

基于区域间协同优化的立体匹配算法

郑志刚¹ 汪增福¹

摘要 提出了一种基于分割区域间协同优化的立体匹配算法. 该算法以图像区域为匹配基元, 利用区域的彩色特征以及相邻区域间应满足的平滑和遮挡关系定义了区域的匹配能量函数, 并引入区域之间的合作竞争机制, 通过协同优化使所定义的匹配能量极小化, 从而得到比较理想的视差结果. 算法首先对参考图像进行分割, 利用相关法得到各分割区域的初始匹配; 然后用平面模型对各区域的视差进行拟合, 得到各区域的视差平面参数; 最后, 基于协同优化的思想, 采用局部优化的方法对各区域的视差平面参数进行迭代优化, 直至得到比较合理的视差图为止. 采用 Middlebury test set 进行的实验结果表明, 该方法在性能上可以和目前最好的立体匹配算法相媲美, 得到的视差结果接近于真实视差.

关键词 立体匹配, 图像分割, 协同优化, 平面拟合, 均值偏移
中图分类号 TP391

A Region Based Stereo Matching Algorithm Using Cooperative Optimization

ZHENG Zhi-Gang¹ WANG Zeng-Fu¹

Abstract This paper presents a stereo matching algorithm based on inter-regional cooperative optimization. This algorithm uses regions as matching primitives and defines the corresponding region cost functions for matching by utilizing the color statistics of regions and the constraints on smoothness and occlusion between adjacent regions. In order to obtain a more reasonable disparity map, a cooperative optimization procedure is employed to minimize the matching costs of all regions by introducing the cooperative and competitive mechanism between regions. Firstly, a color based segmentation method is used to segment the reference image into regions with homogeneous color. Secondly, a local window-based matching method is used to determine the initial disparity estimates of each image point. And then, a plane fitting technique is applied to obtain the parameters of disparity plane corresponding to each image region. Finally, under a framework of inter-regional cooperative optimization, the disparity plane parameters of all regions are iteratively optimized by a local optimization method until a reasonable disparity map is obtained. The experimental results based on Middlebury test set indicate that the performance of our method is competitive with the best stereo matching algorithms and the disparity maps recovered are close to the ground truth data.

Key words Stereo correspondence, image segmentation, cooperative optimization, plane fitting, mean shift

立体匹配是计算机视觉中的一个经典问题. 多年来, 研究者们提出了大量的算法试图解决这个问题. 但由于问题本身的病态性, 目前还没有比较完美的解决方案. Scharstein 等对到目前为止所提出的各种主要方法进行了较全面的综述^[1]. 近年来, 由于图割 (Graph-cut, GC)^[2-3] 和信念传播 (Belief-propagation, BP)^[4-5] 等全局优化方法的成功应用, 使立体匹配在精度和准确性方面得到较大的提高. 对立体视觉中的两大难题: 遮挡和少纹理区域问题, 研究者们也进行了深入的研究, 并提出了一些有针对性的解决方案. 其中, 基于区域分割的方法^[6-10] 对于解决缺少纹理区域问题效果比较好. 这些算法首先利用图像的颜色信息对图像进行分割, 再利用

虽不是很准确但相对比较快速的匹配方法 (例如相关法、动态规划法等) 得到初始的视差估计, 然后采用平面模型对每个分割区域的视差进行拟合, 完成对缺少纹理区域的推理估计. 由于 GC 和 BP 等优化算法都是基于“贴标签 (Labelling)”的优化算法, 为了限制最终优化时的标签数, 这些算法在优化前通常需要在视差平面参数空间中对区域进行聚类分层以减少最终优化时的标签数, 但这样做会使得算法的通用性降低, 而且最终优化时被用来作为标签的视差平面参数是不变化的, 如果最终优化前视差平面参数估计不准确, 则其在优化过程中不能被修正. 此外, 文献 [11] 提出了一种对称的立体匹配算法. 该算法的最大特点是在区域的层次上平等地处理左右图像, 从而比较容易确定和处理遮挡区域, 较好地解决了场景中的遮挡问题.

为了使算法不受最终标签数的限制, 并能最终得到每个区域的最优平面参数, 本文提出了一种基于区域间协同优化的立体匹配算法. 协同优化算法^[12] 利用优化子目标之间合作竞争关系达到优化

收稿日期 2007-11-04 收修改稿日期 2008-10-23
Received November 4, 2007; in revised form October 23, 2008
国家自然科学基金 (60455001, 60875026) 资助
Supported by National Natural Science Foundation of China (60455001, 60875026)

1. 中国科学技术大学自动化系 合肥 230026
1. Department of Automation, University of Science and Technology of China, Hefei 230026
DOI: 10.3724/SP.J.1004.2009.00469

总匹配能量的目的. 如果把每个区域的优化视为优化子目标, 并利用区域之间的合作竞争关系对每个区域的匹配能量进行协同优化, 则可以在最小化单个区域能量的同时使总匹配能量也逐步趋向极小值. 文献 [12] 的协同优化算法对子目标的优化依赖于局部优化方法, 而局部优化很容易陷入局部极小值, 因此, 为了获得全局最优解, 该算法本质上需要有一个较好的优化起点. 与此不同, 本文所提出的协同优化算法虽然在优化前也需要利用初始匹配的结果对每个区域的视差平面参数进行估计, 且好的初始匹配会对最终获得正确的视差信息起到积极的正面效果. 但是, 本文算法的最终结果不仅和每个区域的初始视差有关, 还和所有区域总的视差分布有关. 特别地, 如果由初始匹配得到的视差平面参数的估计不准确, 可以在后续的优化过程中利用其邻域的相关信息修正.

