

基于斜平面平滑优化的半全局立体匹配

针对什么问题, 提出什么算法 (论文的题目)

肖进胜¹, 田 红¹, 邹文涛¹, 匡雨丽¹, 谢红刚²

(1. 武汉大学电子信息学院, 湖北武汉 430072; 2. 湖北工业大学电气与电子工程学院, 湖北武汉 430068)

指出某个现有方法的问题

该算法实现了什么效果

摘 要: 基于斜平面平滑的半全局匹配算法虽然能很好地处理深度不连续区域, 减少视差断裂现象, 具有很好的视差平滑效果, 但此方法计算量大, 并且若输入的左右图像对不是理想校正的, 最终得到的视差图中可能会出现黑色小方块。针对上述问题, 本文提出了一种基于斜平面平滑优化的双目立体视觉算法, 优化了视差斜平面拟合过程, 提高了斜平面平滑的运算速度; 另, 通过对初始视差图中无效的视差值进行临近插值填充, 很好地解决了视差图中的方块现象; 同时考虑到同一分割块内的像素梯度值应比较接近, 对图像分割时用到的能量函数增加梯度项, 使分割区域更加合理。实验结果表明, 本文算法能取得更好的视差图, 主观效果得到改善, 运算速度提高约 40%。

关键词: 半全局匹配; 斜平面平滑; 最临近插值; 梯度信息

中图分类号: TP391.41

文献标识码: A

文章编号: 0372-2112 (2018) 08-4835-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2018.08.006

用实验说明本方法的有效性

另外两个改进点是怎么做的, 实现了什么效果

Semi-global Matching Algorithm Based on Improved Slanted Plane Smoothing

XIAO Jin-sheng¹, TIAN Hong¹, ZOU Wen-tao¹, KUANG Yu-li¹, XIE Hong-gang²

(1. Electronic Information School, Wuhan University, Wuhan, Hubei 430072, China;

2. School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, Hubei 430068, China)

Abstract: The semi-global matching algorithm based on slanted plane smoothing can achieve good results in disparity discontinuous area. It can reduce disparity fracture phenomenon and makes the disparity image more smoothing. However, the slanted plane smoothing algorithm has huge computing quantity and long run time. If the input images are not very ideal, the final disparity image could have black small blocks. Aim at the above problems, binocular stereo vision algorithm base on improved slanted plane smoothing is proposed in this paper. The proposed algorithm not only increases the speed of calculation but also fills the invalid disparity in original disparity image using the nearest interpolation algorithms. Considering that the gradients of the pixels in the same segmentation should be close, the gradient energy is supplemented to optimize the energy function. So the proposed algorithm segments the image more reasonably. The experiments demonstrate that the proposed algorithm has a better disparity result. The subjective quality is improved and the time consuming is also increased by 40%.

Key words: semi-global matching; slanted plane smoothing; nearest interpolation; gradient information

1 引言

立体匹配一直是双目立体视觉的关键技术。传统的立体匹配算法可分成局部、全局和半全局三种。局部匹配算法比较待匹配点一定范围内的局部特性来进行匹配。Bigone 等人^[1]把图像的边缘信息作为匹配特征进行局部匹配, 然而没有考虑到各匹配窗口内图像特征的差异。全局匹配算法能较好地处理弱纹理、深度不连

续区域和遮挡区域。但在全局范围内求最优解非常耗时, 且不一定能得到最优解。马宁等扩展相位相关的小基高比立体匹配, 获得高精度亚像素级视差以满足小基高比摄影测量需求^[2]。2008 年, Hirschmüller 提出了半全局匹配方法^[3], 利用局部匹配的思想计算匹配代价, 再用全局匹配的动态规划思想从不同方向的扫描线上进行匹配代价的聚合, 然后选择最小代价体得到最终的视差。

收稿日期: 2017-05-05; 修回日期: 2018-04-19; 责任编辑: 蓝红杰

基金项目: 国家自然科学基金 (No. 61471272); 湖北省自然科学基金 (No. 2014CFB585, No. 2016CFB499)

