差图中的稳定点(1)和不稳定点(0)。以上几种像素点的分类方法所得结果可作为本文构建 PatchMatch 优化算法所需候选标签集生成和错误点视差优化的先验知识。

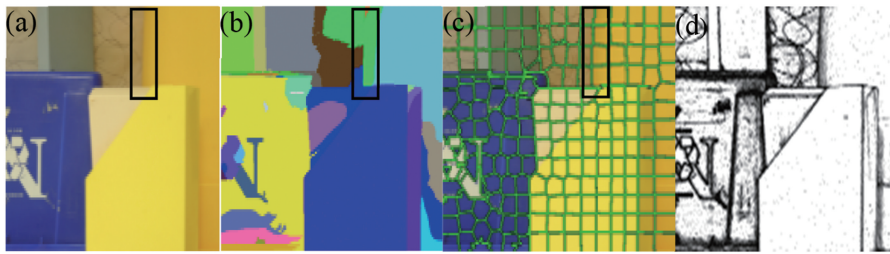


图6 多分类信息图。(a) Plastic;(b) Meanshift;(c) SNIC;(d)纹理结构图

Fig. 6 Multi-category graphs. (a) Plastic; (b) Meanshift; (c) SNIC; (d) texture structure graph

2.3.2 基于像素分割类的标签生成机制

利用2.3.1节的 Meanshift 和 SNIC 像素分类先验知识和 PatchMatch 算法优化原理,改进了 LocalExp 算法的标签生成机制:

- 1) 传播机制。在局部窗口的中心区域的对角线上选取1~2个像素点的3D标签,加入更新候选标签集合。
- 2) 局部最优标签机制。利用局部随机一致性

算法(LO-RANSAC)^[21]拟合局部最优视差平面参数,并将该参数加入到候选标签集合中。

- 3) 基于像素分类的标签机制。在像素分类集合 S 中构造同名像素点集合,并在其中随机选取一定数量的样本点用于 LO-RANSAC 拟合新标签,并加入候选标签集合中。
- 4) 视图传播机制。选取在其他视图下对应区域的有效视差标签加入到候选标签集合中。

5) 随机扰动细化机制。对当前窗口内选取的一个标签量增加归一化的扰动量,细化标签值,直至扰动量足够小,并将扰动标签加入候选标签集合中。

图 7 是采用上述标签生成机制后,基于 PatchMatch 和 GC 优化算法的求解视差的算法流程。针对每个图像细胞单元,首先构建候选标签集合 $S_{et}(\alpha_{ij})$,选取集合中最小化能量函数的标签 L'_p 作为

更新像素点的标签值。最小化能量函数表达式为

$$L^{l+1} = \underset{L'}{\operatorname{argmin}} E[L' | L'_p \in S_{et}(\alpha_{ij})], \quad (22)$$

式中: l 为迭代次数; L' 为最小化能量函数的标签集合; E 为能量函数。

图 8 对比了改进标签生成机制和原方法在一次迭代后的错误率,结果表明,相对于 LocalExp,本文方法的错误率更低。

Input: pixel classification S ; current L ; perturbation size: Δ ; target cell (i, j) ;
 Output: updated L (only labels $p \in R_y$);

1. Repeat1: expansion proposer (propagation);
2. $\alpha_y \leftarrow L_p$ with randomly chosen $p \in C_y$;
3. $L^{l+1} = \underset{L'}{\operatorname{argmin}} E\{L' | L'_p \in \{\alpha_y, L_p\}, p \in R_y\}$;
4. Until propagation times;
5. Repeat2: RANSAC proposer(optional);
6. $\alpha_y \leftarrow \operatorname{RANSAC}(L, C_y)$ and $\alpha_y \leftarrow \operatorname{RANSAC}[L, S(\text{pixel is } C_y)]$;
7. $L^{l+1} = \underset{L'}{\operatorname{argmin}} E\{L' | L'_p \in \{\alpha_y, L_p\}, p \in R_y\}$;
8. Until optional times;
9. Repeat3: randomization proposer (refinement);
10. $\alpha_y \leftarrow L_p$ with randomly chosen $p \in C_y$;
11. $\alpha_y = \alpha_y + \Delta$.
12. $L^{l+1} = \underset{L'}{\operatorname{argmin}} E\{L' | L'_p \in \{\alpha_y, L_p\}, p \in R_y\}$;
13. $\Delta = \Delta / 2$;
14. Until refinement times or Δ is sufficiently small;
15. GC optimization of Eq. (1).