本文算法是针对使用标准配置的两台摄像机所拍摄的立体图对而设计的, 由三个主要步骤构成: 1) 通过相关法得到初始视差图; 2) 采用一种鲁棒的平面拟合方法对每个区域的视差进行平面拟合得到视差平面参数; 3) 根据所定义的区域匹配能量, 综合考虑各区域平面参数的变化对总匹配能量的影响, 采用协同优化算法使每个分割区域的匹配能量达到极小值, 从而得到比较理想的视差结果. 算法的流程图如图 1 所示. 由于每一步骤上发生的错误在很大程度上可以在后续步骤中得以修正, 因此本文算法对实拍图像也表现出很好的鲁棒性.

本文的结构如下: 第 1 节介绍得到初始视差图的方法; 第 2 节介绍协同优化的基本原理; 第 3 节介绍本文提出的基于区域间的协同优化算法; 最后给出实验结果并给出结论和展望.

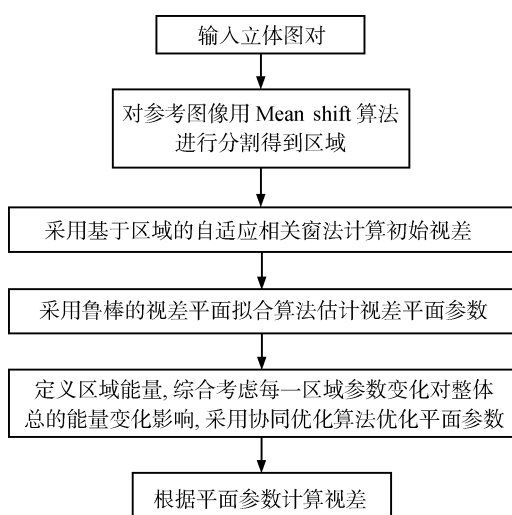


图 1 算法流程图

Fig. 1 The flowchart of the algorithm

1 初始视差图的获得

1.1 区域分割和基于区域的自适应相关匹配

本文算法使用 Mean shift 算法^[13] 通过对输入彩色图像的色彩域和空间域特征进行聚类达到对图像进行分割的目的. 由于基于区域的匹配算法通常假设不同颜色的区域边缘与视差跳跃边缘相一致, 为了使这一假设尽可能得到满足, 实际中采用“过分割”算法以得到尽可能多的分割区域.

基于相关窗的匹配方法是一种效率非常高的经典立体匹配算法. 但由于它对纹理较少的区域和景物边缘区域效果不理想, 因此研究者提出了可变窗算法^[14-16] 和颜色加权算法^[17] 等很多改进算法. 其中, 颜色加权算法取得的效果比较好, 但它的速度较慢, 无法体现相关法高效的优点. 文献 [18] 提出了一种基于分割区域的自适应相关方法. 该算法首先对参考图像进行分割, 并基于分割结果在相关窗内通过对同一区域像素加权达到可变窗的效果, 较好地去除了相关法的边缘效应; 与此同时, 通过对每个区域的能量进行迭代计算, 保留了相关法高效的特点. 在进行视差选择时, 本文采用最简单的赢者通吃 (Winners-take-all, WTA) 算法, 由此会带来一定的误差, 但这些误差可由后续步骤消除. 图 2 给出了对 Tsukuba 立体图像对进行处理的结果. 其中, 图 2(a) 和 2(b) 分别是对左图像的分割图像和经过初匹配后得到的视差图.

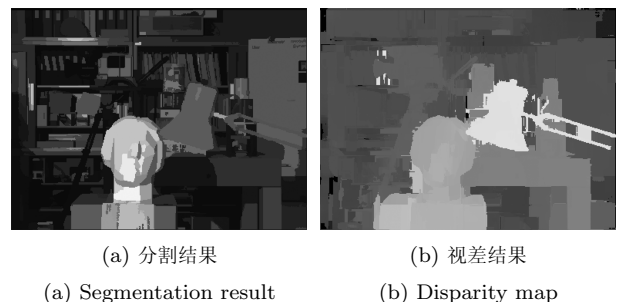


图 2 自适应相关窗立体匹配算法对 Tsukuba 立体图像对得到的分割结果和视差图

Fig. 2 The segmentation result and disparity map of the Tsukuba stereo pair by using the adaptive correlation window stereo matching algorithm

1.2 基于投票的鲁棒视差平面拟合算法

分割区域的视差平面模型用函数

$$d(x, y) = ax + by + c \quad (1)$$

表示. 其中, x, y 为图像坐标, a, b, c 为平面参数. 拟合时, 采用匹配的可靠性为权值, 最终得到的平面法向方向 (n_x, n_y, n_d) 是矩阵 A 的最小特征根所对应的特征向量.

$$A = \begin{bmatrix} \sum_{i=0}^N w_i x_i^2 & \sum_{i=0}^N w_i x_i y_i & \sum_{i=0}^N w_i x_i d_i \\ \sum_{i=0}^N w_i x_i y_i & \sum_{i=0}^N w_i y_i^2 & \sum_{i=0}^N w_i y_i d_i \\ \sum_{i=0}^N w_i x_i d_i & \sum_{i=0}^N w_i y_i d_i & \sum_{i=0}^N w_i d_i^2 \end{bmatrix} \quad (2)$$

其中, w_i 是反映匹配可靠性的权值, x_i 、 y_i 是匹配点的图像坐标, d_i 是匹配点的视差, N 是该区域内的匹配点数。

由于在计算初始视差图时不可避免地会带来一定的误差, 而常用的平面拟合算法恰恰很容易受误差点的影响, 所以在进行平面拟合之前一般需要排除数据中的干扰点 (离群数据)。随机抽样一致性 (Random sample consensus, RANSAC) 算法可用于排除干扰点, 但由于 RANSAC 算法的结果依赖于初始点的选择, 而初始点的选择具有一定的随机性, 这使得该算法得到的结果并不稳定。同时, RANSAC 算法还需要设定待拟合点到估计平面的距离阈值, 阈值设定的合理与否也对结果有较大的影响。为了克服 RANSAC 算法的缺陷, 本文采用了一种更为鲁棒的基于投票的改进方法。