目前国内外对双目立体视觉的研究方向主要分两个方面:一方面是对传统立体匹配算法的改进;另一方面是用最近热门的深度学习方法对左右图集进行训练得到深度信息. Jure Zbontar 在 2015 年提出用卷积神经网络计算图像块的相似度^[4],进行半全局匹配以及后续优化得到深度信息. 近年深度学习在立体视觉领域的运用越来越普遍^[5-6]. 这类方法得到的视差图虽然匹配精度高,但对设备要求高,训练时间长. 并且视差效果对训练图像集的选择有很强的依赖性. Zhang 提出二进制的立体匹配算法 BSM^[7],匹配代价体通过二进制位运算得到,运算速度快,但主观及客观效果存在很大缺陷. Gorkem Saygili 等人提出了一种自适应权重相似度测量算法^[8],不同的图像区域使用不同的相似度测量. Lee 在 2015 年提出 RWR^[9],把匹配代价聚合成超像素代价,用随机游走算法更新超像素之间所有估计视差值的匹配代价. 储珺^[10]提出基于线性滤波的树形结构动态规划立体匹配算法,能有效地改善传统动态规划算法中存在的条纹现象,减少弱纹理区域和边缘区域存在的误匹配. Yamaguchi 等人提出斜平面平滑模型^[11],根据图像的不同特征信息对图像进行分割并拟合. 该算法能很好地解决传统半全局匹配算法中的视差断裂现象,改善遮挡问题的处理. 但该算法计算速度慢,偶尔还会出现过平滑现象,边缘信息保持得不够完整.

针对斜平面平滑算法的缺陷,本文提出了一种改进的基于斜平面平滑优化的双目立体匹配算法. 通过分析斜平面平滑算法中视差斜平面拟合的时间复杂度,优化计算过程,提高运算速度,改善过平滑现象. 对初始视差图中无效的视差值进行最临近插值填充,很好地解决了视差图中的方块现象. 同时考虑到同一分割区域内像素点的梯度值应该比较接近,对图像分割时用到的能量函数进行优化,引进梯度信息,使分割得更加完整,边缘信息保持得更好. 实验结果表明,改进的斜平面平滑算法能使图像的主观及客观效果得到改善,运算速度有很大提高.

2 斜平面平滑半全局立体匹配

斜平面平滑算法^[11]先将图片分割成多个块,这些块又称超像素. 视差值的计算以块而非像素点为单位进行. 在这个块内可以建立视差斜平面,该平面保存了分割块内像素点的拟合视差值. 对于分割块 i 内的视差平面 θ_i 中任意像素点 p ,通过 A_i 、 B_i 、 C_i 三个参数,其拟合视差值可如下式定义:

$$\hat{d}(p, \theta_i) = A_i p_x + B_i p_y + C_i \quad (1)$$

斜平面平滑算法的基本思想用能量的形式来描述. 将分割块内初始视差距离拟合视差斜平面较远的

像素点定义为外点,其余点为内点. 设 f 标志是否为外点. 设 o 为相邻分割块的分配标签, q 为 p 的相邻像素点, s_p 、 s_q 是 p 、 q 所属的分割块索引. 总能量函数定义如下:

$$E = E_c + E_p + E_d + E_s + E_l + E_b \quad (2)$$

这个能量函数并非一次性定义并求解,下文将分步骤介绍该式各部分原理.

第一项: $E_c = \sum_p E_{col}(p, c_{s_p})$,色彩能量是边界点的 $L^* a^* b$ 三分量和相邻像素点所属分割块的平均 $L^* a^* b$ 三分量差值的平方和.

第二项: $E_p = \lambda_{pos} \sum_p E_{pos}(p, \mu_{s_p})$,位置能量是边界点的像素坐标和相邻分割块的平均像素坐标的距离平方.

第三项: $E_d = \lambda_{depth} \sum_p E_{depth}(p, \theta_{s_p}, f_p)$,深度能量,当像素点是内点时是初始视差值和拟合视差值的差值平方和. 否则深度能量为常数.

第四项: $E_s = \lambda_{smo} \sum_{\{i, j\} \in N_{smo}} E_{smo}(\theta_i, \theta_j, \rho_{i, j})$,平面平滑能量用于衡量相邻分割块的情况. 共面的分割块将合并,视差断裂的分割块之间形成边界;被遮挡的分割块区域具有更小的视差.

第五项: $E_l = \lambda_{com} \sum_{\{i, j\} \in N_{com}} E_{prior}(o_{i, j})$,复杂度能量鼓励两个区域共面,使图片的复杂度降低. 其中,遮挡的代价 > 视差较链的代价 > 共面的代价.

第六项: $E_b = \lambda_{bou} \sum_{\{p, q\} \in N_b} E_{bou}(s_p, s_q)$,为边界长度能量. 对于相邻且不属于同一分割块的点增加边界长度能量.