图 7 改进的局部扩张运动算法

Fig. 7 Improved local expansion movement algorithm

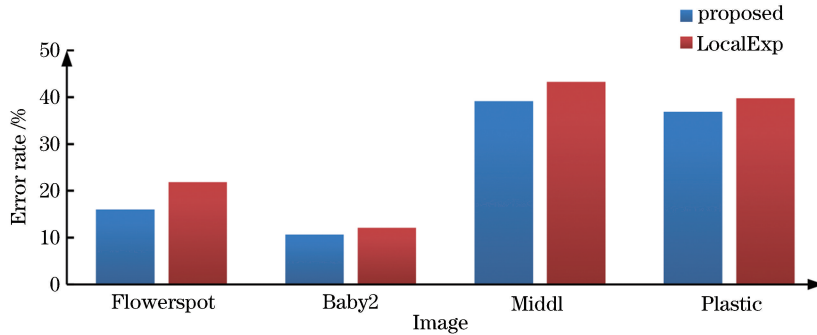


图 8 对比改进标签生成机制和原方法在一次迭代后的错误率

Fig. 8 Comparison of error rates between improved generation mechanism and the original method in one iteration

2.3.3 基于像素分割、纹理及一致性分类的视差优化

从图 8 对比实验结果可知,虽然误匹配率相对于原算法有所降低,但仍较高。为此,采用一种基于像素分类信息的填充方法进一步优化视差。首先通过 LRC 分类器分离不稳定点,其次,依据纹理分类器对有纹理和无纹理区域的不稳定点分别进行处理。

图 9(a)表示位于纹理区域像素点的最近邻域填充方法。以图中标号 5 的不稳定更新为例,分别

向左和向右找到第 1 个位于像素分类器中同分割区域的稳定点。为了避免前景膨胀,将稳定点中使错误点视差更小的标签值赋予该点。图像中无/弱纹理区域一般结构简单且对称,为了有效扩大搜集范围并获得更多的稳定视差点,如图 9(b)所示,采用横竖方向的交叉线法提取一定数目的稳定点作为权重中值滤波像素点集合,用于优化错误点。该方法在当前错误像素点上下左右衍生一定范围以提取稳定点,在各方向选取平均化个数的稳定点以防止选

取过多的不同区域点,通过累计稳定点个数自适应控制衍生搜索的臂长大小。最后采用权重中值选取标签用于更新无/弱纹理区域的不稳定点视差标签值。

图 10 是所提方法和 LocalExp 算法经过 4 次迭代视差优化的视差变化过程图。图 10(a)为 LocalExp 运行结果图,其中方框 1,2,3 代表三个不同区域,图中的黑色点代表视差匹配错误点。在迭

代过程中,方框 1 中部分点可在 LocalExp 算法迭代优化后得到正确的视差值,对于方框 2,3 中的像素点,由于能量函数陷入局部最小值且传统的视差优化手段对于陷入局部最小值的区域无法有效进行滤波处理,因此得到的视差标签图失效。图 10(b)是所提方法的迭代过程图。从图中黑色点数目可知,在初始状态下和迭代过程中,错误区域误匹配率不断降低,最终获得较准确的视差图。