考察式 (1), 通过对同一行的一对点计算 $(\delta d)/(\delta x)$, 可以得到平面参数 a 的一个估计, 对区域中同一行中的所有点进行相似的计算可以得到参数 a 的若干估计值; 将这些估计值在一维的 a 参数空间进行投票, 并对投票结果进行高斯平滑后从中选出得票最多的值作为参数 a 的最终估计。同样地, 对同一列的点通过计算 $(\delta d)/(\delta y)$, 可以得到参数 b 的估计。在计算出参数 a 、 b 后, 由式 (1) 可算出 c , 采用类似的投票方法可确定参数 c 。

采用上述算法, 可以有效地去除离群数据的影响, 使结果具有鲁棒性。为了对上述投票方法进行评估, 本文给出了在相同参数设置条件下分别采用 RANSAC 方法和投票方法进行拟合得到的 Tsukuba 图像对的视差与标准视差的误差比较结果。从图 3 所示的结果可以看出, 本文提出的投票方法不仅拟合的错误率低, 而且结果非常稳定; 而 RANSAC 方法虽然拟合的平均错误率也较低, 但性能不如投票法稳定。

图 4 是利用投票法对视差平面拟合后得到的视差图。实验结果表明, 经过以上处理, 初始视差图中的噪声得到了有效的抑制。

2 协同优化算法的基本原理

所谓协同优化^[12] 是将一个复杂的目标首先分解成若干个简单的子目标, 然后在考虑子目标之间

所具有的关联关系的基础上, 对这些子目标进行综合优化。在优化每一子目标函数的时候, 综合考虑与之有关联关系的其他子目标函数的优化结果, 使相互关联的子目标函数之间的优化结果保持一致。协调优化的目标是在保证每个子目标取得最优解的前提下, 使包含在不同子目标函数中的同一变量的取值保持一致。一般而言, 这样的解就是所求问题的最优解。特别地, 如果在相互关联的子目标当中的某些子目标具有性质良好的局部最优解, 则协同优化的结果大多可以最终得到全局最优解。

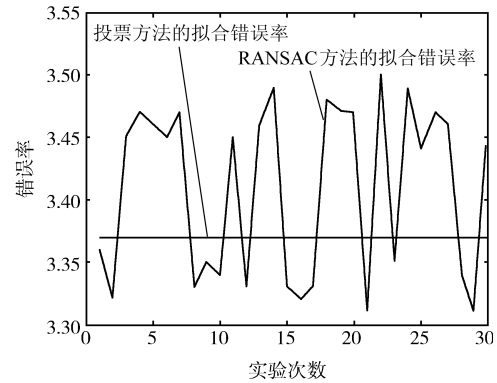


图 3 RANSAC 方法和投票方法平面拟合结果的比较
Fig. 3 The comparison of the plane fitting results based on the RANSAC algorithm and the voting algorithm



图 4 采用基于投票的平面拟合法得到的视差图
Fig. 4 The disparities obtained by the plane fitting algorithm based on voting

具体言之, 设目标函数是一个 n 元函数 $E(x_1, x_2, \dots, x_n)$, 简写成 $E(x)$, 则协同优化算法先将其分解成 m 个简单的子目标函数

$$E(x) = E_1(x^1) + E_2(x^2) + \dots + E_m(x^m) \quad (3)$$

其中, x^i , $i = 1, 2, \dots, m$ 表示第 i 个子目标的变量集合, 它是 x 的一个子集。显然, 如果单独优化每一子目标函数, 则一般情况下对各个子目标单独优化的结果之间很难取得一致。例如, 考虑变量 x_i , 一般而言, 它被包含在若干个子目标中。由于优化时没有考虑到子目标之间的关联, 则对各个相关的子目标单独优化得到的变量 x_i 的最优解之间可能并不

一致. 为了使各子目标函数之间的优化结果能够保持一致, 协同优化算法在优化每一子目标函数 $E_i(x)$ 的时候, 需考虑与之相关联的其他子目标函数的优化结果 $\min_{x^j} E_j(x)$, $j \neq i$ 对其的影响.

具体做法是进行迭代计算, 通过数值加权的方式, 利用其他子目标函数的优化结果对每一个子目标函数进行如下修正

$$(1 - \lambda_i)E_i(x) + \lambda_i \sum_{j \neq i} w_{ij} E_j(x) \quad (4)$$

这里, λ_i , w_{ij} 为加权系数, 满足 $0 \leq \lambda_i \leq 1$, $w_{ij} \leq 1$. 其中, λ_i 表示合作强度, 而 w_{ij} 则用于刻画传播的力度.

假设程序已进行 $k-1$ 次迭代, 则协同优化算法在第 k 次迭代时所采用的迭代公式为

$$E_i^{(k)}(x) = \min \left((1 - \lambda_i^{(k)})E_i^{(k-1)}(x) + \lambda_i^{(k)} \sum_{j \neq i} w_{ij} E_j^{(k-1)}(x) \right), \quad i = 1, 2, \dots, m \quad (5)$$

不断进行上述迭代过程, 直至算法收敛或执行完规定的迭代次数. 由于对每一个区域的优化结果会在下一次迭代中向周边传播, 经过若干次迭代后每一个优化变量在相关子目标函数中的最终取值将会取得一致.

在我们的算法中, i 表示左图像中根据 Mean shift 算法得到的第 i 个区域, 而 j 表示左图像中与 i 相邻的区域. 每一个图像区域被作为一个需要优化的子目标, 协同优化的目的是得到作为上述 m 个子目标变量集合的区域视差参数. 这样, 若区域分割的结果共得到 m 个区域, 则需要协同优化的子目标的个数为 m . 如图 5 所示, 每一个区域都是一个优化子目标. 例如, R_6 是一个优化子目标, 在优化该子目标时, $i = 6$, 而 $j = 2, 3, 5, 7, 9$. 同理, R_7 也是一个优化子目标, 在优化该子目标时, $i = 7$, 而 $j = 3, 4, 6, 8, 9, 10$.