这些能量结合后就构成了整体能量,在计算时,考虑各项能量尽量减小. 该算法先根据图像的 E_c 、 E_p 和 E_b 对均匀分割的图像进行调整;其次结合半全局匹配得到的初始视差,对调整后的分割块进行视差斜平面初步拟合;使用初次拟合得到的 E_d 对分割区域进行二次调整,最后根据分割块、边界点的类型构成拟合平面的 E_s 进行能量函数最小化,对拟合平面的参数进行再次调整,使 E_s 最小,最终生成稠密、平滑的视差图.

3 改进的斜平面平滑算法

斜平面平滑算法考虑到遮挡区域和视差断裂区域,对不同的区域进行视差斜平面拟合. 该算法用能量函数最小化的思想进行分割,能够得到更加完整的分割图像. 虽然该算法能更好地解决视差断裂现象,但是该算法运算量大,时间复杂度高,而且图像分割中并没有考虑到图像的梯度信息. 如果输入的左右图像对不是严格校正的,得到的初始视差图就会有很多误匹配

点, 这些误匹配点经过视差斜平面拟合后就会形成“黑方块”。

针对上述问题, 我们通过分析斜平面平滑算法中视差斜平面拟合的时间复杂度, 减少计算量。同时对能量函数用梯度信息进行优化, 使分割得更加完整, 边缘信息保持得更好。并对初始视差图中无效的视差数据进行插值填充, 以解决视差图中的方块现象。改进的斜平面平滑算法流程图见图 1, 其中方框部分是本文改进的地方。其算法描述见算法 1。

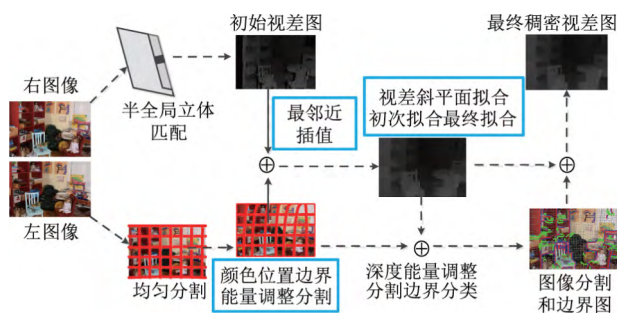


图1 本文改进的斜平面平滑算法流程图

算法 1 改进后的斜平面平滑算法

1. 输入左右图像对, 左图像为参考图像, 右图像为匹配图像;
2. 利用半全局立体匹配对左右图像进行匹配, 得到初始视差图;
3. 按照超像素数均匀分割左图像。从分割区域的边界处提取出边界点, 计算边界点和它相邻分割块的 E_c 、 E_p 、 E_b 以及梯度能量 E_g (定义见式(9))。调整 p 所属的分割块使 p 所在的分割块能量最小;
4. 对初始视差图中的无效视差数据进行插值填充, 并利用最小二乘法进行视差斜平面初步拟合, 使拟合平面和初始视差值产生的 E_s 最小;
5. 考虑初次拟合得到的 E_d , 对分割区域进行二次调整。将边界点分成遮挡点、视差断裂点和共面点三类, 通过边界点衡量相邻分割块的关系;
6. 完成对边界类型标签分配后, 进行平滑操作, 利用边界类型信息对视差斜平面进行不同的平面拟合, 对拟合平面的参数进行再次地调整, 使 E_s 最小, 生成最终的视差图。

3.1 斜平面平滑算法运算速度的优化

通过对斜平面平滑算法的分析, 作者发现对视差斜平面初步拟合和最终拟合这两部分耗时最大。初步拟合中, 原算法结合了最小二乘法的 RANSAC 算法^[12]来对初始视差图进行视差斜平面的拟合。从分割块内随机选取三个不同点确定视差平面, 根据初始视差与拟合视差平面的距离来区分内点和外点。通过多次迭代找到内点最多的那个平面。

确定所有内点后, 使用最小二乘法对视差斜平面进行再次拟合。最小二乘法即是让方程组计算出来的数据跟实际视差数据的残差平方和 S 最小:

$$S = \sum_{i=0}^{n-1} (a_0 x_i + a_1 y_i + a_2 - z_i)^2 \quad (3)$$

其中 $i=0, 1, \dots, n-1$, n 表示该分割块内点的个数, z_i 表示分割块中内点 $p_i(x_i, y_i)$ 的初始视差值。通过下述方程组:

$$\begin{cases} a_0 \sum x_i^2 + a_1 \sum x_i y_i + a_2 \sum x_i = \sum x_i z_i \\ a_0 \sum x_i y_i + a_1 \sum y_i^2 + a_2 \sum y_i = \sum y_i z_i \\ a_0 \sum x_i + a_1 \sum y_i + a_2 n = \sum z_i \end{cases} \quad (4)$$