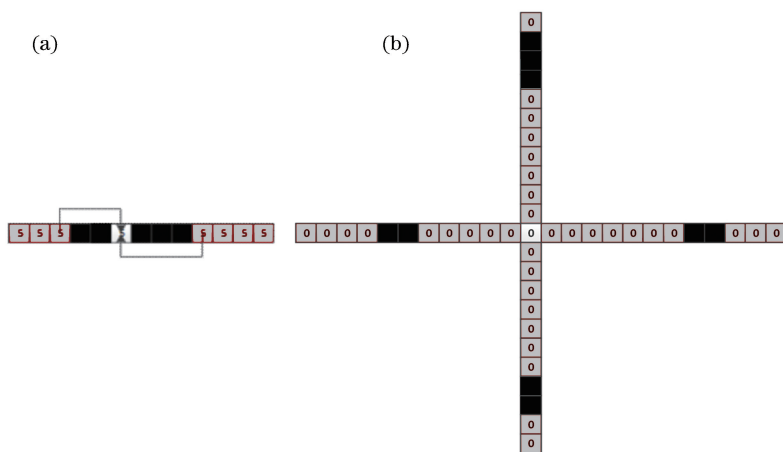


图 9 填充方法。(a)纹理区域;(b)弱纹理分区域。红色为稳定点,黑色为不稳定点,白色为剔除点

Fig. 9 Filling methods. (a) Texture region; (b) texture-less region. Red regions represent stable points, black regions represent unstable points, and white regions represent culling points

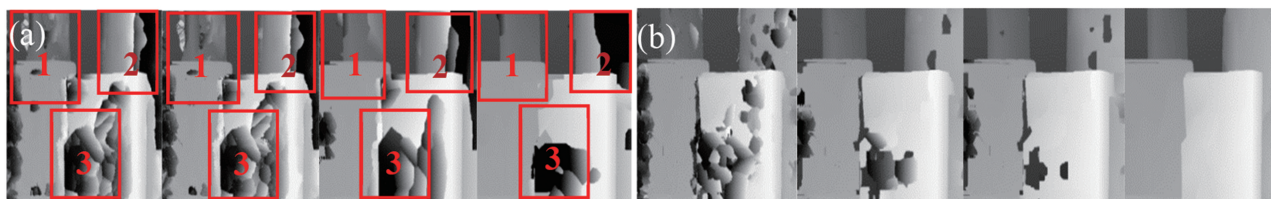


图 10 迭代优化过程。(a) LocalExp;(b)提出的算法

Fig. 10 Iterative optimization process. (a) LocalExp; (b) proposed algorithm

2.4 算法流程

图 11 是本文算法流程图,其中圆 1 代表算法初始化环节,即完成 IPMS 算法中的视差标签初始化和构建像素分类器用于算法迭代优化与视差优化;圆 2 为算法迭代优化环节,即对于每个局部窗口单元采用基于本文改进的标签生成机制的局部扩张运动算法来估计像素点视差;圆 3 代表满足迭代终止后的输出视差图,左下角为算法中数据流动图例说明。

图 12 给出本文算法优化流程说明。1)算法初始化:行 1~5 分别是 2.2~2.3 节中阐述的多层窗口模型尺寸大小、利用 meanshift、SNIC 和纹理分类器 $T(p)$ 构建原始像素点分类集合 S 、迭代优化算法 1 中的 propagation、optional 和 refinement 更新次数

以及其他的实验参数设定。2)算法迭代循环:行 8~9 表示从窗口尺寸集合 $G = \{10 \times 10, 20 \times 20, 35 \times 35\}$ 从小到大选取窗口大小执行 2.2.1 节中阐述的交叉窗口模型和 2.2.2 节的 IPMS 算法初始化(该标签初始化算法只在首次迭代中执行,其余迭代过程中视差标签从上次迭代结果选取)。行 11 代表同组号窗口并行执行 Algorithm 1-局部扩张运动算法迭代更新各自窗口内像素点的标签值,其中 $\Delta = \{\Delta'_d, \Delta'_n\}$ 为视差和法向量的扰动量,用于细化视差标签^[14]。行 14 表示在一次迭代后,采用 2.3.3 节提出的方法进行错误点视差填充;行 16 处表示在一次窗口尺寸迭代完成后,缩小扰动范围 Δ 来改变视差标签细化精度。行 17 为迭代优化算法停止。行 18 执行后处理方法^[14],以进一步提高视差精度。