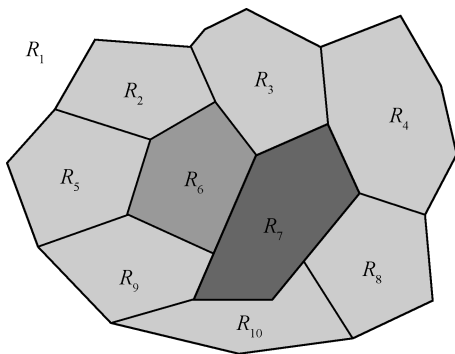


图 5 区域示意图

Fig. 5 The sketch map for optimization of sub-targets

3 基于区域间协同优化的立体匹配

在进行协同优化时, 为了评价当前图像区域 i 的视差平面参数是否合理, 需要定义一个能量函数 E_i , 该能量函数由三部分组成

$$E_i = E_{\text{data}} + E_{\text{occlude}} + E_{\text{smooth}} \quad (6)$$

式 (6) 中的第一项 E_{data} 是数据能量, 定义为

$$E_{\text{data}} = \max_{p \in V_l, q \in V_r} (|r(p) - r(q)|, |g(p) - g(q)|, |b(p) - b(q)|) \quad (7)$$

其中, V_l 和 V_r 分别表示当前区域在左右图像上的可见像素集, $p \in V_l$, $q \in V_r$ 为左右图像上相匹配的两个对应像素, r , g , b 表示相应像素的 RGB 颜色值. 各像素的 RGB 颜色值由以下方法得到: 根据所估计的当前区域的视差平面参数计算与参考图像 (本文选为左图像, 下同) 上的像素 p 对应的被匹配图像 (本文选为右图像, 下同) 上的像素 q , 并获取其 RGB 颜色值. 由于计算得到的视差是浮点型的, 所以像素 q 一般并不正好处于整数像素位置上, 故其 RGB 颜色值一般不能直接从图像中获得. 在本文的算法中, 其值由被匹配图像上 q 的四个近邻像素的 RGB 颜色值经过插值得到. 计算数据能量时考虑了可视性准则^[9], 即要求用来计算数据能量的像素在左右图像对中对都是可见的.

E_{occlude} 是遮挡能量, 这一项对于算法是必要的, 否则将产生一些无意义的结果. 例如, 将整个区域的视差都设为 0 将可能导致能量最小. 遮挡分为左遮挡和右遮挡两种情况. 如图 6 (见下页) 所示, 考虑参考图像上的两个相邻区域 A 和 B , A 在 B 的右边. 当区域 A 在相邻边界处的视差大于其左邻域 B 的视差时, 那么根据当前的视差计算结果将 A 和 B 映射到右图像时, B' 的一部分 (图中的区域 D) 将被 A' 所遮挡 (左遮挡), 而当区域 A 在相邻边界处的视差小于其左邻域 B 的视差时, 映射到右图像后的两个区域 A' 和 B' 之间将会出现空隙 (图中的区域 C), 相当于右图像上区域 A' 有一部分在左图像上被遮挡了 (右遮挡). 在我们的算法中, 如下设置遮挡能量: 对参考图像而言, 当一个区域被其他区域遮挡时, 该区域要加上左遮挡能量; 同理, 对被匹配图像而言, 当一个区域被其他区域遮挡时, 该区域要加上右遮挡能量. 实际中, 被遮挡区域的检测可如下进行: 首先根据当前的视差计算结果将左图像上相关的各区域映射到右图像, 然后在右图像上检查映射情况. 如果某个像素被重复映射, 则将该像素记为左遮挡; 如果某个像素未被映射, 则将该像素记为右遮挡.

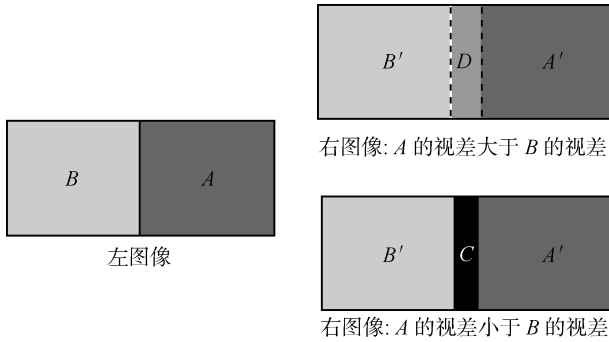


图 6 遮挡示意图

Fig. 6 The sketch map of occlusion

综上所述,可以得到遮挡能量的形式为

$$E_{\text{occlude}} = (|Occ_L| + |Occ_R|)\lambda_{\text{occ}} \quad (8)$$

其中, Occ_L 、 Occ_R 分别表示左遮挡和右遮挡像素的个数, 而 λ_{occ} 表示所设置的遮挡惩罚常量.

E_{smooth} 是平滑能量, 定义为

$$E_{\text{smooth}} = \sum_{p \in B_c, q \in N} \begin{cases} \lambda_{\text{disp}}, & \text{若 } |d(p) - d(q)| \geq 1 \\ 0, & \text{其他} \end{cases} \quad (9)$$

这里, B_c 表示参考图像上当前区域的边界点集, N 表示和 B_c 近邻的其他区域上的边界点集, $p \in B_c$ 、 $q \in N$ 为四连通意义上的两个近邻像素点, $d(p)$ 、 $d(q)$ 为像素 p 、 q 的视差, 而 λ_{disp} 为所设置的平滑惩罚常量, 其中的判别式 $|d(p) - d(q)| \geq 1$ 用于判断当前区域上一个边界点是否为视差不连续的点. 作为附加条件, 要求 p 不是一个遮挡边界点. 显然, 式 (8) 给出了对当前区域上具有不连续视差且不属于遮挡边界的边界点所施加的惩罚能量之和.

由式 (6) 定义的准则函数可知, 区域的匹配能量越小, 说明所得到的视差平面参数越合理.

计算完每个区域的能量后, 选用局部优化方法 (例如 Powell 局部优化方法), 按照式 (5) 对所有区域的视差平面参数进行迭代优化. 具体的优化流程如图 7 所示.