解出视差斜平面的系数 a_0, a_1, a_2 , 使得 S 最小, 得到分割块的初步拟合视差平面 $z = a_0 x + a_1 y + a_2$ 。分割区域内像素点的视差值即可通过拟合平面来计算。

设输入图像的长和宽分别为 H 和 W , 图像分割后有 M 个分割块, 每个分割块中有 N 个像素点, 则有 $H \cdot W = M \cdot N$ 。设视差斜平面初步拟合过程中剔除外点的计算量为 T , 内点视差数据代入式(4)求得视差斜平面参数的计算量为 L , 因此遍历一个分割块内所有点的计算量为 $T \cdot N + L$, 对整幅图像进行初步拟合的计算量为 $M \cdot (T \cdot N + L)$ 。

考虑到后续优化中增加了对半全局匹配初始视差的插值步骤, 分割块内像素点的视差值极少出现异常。因此再用 RANSAC 方法检测并剔除外点不仅没有必要, 反而会增加计算量。因此我们改进后的斜平面平滑算法直接用分割块内像素点的视差数据进行初步的视差斜平面拟合。根据上述分析, 改进后的算法该部分计算量不超过 $M \cdot L$, 减少了 $M \cdot T \cdot N$, 加快了运算速度。

其次, 在斜平面最终拟合过程中, 通过不断利用边界分类信息进行拟合迭代, 计算量相当大。本文使用 Middlebury 平台中的评价指标 bad2.0 (其具体定义见 4.1) 对每次迭代输出的视差图进行测试。本文设计终止迭代函数来对平面拟合的效果与速度进行平衡。具体公式如下:

$$\frac{E_i - E_{i+1}}{E_i} \leq \varepsilon \quad (5)$$

其中 E_i 和 E_{i+1} 分别表示第 i 次和第 $(i+1)$ 次迭代后视差图的 bad2.0。在本文实验中设置为 0.01。迭代效果参考图 2、图 3。对最终视差斜平面拟合迭代 4 次后, 视差图的平滑效果已到达最优, 继续增加迭代次数, 椅子的边缘反而会出现锯齿现象, bad2.0 值趋于平稳。

对视差斜平面的初步拟合和最终拟合两部分优化后进行测试, 测试平台为 Intel i3-3220 处理器, 4G 内存, Windows7 32 位操作系统, 表 1 给出具体的运算时间对比。可以看出, 改进后的斜平面平滑算法运算时间减少了 40%。如图 4 所示, 图 b 和图 c 优化前后的视差图主观效果比较接近。

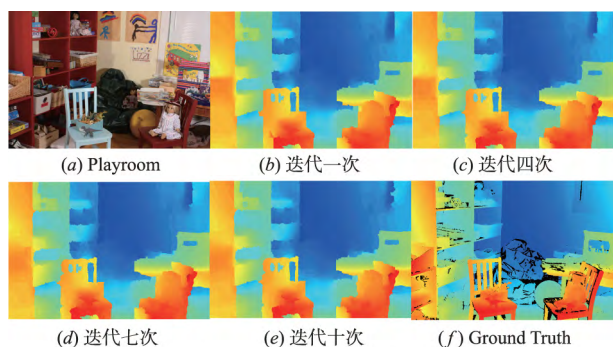


图2 迭代次数对最终视差图的主观影响

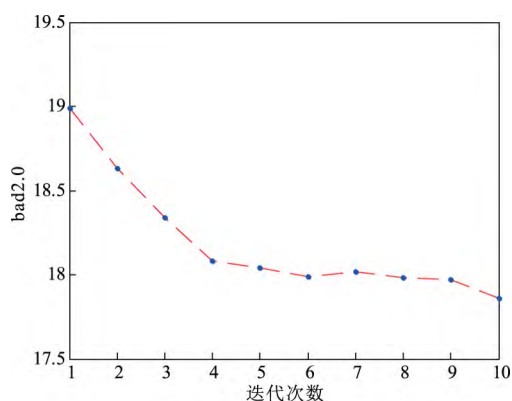


图3 Playroom视差图和迭代次数的关系

表1 运算时间对比(单位:s)

图像	大小	优化前	优化后
Piano	707 × 481	32.838	19.271
Playroom	699 × 476	33.152	19.283
Playable	680 × 463	31.138	18.860
Adirondack	718 × 496	31.918	19.626
Vintage	722 × 480	32.324	19.00
Motorcycle	741 × 497	34.922	19.429