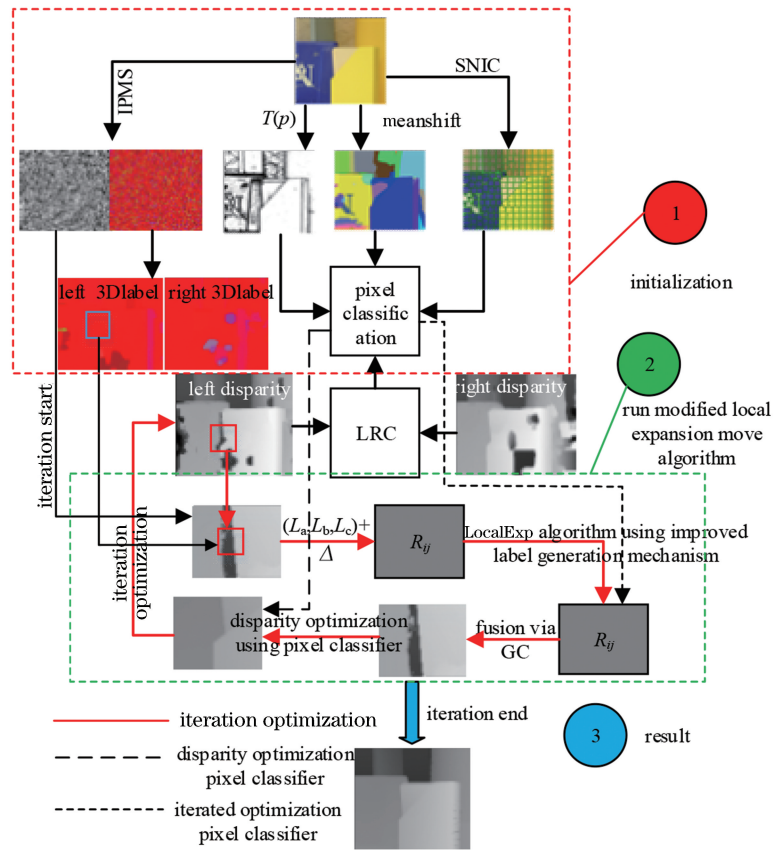


图 11 本文算法流程图

Fig. 11 Flow chart of proposed algorithm

Algorithm 2: algorithm optimization procedure

1. Define three levels of grid structures: $G = \{G_1, G_2, G_3\}$;
2. Initialize pixels classification (S) by SNIC and meanshift and texture function $T(p)$.
3. Initialize perturbation size $|\Delta|$;
4. Initialize the generation proposals cycles.
5. Repeat
 6. for1 each grid structure G do
 7. for2 each disjoint groups $k=0, \dots, 16$ at small grid or $k=0, \dots, 35$ at big grid do
 8. Produce cross window models with grid size at 2.2.1
 9. Initialize the solution L by IPMS at 2.2.2 in first iteration.
 10. for3 each cell (i, j) in the group K do [in parallel];
 11. Do improved local expansion move $(S, L, \text{cell}(i, j), |\Delta|)$ at 2.3.2;
 12. endfor3;
 13. endfor2;
 14. Refinement the solution D by filling error disparities using the disparity refinement at 2.3.3;
 15. endfor1;
 16. $|\Delta| = |\Delta|/2$;
17. Until convergence and stop repeat.
18. Do post processing.

图 12 本文算法优化进程

Fig. 12 Optimization procedure of our algorithm

3 实验结果及分析

实验环境为: Windows 7, 64 位系统, Intel(R) Core(TM) i5-3210M, CPU 主频为 2.50 GHz, 4 核, 内存为 8 GB。为了验证本文算法对无/弱纹理区域的适用性与准确性以及纹理丰富区域的通用性, 实验数据采用 Middlebury 大学的立体视觉数据库中 Midd2003 Datasets、Midd2006 Datasets。该数据集含有真实的视差数据, 可用于衡量算法的误差精度, 以及有效验证算法的正确性。本文实验参数为 $\epsilon = 0.01$, $\epsilon_{\text{dis}} = 0.8$, $N = 5$, $\gamma_n = \gamma_{\text{cn}} = 10$, $\tau_{\text{col}} = 0.2$, $\tau_{\text{grad}_x} = \tau_{\text{grad}_y} = 0.7$, $\lambda = 7$, $\alpha = 0.7$, $R = 20$, $\beta = 0.4$ 。