当一个区域的平面参数变化时, 它不仅会影响本身区域的能量, 同时也会对邻域的能量产生影响. 根据协同优化原理, 对每个区域优化时考虑其邻域的能量. 这里的邻域包括所有可能带来能量变化的区域, 即在视差搜索范围内的所有邻域.

初始匹配时生成的视差并不要求都是正确的, 实际上要求初始匹配的视差完全正确也是不合理和不可能. 这里只要求初始匹配的视差是基本正确的. 由于视差计算和遮挡计算之间的关系是互相依存的, 只有当视差正确时计算得到的遮挡才是正确的, 因此, 由具有错误视差参数的区域计算得到的遮

挡也是不正确的, 但只要其邻域的视差平面参数基本正确, 利用邻域之间的配合, 经过数次迭代之后, 错误区域的视差参数就可以被修正. 作为一个实例, 图 8 (见下页) 显示了将协同优化算法应用于标准图像对 Tsukuba 的情况. 如图 8 所示, 在初始匹配时, 一些区域的视差明显是不正确的 (参见图 4), 但随着协同优化算法的迭代进行, 通过邻域间的相互配合, 算法的总匹配能量 e 迅速下降, 相应区域的视差平面参数也基本上都被修正为正确的参数. 图 9 (见下页) 给出了迭代优化过程中总匹配能量下降的情况. 可以看到, 总匹配能量的下降速度较快, 经过 4 次迭代后已基本达到理想的水平.

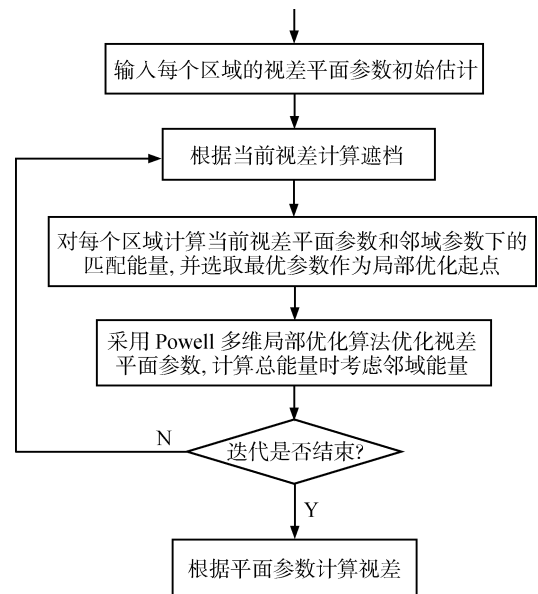


图 7 协同优化算法的流程图

Fig. 7 The flowchart of the cooperative optimization algorithm

另外, 基于本文算法所得到的视差结果与图像的初始分割有一定的关联. 总体上来说, 分割得到的区域边缘和实际的视差边缘越一致, 所得到的视差结果就越理想. 为了调查不同分割结果对最终视差计算的影响, 以 Tsukuba 立体图对为例, 进行了不同分割情况下的视差计算, 结果如图 10 (见下页) 所示. 其中, 采用 Mean shift 算法^[13] 完成对参考图像的分割. 分割时的色彩距离和空间距离参数分别为 σ_r 、 σ_s . 上述两个参数的取值越小, 表示分割越细. 实验结果表明, 虽然最终的视差结果与图像的初始分割之间存在一定的关联, 算法的匹配性能在一定程度上受分割结果的影响. 但算法总体的匹配效果是令人满意的. 从实验结果可以看出, 适度的过分割对获得正确的视差信息是有利的. 分割过粗会在一定程度上影响立体匹配的精度, 而分割过细则有可能导致局部的误匹配. 对 Middlebury 网站提

供的立体图对集合所进行的实验结果表明, $\sigma_r = 7$, $\sigma_s = 5$ 是较理想的选择.

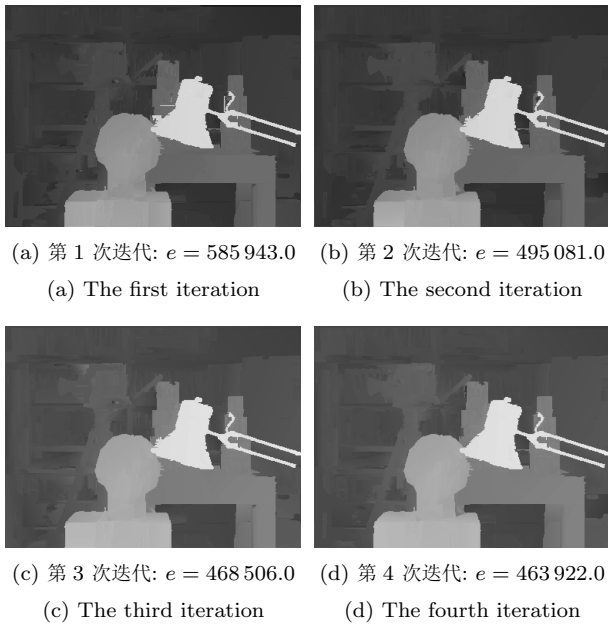


图 8 将协同优化算法应用于标准图像对 Tsukuba 的情况 (经过协同优化后修正了原来错误的匹配区域, 其中, (a)~(d) 是第 1 次到第 4 次迭代得到的视差图)

Fig. 8 The cooperative optimization results: disparity map of Tsukuba stereo image pair