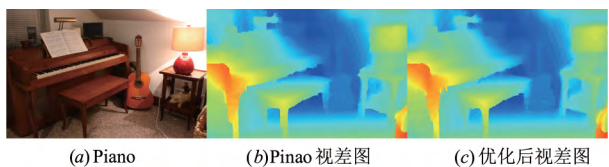


图4 运算速度优化前后的视差图效果对比

3.2 初始视差图的插值

如果输入的左图右图并不理想,使得半全局立体匹配存在很多误匹配点,经过左右一致校验后,这些误匹配点就会被强制视差值为0。如图5所示,可以看出,用半全局立体匹配(SGM)算法得到的视差图会出现一大片误匹配区域,视差值被置为0。分割块内的无效视差值,会对最终拟合的视差平面产生很不利的影响。

为了尽可能地不增加斜平面平滑(SPSS)算法的运算时间,我们使用执行效率高且占用内存小的最邻近

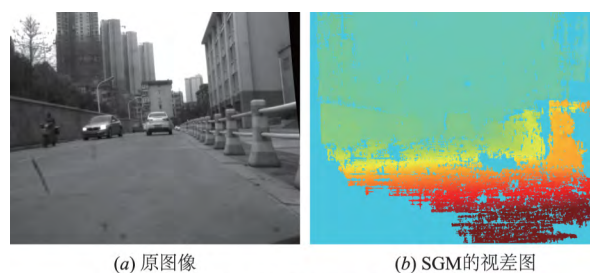


图5 原图和SGM得到的视差图

插值法。通过水平方向和垂直方向的二次遍历,将距离待求像素点最近的有效视差值赋值给待求像素点。插值填充后,利用插值后的视差数据能很好地拟合视差斜平面,减少误匹配点对视差图的干扰,使得视差结果精度更高。如图6所示,插值填充后视差图中的黑色方块被移除,可以得到更加平滑精确的视差图。

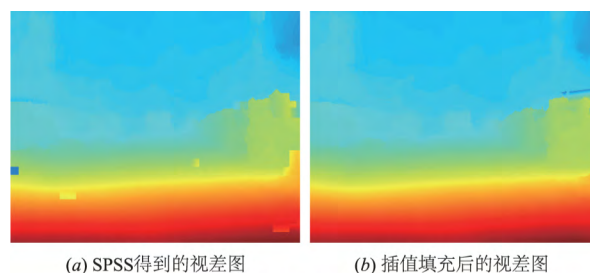


图6 视差图和插值填充后的视差图

3.3 基于梯度信息的能量函数优化

斜平面平滑算法中用于图像分割的能量函数式(2)并没有考虑到图像的梯度信息。考虑到边界处的梯度值会发生突变,利用叠加了图像梯度信息的能量函数进行分割,分割区域更加合理。

本文使用计算简单效率高的局部差分 Roberts 算子来提取边缘信息,其边界定位比较准。简化的计算公式为:

$$G(x, y) = |I(x, y) - I(x+1, y+1)| + |I(x+1, y) - I(x, y+1)| \quad (6)$$

其中 $I(x, y)$ 表示像素点 (x, y) 的灰度强度。斜平面平滑算法中用于图像分割的能量有色彩能量、位置能量、深度能量以及边界长度能量。这里将 Roberts 算子提取出的图像梯度信息叠加在能量函数中。改进后的分割能量函数如下:

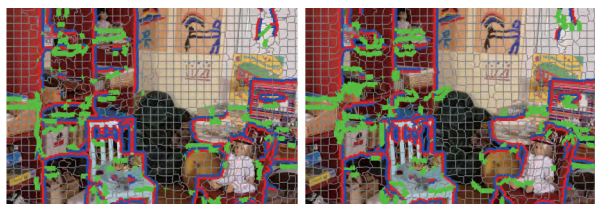
$$S_p = \operatorname{argmin}\{E_c + E_p + E_d + E_b + E_g\} \quad (7)$$

其中梯度能量表示为:

$$E_g = \lambda_{gra} \sum_p E_{gra}(p, G_{S_p}) = \lambda_{gra} \sum_p \|p - G_{S_p}\|^2 \quad (8)$$