图 13 中列出基于 PatchMatch 优化的匹配算法^[5,14,22-23]和本文提出的算法在 Midd2006 数据集中存在大片无纹理区域的 4 组测试图像

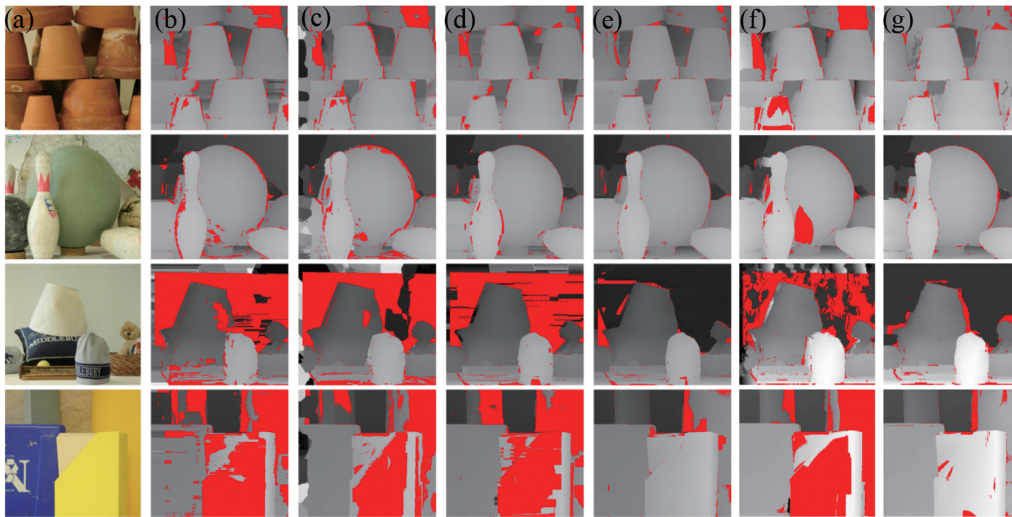


图 13 几种全局 PatchMatch 算法的视差图(匹配错误率大于 1 pixel 的点用红色显示)。(a) Image; (b) PMBP; (c) SPM-BP; (d) GCLSL; (e) PMSC; (f) LocalExp; (g) proposed

Fig. 13 Disparities of some global PatchMatch algorithms (points with error matching rate greater than 1 pixel are shown in red). (a) Image; (b) PMBP; (c) SPM-BP; (d) GCLSL; (e) PMSC; (f) LocalExp; (g) proposed

表 1 中给出几种基于 PatchMatch 优化思想的全局匹配算法在 Midd2006 数据集典型图像中评价结果, 其数值为图像中非遮挡区域的像素点视差错误值大于 1 pixel 的误匹配率结果。从表 1 中可知, 本文算法取得了最低的平均误匹配率 2.59%, 相对 LocalExp 算法的平均误匹配率 (11.66%), 在所有测试的图像中, 只有 Rocks 图像的误匹配率高于 LocalExp。本文算法在纹理较为丰富的测试图像 Cloth1、Cloth4 中获得的最低错误匹配率分别为 0.59% 和 0.89%, 且在有无/弱纹理区域的 Midd1、Plastic 图像中, 取得最低的误匹配率, 在所有存在无纹理区域的图像中都取得较低

Flowerpots、Bowling2、Midd1、Plastic 的视差图可视化结果, 其中红色点代表视差错误大于 1 pixel 的点。从图 13 可知, 在采用没有结合 CNN 代价作为匹配代价的 PatchMatch 优化的全局匹配算法 PMBP^[5]、LocalExp^[14]、SPM-BP^[22] 和 GCLSL^[23] 得到的视差图中, 出现了大量错误点。得益于可靠的初始化标值值和基于像素分类的迭代计算和视差优化, 本文算法在任何一个测试图像上的实验效果都不存在较大区域的错误点。PMSC^[7] 由于利用 CNN 作为匹配代价进行计算, 获得了较好的实验效果, 但其需要通过数据集驱动来训练深度网络结构并计算初始匹配代价值。相比之下, 本文算法未采用 CNN 网络模型却取得与 PMSC 接近的视觉效果, 在弱纹理区域的结果甚至更优。