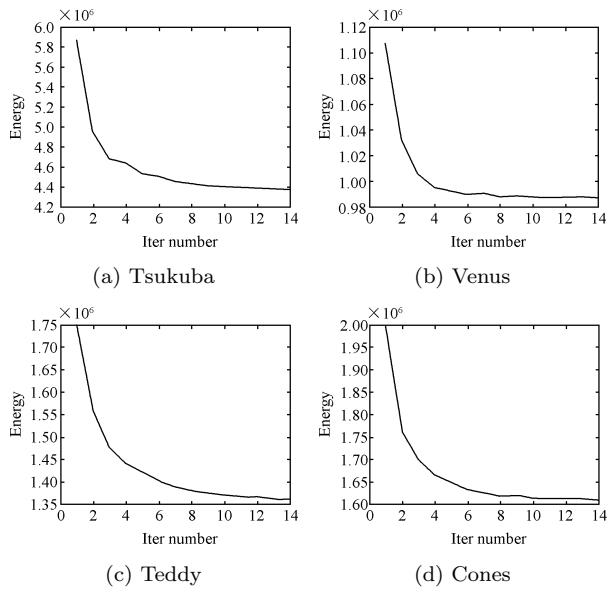


图 9 将协同优化算法应用于 Middlebury 标准图像对能量下降的情况

Fig. 9 The curve of total energy of the stand image pairs on Middlebury website obtain by our algorithm

4 实验结果

为了验证算法的有效性, 用 VC 8.0 实现了本文

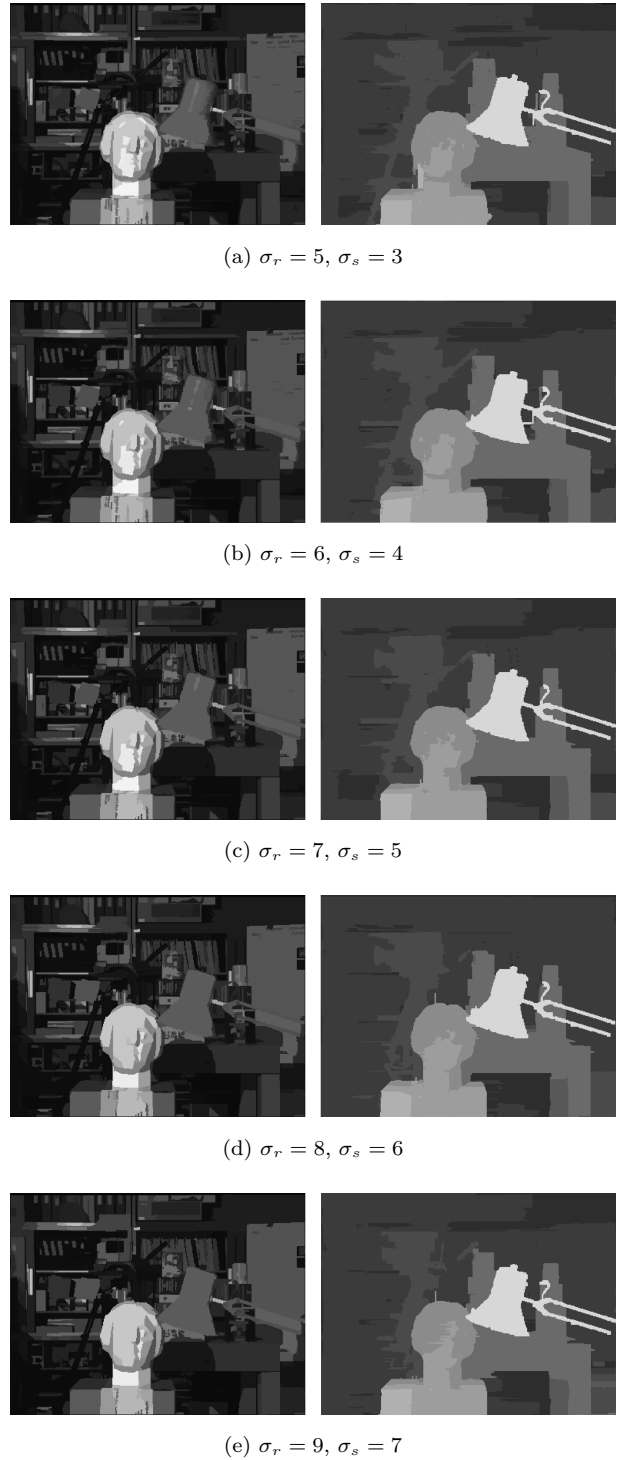


图 10 不同色彩距离和空间距离参数选择下的区域分割结果和对应的视差结果

Fig. 10 The results of segmentation and disparity map on different parameters of mean shift segmentation algorithm

的算法, 并采用目前学术界公认的由 Middlebury 网站 (<http://vision.middlebury.edu/stereo>) 发布的标准立体图片对算法进行了实验和评测. 图 11

给出了采用本文算法得到的若干标准立体图片的视差图. 为了验证算法的鲁棒性, 我们还采用实际拍摄的图像对本文算法进行了评测. 图 12 (见下页) 是所使用的实验图片和根据本文算法计算得到的视差图. 从图中可以看出, 本文算法对实际图像处理结果也令人满意, 较好地恢复了景物边缘的视差.

目前算法的运行时间与迭代次数有关, 在 Pentium 1.6 G 的 CPU 上对 Tsukuba 图像对迭代 4 次的时间约为 20 秒, 其中 Mean shift 图像分割的时间约为 8 秒.

实验结果表明, 本文算法在性能上可以和目前最好的算法相媲美. 具体比较结果示于表 1 (见下页). 其中, 对 Tsukuba 和 Venus 图像的全局匹配正确率排名第一. 实验中选取左图像为参考图像, 参数的设定^[12]为

$$\lambda_i = 0.5, \quad \lambda_{\text{disp}} = 5, \quad \lambda_{\text{occ}} = 5$$