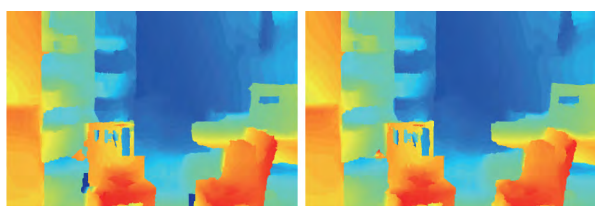
G_{S_p} 代表相邻像素点所属分割块内的平均梯度值。梯度能量是边界点的梯度值和相邻像素点所属分割块的平均梯度值的差值平方和。

作者对改进能量函数的斜平滑算法进行测试, 实验结果如图 7、8 所示。图 7 是图像分割示意图, 其中带颜色的粗线条表示边界线。对比 (a) 和 (b), 可以看出 (b) 中球的边界信息更加丰富。



(a) 原分割图 (b) 改进能量函数后的分割图

图7 改进能量函数前后的图像分割示意图



(a) 原视差图 (b) 改进能量函数后的视差图

图8 改进能量函数前后的最终视差图

图 8 是运算速度优化后, 改进能量函数前后的视差图。图 8 (b) 相对于图 8 (a) 而言, 两个凳子左边的误配区域消失, 图像分割更加完整。

作者还使用 Middlebury 平台上的 bad2.0 这一客观指标进行对比测试。从表 2 中可以看出, 能量函数在增加了梯度信息后, 像素点的误差比例 bad2.0 下降了约 1%, 匹配效果更好。

表 2 改进能量函数前后 bad2.0 对比(单位: %)

图像	大小	优化前	优化后
Piano	707 × 481	11.35	10.80
Playroom	699 × 476	14.30	13.32
Playable	680 × 463	10.11	9.24
Adirondack	718 × 496	4.10	4.04
Vintage	722 × 480	5.06	4.86
Motorcycle	741 × 497	7.56	6.95

4 实验结果与分析

衡量各双目立体匹配算法的匹配效果, 目前应用较多的是 Middlebury 测试平台^[13], 因此本文也使用 Middlebury 测试平台来验证算法的有效性。

4.1 实验结果评估方法

该平台上常用的评价指标有误差比例和平均误差。

误差比例: 视差值误差在一定范围内的像素点占全部像素点的比例, 有 bad0.5、bad1.0、bad2.0 和 bad4.0 四种。本文采用 bad2.0 指标。

$$\text{Bad}\delta_d = \frac{1}{N} \sum_{(x,y)} (|d(x,y) - d_T(x,y)| > \delta_d) \quad (9)$$

其中 δ_d 是误差阈值, $d_T(x,y)$ 是该像素点的实际视差值, $d(x,y)$ 是各匹配算法计算得到的视差值, N 是像素点数量。Middlebury 平台还按照不同的区域计算视差误差比例, nonocc, all, disc 分别是非遮挡区域, 全部区域, 深度不连续区域的视差误差百分比。

平均误差是所有像素点的视差误差的平均值。

$$\text{Avgerr} = \frac{1}{N} \sum_{(x,y)} |d(x,y) - d_T(x,y)| \quad (10)$$

4.2 实验结果对比与分析

本文选用 Middlebury 测试平台上 2014 年图像数据库进行了大量的测试, 限于篇幅, 这里只给出有代表性的 Playtable、Piano、Playroom、Adirondack 四个场景的测试结果。本文对比使用 BSM^[7]、SGM^[3]、Elas^[14]、RWR^[9]、MC-CNN^[4]、及斜平面平滑算法 SPSS^[11]。Elas 算法是 Middlebury 平台自带的算法, 该算法把待匹配图像的特征点作为支撑点形成三角进行立体匹配。图 9 至 11 是视差主观效果对比, 用不同的色调表示不同的视差值, 红色视差值较大, 深度小; 变成蓝紫色, 视差值变小, 深度变大。