的误匹配率。

表 2 是本文算法与 LocalExp 在测试图像 Midd2003 中的实验测试结果对比, Rate 代表视差图在非遮挡区域(nonocc)、全部区域(all)以及视差不连续区域(disc)的误匹配率。本文算法相对于 LocalExp 算法在 Teddy、Cones 上取得更低的误匹配率, 由原来的 8.153% 降低到本文的 7.143%。在表 2 中的测试图像的视差图上框出视差细节对比区域, 两种方法中获取视差效果更好的用虚线方框圈出, 反之用实线方框圈出。对比方框内的细节可以发现, 本文方法在图像 Teddy 和 Cones 的 disc 区域上获得更好的边缘精度, 且在非遮挡区域获得更好

的平滑视差效果。但对于第 1 行 Tsukuba 图像匹配效果较差,因为其局部场景结构性过于复杂、图中存在阴影覆盖的影响以及图像视差变化范围小,所以在三个区域的匹配错误率有小幅上升。尽管本

文算法在图像 Tsukuba、Venus 上的匹配精度降低,但整体上仍取得与 LocalExp 接近的平均错误率,LocalExp 和本文算法的平均错误率分别为 5.97% 和 6.01%。

表 1 几种基于 PatchMatch 优化的匹配算法在 Midd2006 数据集中非遮挡区域的匹配结果
(阈值为 1 pixel,最好的实验结果用粗体显示)

Table 1 Matching results of some PatchMatch based stereo algorithms in Midd2006 datasets with nonocc regions
(threshold is 1 pixel, and the best results are shown in bold)

Image	PMBP	SPMBP	GCLSL	PMSC	LocalExp	Proposed
Aloe	4.51	6.75	3.21	3.06	3.92	3.25
Baby1	4.10	3.27	2.21	1.98	2.74	1.34
Baby2	4.77	3.97	2.08	1.05	5.48	1.41
Baby3	4.77	3.92	3.07	3.12	6.56	2.73
Bowling1	14.10	12.10	4.14	2.06	5.37	2.46
Bowling2	4.64	5.27	2.19	1.45	6.44	1.99
Cloth1	1.68	1.17	0.71	0.60	0.78	0.59
Cloth4	3.10	2.20	1.75	1.87	0.99	0.98
Flowerpots	9.28	8.80	4.60	2.49	10.89	4.01
Lampshade1	13.50	8.67	12.60	1.50	5.96	2.14
Lampshade2	16.50	17.20	10.00	0.99	21.70	1.08
Midd1	37.40	37.40	34.90	12.80	30.80	6.07
Midd2	38.40	33.20	32.90	4.37	25.60	4.71
Monopoly	42.40	32.70	21.10	3.46	28.04	4.97
Plastic	44.80	35.20	43.90	4.40	40.10	3.34
Rocks1	4.15	2.60	2.19	1.80	1.60	2.34
Wood1	1.52	4.19	0.48	0.73	1.30	0.71
Average	14.68	12.85	10.70	2.80	11.66	2.59







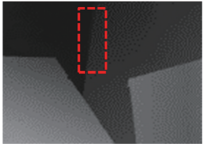
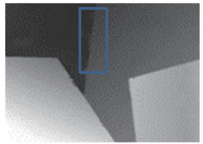
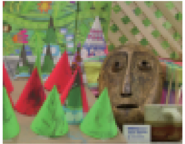
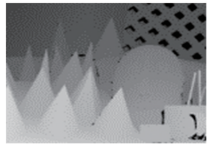