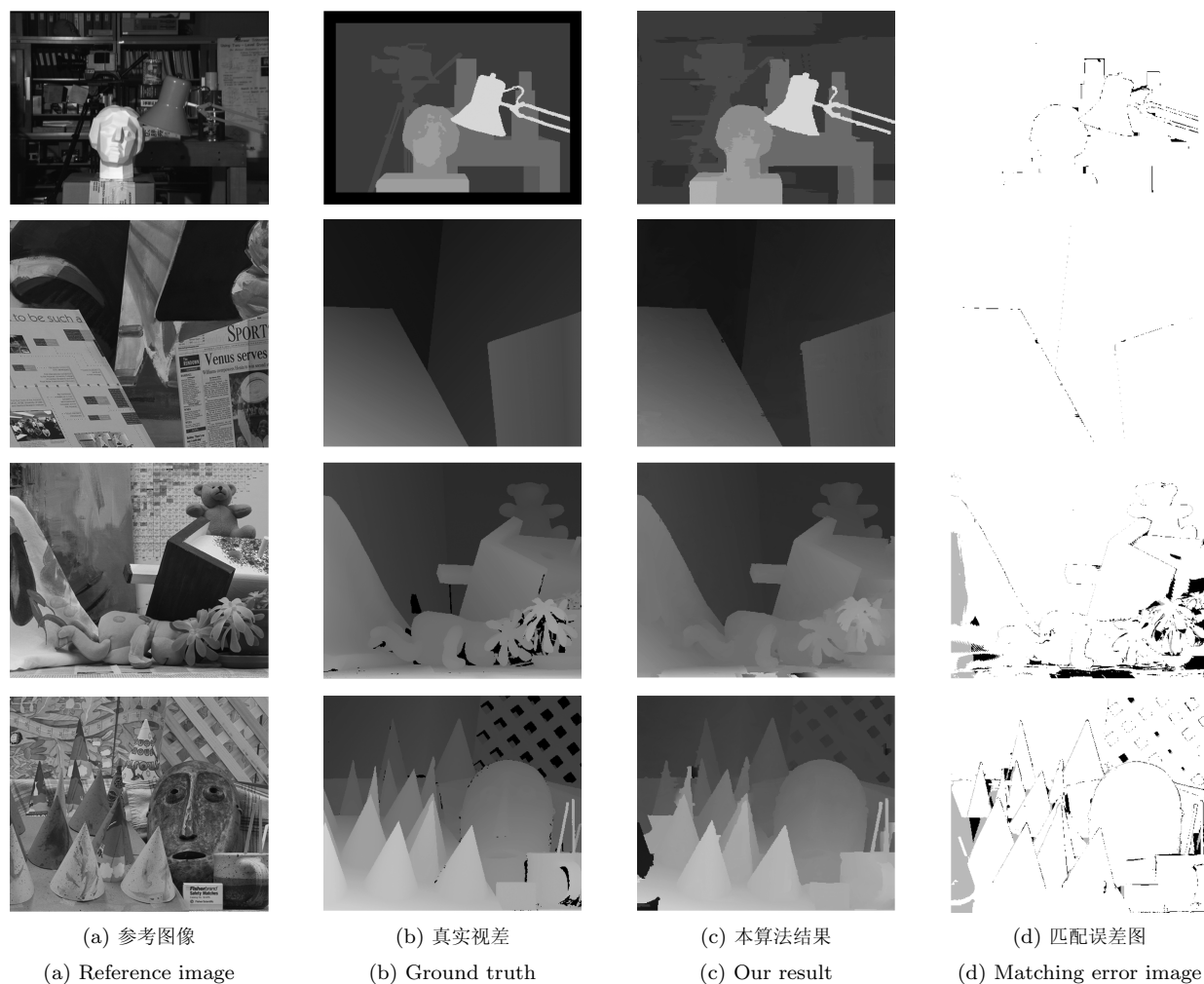


图 11 对标准图像 Tsukuba、Venus、Teddy、Cones 的处理结果, 当恢复视差与真实视差差别大于 1 个像素时视为匹配错误

Fig. 11 The disparity maps of the standard stereo image pairs on Middlebury website obtain by our algorithm

$$\sum_{i=1}^n w_{ij} = 1, \quad j = 1, 2, \dots, n$$

其中, λ_i 表示合作强度, λ_i 越大, 表示协调优化过程中邻域的作用越大, 反之, 则表示当前区域自身的作用更强. 本文中, λ_i 选为 0.5, 表示将自身和邻域的作用同等看待. 此外, w_{ij} 是传播系数, 它刻画了协调优化过程中每一个相邻区域对当前区域优化结果的影响. 本文中采用了等权重的方案, 即将与当前区域相邻的所有邻域同等看待. 实验结果表明, 本文算法所采用的固定参数选择方案对实验集中所包含的各种场景的立体图对均有上佳的表现, 显示出了很强的鲁棒性. 当然, 如果可以利用各区域视差参数的先验知识, 则可望取得更好的结果.

5 结论和展望

本文提出了一种基于区域的协同优化立体匹配

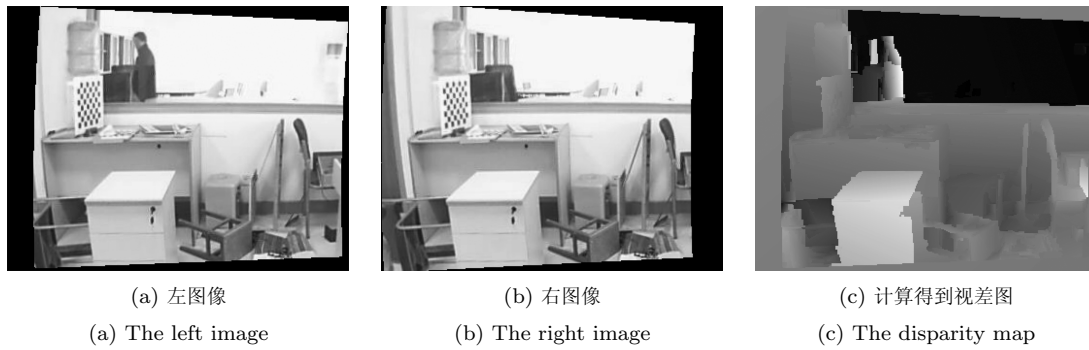


图 12 对实拍图像处理结果, 图像经过校准规则化

Fig. 12 The disparity maps of stereo image pairs taken in office environments

表 1 Middlebury 网站对标准图像的评测排名. 其中, 每种算法对应的第一行表示错误率, 第二行表示在所有参加评估的算法中该项目的排名, 粗体字表示本项目排名第一 (限于篇幅, 本表中仅列出了有代表性且性能较优的几种方法的测评结果¹.)