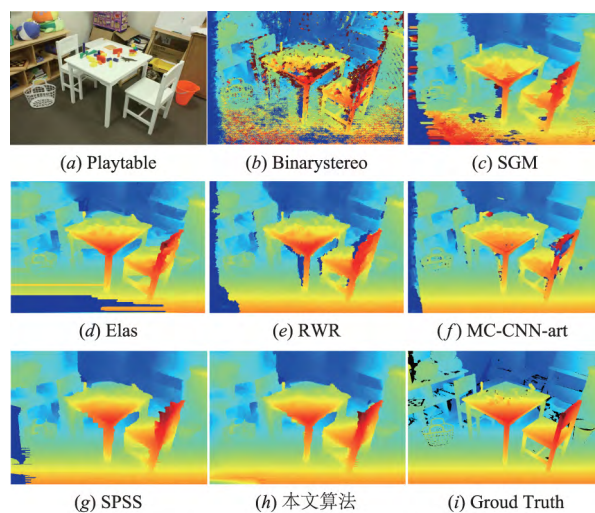


图9 Playtable场景的视差对比图

从主观效果来看, 本文算法具有较少的图像噪点, 整体视差效果最好。Elas 算法对噪点有所改善, 也进行了平滑约束, 但是有严重的拉伸现象, 如图 10 所示。而 RWR 算法对图像边缘部分处理效果不好, 如图 11 中椅子的边缘部分和图 9 中桌子的边沿部分。MC-CNN 算法虽然能很好地保持边缘信息, 得到相对稠密和精确的视差图, 但是算法训练时间长, 对测试样本有较高要求, 如果测试图片与训练图片集的场景差别很大, 视差图效果并不理想。而本文算法与原始斜平面平滑 SPSS 算法相比, 视差图的边界更接近 Ground Truth, 图像两侧视差效果更好。

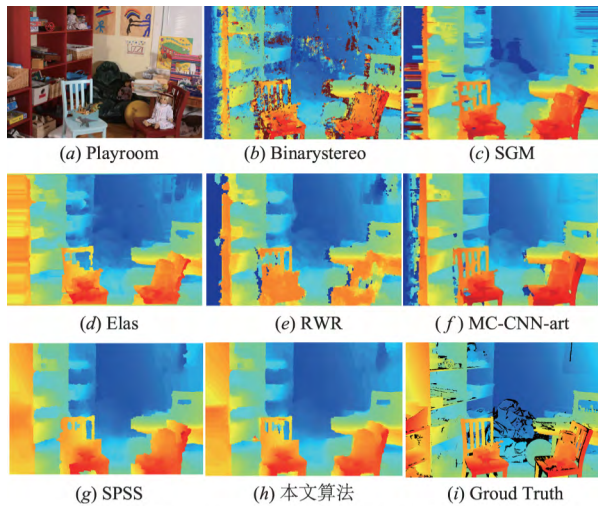


图10 Playroom场景的视差对比图

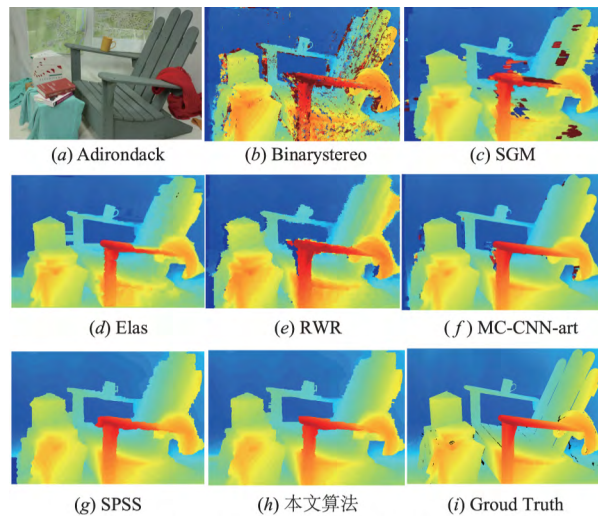


图11 Adirondack场景的视差对比图

本文还使用客观指标误差比例 bad2.0 和平均误差 Avgerr 来对比视差精度. 我们按非遮挡区域 (Nonocc) 和图像所有区域 (All) 分别进行对比, 客观指标对比结果如表3 表4.

表3 客观指标 bad2.0 对比 (单位: %)

	Playtable		Playroom		Adirondack	
	Nonocc	all	Nonocc	all	Nonocc	all
BSM	55.29	35.35	25.21	26.47	10.48	14.18
SGM	38.8	44.2	26.3	34.5	14.9	18.7
Elas	16.96	18.01	16.6	17.98	6.12	5.34
RWR	17.31	10.99	18.40	24.11	9.20	6.42
MC-CNN	15.1	23.3	14.6	25.6	5.15	10.0
SPSS	11.57	9.34	13.17	12.16	4.54	3.81
proposed	9.34	9.24	13.62	13.32	4.38	4.04

表4 客观指标 Avgerr 对比 (单位: %)

	Playtable		Playroom		Adirondack	
	Nonocc	all	Nonocc	all	Nonocc	all
BSM	13.63	9.7	13.53	9.11	3.58	5.42
SGM	15.7	18.5	4.24	6.15	1.85	2.46
Elas	2.40	4.55	2.76	1.90	1.09	1.88
RWR	4.69	1.54	2.72	2.85	1.59	1.02
MC-CNN	3.11	9.71	2.27	18.3	0.76	4.24
SPSS	2.95	1.32	2.66	1.55	1.37	1.04
proposed	1.05	1.15	2.42	1.52	1.28	1.06

各对比算法的平均运行时间列表如表5.