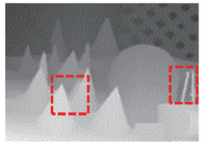




4 结 论

提出一种基于像素类别优化的全局匹配算法。与多种相近算法比较,本文算法在 Midd2006 图像数据集上取得了最低的平均错误匹配率,相对于 LocalExp 算法降低了 9.07%,而且在存在大量无纹理区域图像 Midd1、Plastic 上取得了最低的误匹配率,实验验证以下结论:本文算法能够有效改善图像

中弱纹理和无纹理区域的误匹配问题。但本文算法还存在一些不足:1)利用不同分类方法构建图像中像素点的类别信息,其作为一种人工设计的方法,对不同分割方法的准确性依赖较大;2)由于迭代过程中加入视差优化算法,运行比 LocalExp 耗时。在接下来的工作中,将研究以深度学习和传统方法相结合的方式获取可靠的全局像素信息,以提升匹配算法精确度并进一步缩短算法的运行时间。

表 2 本文算法与 LocalExp 在 Midd2003 数据集的错误率对比(阈值为 0.5 pixel)

Table 2 Comparison of error rates between proposed method and LocalExp on Midd2003 datasets (threshold is 0.5 pixel)

Image	GroundTruth	LocalExp	Rate	Proposed	Rate
			4.08 4.73 9.71		5.13 6.69 10.6
			0.35 0.56 3.30		0.74 0.87 5.94
			3.46 8.65 9.72		3.01 8.26 8.98
			5.16 7.73 14.2		4.35 7.06 11.2

参 考 文 献

- [1] Ma R H, Zhu F, Wu Q X, *et al.* Dense stereo matching 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] 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 1, 2011, Dundee. Durham: BMVA Press, 2011: 14.
- [3] Barnes C, Shechtman E, Finkelstein A, *et al.* PatchMatch: a randomized correspondence algorithm for structural image editing[J]. *ACM Transactions on Graphics*, 2009, 28(3): 24.
- [4] Luus R, Jaakola T H I. Optimization by direct search and systematic reduction of the size of search region [J]. *AICHe Journal*, 1973, 19(4): 760-766.
- [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] 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 (CVPR)*, June 23-28, 2013, Portland, OR, USA. New York: IEEE, 2013: 1854-1861.
- [7] Heise P, Klose S, Jensen B, *et al.* PM-huber: PatchMatch with huber regularization for stereo matching [C] // *2013 IEEE International Conference on Computer Vision (ICCV)*, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE, 2013: 2360-2367.
- [8] 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.
- [9] Žbontar J, LeCun Y. Stereo matching by training a convolutional neural network to compare image patches [J]. *Journal of Machine Learning Research*, 2016, 17: 1-32.
- [10] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(9): 1124-1137.
- [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] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. *International Journal of Computer Vision*, 2004, 59(2): 167-181.

- [13] Yang Q X. A non-local cost aggregation method for stereo matching [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 1402-1409.
- [14] Tani ai 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.
- [15] Chipot M, March R, Rosati M, *et al.* Analysis of a nonconvex problem related to signal selective smoothing[J]. Mathematical Models and Methods in Applied Sciences, 1997, 7(3): 313-328.
- [16] 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.
- [17] Olsson C, Ulén J, Boykov Y. In defense of 3D-label stereo [C] // 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2013, Portland, OR, USA. New York: IEEE, 2013: 1730-1737.
- [18] Ahmed S, Hansard M, Cavallaro A. Constrained optimization for plane-based stereo [J]. IEEE Transactions on Image Processing, 2018, 27(8): 3870-3882.
- [19] 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.
- [20] Achanta R, Süssstrunk S. Superpixels and polygons using simple non-iterative clustering [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4895-4904.
- [21] Chum O, Matas J, Kittler J. Locally optimized RANSAC[M] // Michaelis B, Krell G. Lecture notes in computer science. Berlin, Heidelberg: Springer, 2003, 2781: 236-243.
- [22] Li Y, Min D B, Brown M S, *et al.* SPM-BP: sped-up PatchMatch belief propagation for continuous MRFs [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 4006-4014.
- [23] Tani ai T, Matsushita Y, Naemura T. Graph cut based continuous stereo matching using locally shared labels [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1613-1620.