Table 1 The rank in Middlebury website (Here, the first number in each column shows the error rate of the method, the second number denotes the rank, and the bold number was ranked first.)

算法	平均 排名	Tsukuba			Venus			Teddy			Cones		
		nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc
AdaptingBP	3.1	1.11	1.37	5.79	0.10	0.21	1.44	4.22	7.06	11.8	2.48	7.92	7.32
		7	3	8	1	3	1	3	2	3	1	3	2
本文算法	3.1	0.87	1.16	4.61	0.11	0.21	1.54	5.16	8.31	13.0	2.79	7.18	8.01
		1	1	1	2	2	3	7	4	6	4	1	5
DoubleBP	4.1	0.88	1.29	4.76	0.13	0.45	1.87	3.53	8.30	9.63	2.90	8.78	7.79
		3	2	3	4	7	6	2	3	1	5	10	3
OutlierConf	4.8	0.88	1.43	4.74	0.18	0.26	2.40	5.01	9.12	12.8	2.78	8.57	6.99
		2	5	2	8	4	10	5	6	5	3	6	1
SubPix DoubleBP	6.5	1.24	1.76	5.98	0.12	0.46	1.74	3.45	8.38	10.0	2.93	8.73	7.91
		12	14	9	3	8	5	1	5	2	6	9	4
Adapt OvrSegBP	10.9	1.69	2.04	5.64	0.14	0.20	1.47	7.04	11.1	16.4	3.60	8.96	8.84
		24	22	7	5	1	2	15	8	12	12	12	11
SymBP+occ	11.8	0.97	1.75	5.09	0.16	0.33	2.19	6.47	10.7	17.0	4.79	10.7	10.9
		5	13	5	6	5	8	9	7	15	25	22	21

算法. 该算法从不完善的初始视差估计出发, 利用区域间的合作竞争关系, 通过协同优化使区域的匹配能量逐步极小化, 从而得到稠密和可靠的视差估计. 实验结果表明, 利用所提出的算法可以有效地抑制区域间的能量冲突, 最终得到接近于真实视差的视差估计结果.

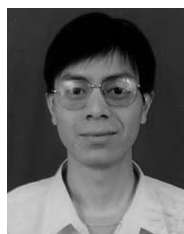
目前, 本文所提出的算法对图像区域分割错误的处理能力尚嫌不足, 如果分割区域边缘与视差跳跃边缘不一致, 将会导致错误匹配, 而且目前的算法对这种错误不具备修正能力; 另外, 目前的算法采用平面模型作为区域的视差模型, 从而限制了算法的通用性. 今后将考虑采用更为复杂的模型, 例如二次曲面、B 样条等以进一步改善视差计算结果.

¹ 更详细的测评结果参见 <http://vision.middlebury.edu/stereo/eval/>

References

- Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 2002, **47**(1-3): 7-42
- Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(11): 1222-1239
- Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(2): 147-159
- Sun J, Zheng N N, Shum H Y. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(7): 787-800
- Felzenszwalb P F, Huttenlocher D P. Efficient belief propagation for early vision. *International Journal of Computer Vision*, 2006, **70**(1): 41-54

- 6 Tao H, Sawhney H S, Kumar R. A global matching framework for stereo computation. In: Proceedings of the 8th International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001. 532–539
- 7 Lin M H, Tomasi C. Surfaces with occlusions from layered stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(8): 1073–1078
- 8 Li H, Chen G. Segment-based stereo matching using graph cuts. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2004. 74–81
- 9 Bleyer M, Gelautz M. A layered stereo matching algorithm using image segmentation and global visibility constraints. *Journal of Photogrammetry and Remote Sensing*, 2005, **59**(3): 128–150
- 10 Birchfield S T, Natarajan B, Tomasi C. Correspondence as energy-based segmentation. *Image and Vision Computing*, 2007, **25**(8): 1329–1340
- 11 Deng Y, Yang Q, Lin X Y, Tang X O. Stereo correspondence with occlusion handling in a symmetric patch-based graph-cuts model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(6): 1068–1079
- 12 Huang X F. Cooperative optimization for energy minimization: a case study of stereo matching [Online], available: <http://arxiv.org/pdf/cs.CV/0701057>, January 9, 2007
- 13 Comanicu D, Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603–619
- 14 Fusiello A, Roberto V, Trucco E. Efficient stereo with multiple windowing. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Juan, Puerto Rico: IEEE, 1997. 858–863
- 15 Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, **16**(9): 920–932
- 16 Veksler O. Stereo correspondence with compact windows via minimum ratio cycle. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(12): 1654–1660
- 17 Yoon K J, Kweon I S. Locally adaptive support-weight approach for visual correspondence search. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2005. 924–931
- 18 Gerrits M, Bekaert P. Local stereo matching with segmentation-based outlier rejection. In: Proceedings of the 3rd Canadian Conference on Computer and Robot Vision. Washington D. C., USA: IEEE, 2006



郑志刚 中国科学技术大学自动化系讲师。2008 年获中国科学技术大学博士学位。主要研究方向为立体匹配技术。

E-mail: zhengzg@ustc.edu.cn

(**ZHENG Zhi-Gang** Lecturer in the Department of Automation, University of Science and Technology of China. He received his Ph.D. degree

from University of Science and Technology of China in 2008. His main research interest is stereo correspondence.)



汪增福 中国科学技术大学自动化系教授。主要研究方向为视听觉信息处理, 智能机器人和模式识别。本文通信作者。

E-mail: zfwang@ustc.edu.cn

(**WANG Zeng-Fu** Professor in the Department of Automation, University of Science and Technology of China.

His research interest covers audio-video information processing, intelligence robots, and pattern recognition. Corresponding author of this paper.)