表5 各算法平均运行时间对比 (单位: s)

算法	运行时间
BSM	890
SGM	12.96
Elas	12.42
RWR	51.48
MC-CNN	67
SPSS	32.72
本文算法	19.38

从表3、4中可以看出, 对于图像全部区域 all, SPSS 算法的 bad2.0 值较小. 由于本文算法是在 SPSS 算法基础上进行运算速度的优化, 速度提高了 40%, 故以视差精度为代价, 小部分图像的 bad2.0 值稍微有所下降, 但本文算法的平均误差 Avgerr 比 SPSS 小. 而对于非遮挡区域 nonocc, 本文算法处理效果整体更好, bad2.0 和 Avgerr 值均较小. 因此, 本文算法适用于遮挡物较少、比较开阔的场景. 而 MC-CNN 算法只是对单幅图像 Adirondack 处理效果较好. 最后, 从表5可以看出, 在算法耗时上, 我们的方法也有较大的优势.

5 结论

针对基于斜平面平滑的半全局立体匹配算法存在的问题, 我们从效果和运算速度这两个方面对基于斜平面平滑算法进行改进. 针对斜平面平滑算法的计算时间长、边缘表现得不突出等问题, 本文提出了一种改进的基于斜平面平滑的双目立体视觉算法. 该算法不仅减少计算量, 提高运算速度, 还对初始视差图中的无效视差数据进行插值填充, 很好地解决了最终视差图中的黑色方块的现象. 同一分割区域内像素点的梯度值应比较接近, 因此我们利用梯度信息对图像分割时的能量函数进行优化, 使图像分割更加完整.

参考文献

- [1] Bignone F, Henricsson O, Fua P, et al. Automatic extraction of generic house roofs from high resolution aerial imagery [A]. Computer Vision-ECCV '96 [C]. Berlin, Heidelberg: Springer LNCS: 1064.
- [2] 马宁, 门宇博, 门朝光, 李香. 基于扩展相位相关的小基高比立体匹配方法 [J]. 电子学报, 2017, 45 (8): 1827 - 1835.
Ma Ning, Men Yubo, et al. A small baseline stereo matching method based on extended phase correlation [J]. Acta Electronica Sinica 2017 45(8): 1827 - 1835. (in Chinese)
- [3] Hirschmuller H. Stereo processing by semiglobal matching and mutual information [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence 2008 30(2): 328 - 341.
- [4] Bontar J, Lecun Y. Stereo matching by training a convolutional neural network to compare image patches [J]. Journal of Machine Learning Research 2015 17(1): 2287 - 2318.
- [5] Kendall A, Martirosyan H, Dasgupta S, et al. End-to-end learning of geometry and context for deep stereo regression [J]. arXiv 1703.04309v1 2017.
- [6] Bailer C, Varanasi K, Stricker D. CNN-based patch matching for optical flow with thresholded hinge embedding loss [A]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. IEEE 2017. 2710 - 2719.
- [7] Zhang K, Li J, Li Y, et al. Binary stereo matching [A]. IEEE 21st International Conference on Pattern Recognition (ICPR) [C]. IEEE 2012. 356 - 359.
- [8] Y Saygili G, Laurens V D M, Hendriks E A. Adaptive stereo similarity fusion using confidence measures [J]. Computer Vision & Image Understanding, 2015, 135 (6): 95 - 108.
- [9] Lee S, Lee J H, Lim J, et al. Robust stereo matching using adaptive random walk with restart algorithm [J]. Image & Vision Computing 2015 37(5): 1 - 11.
- [10] 储珺, 龚文, 缪君, 等. 基于线性滤波的树结构动态规划立体匹配算法 [J]. 自动化学报, 2015, 41 (11): 1941 - 1950.
Chu Jun, Gong Wen, Miao Jun, et al. A Tree structure dynamic programming stereo matching algorithm based on linear filtering [J]. Acta Automatic Sinica 2015 41(11): 1941 - 1950. (in Chinese)
- [11] Yamaguchi K, Mcallester D, Urtasun R. Efficient Joint Segmentation, Occlusion Labeling, Stereo and Flow Estimation [A]. Computer Vision-ECCV 2014 [C]. Springer, 2014. 756 - 771.
- [12] 彭红, 肖进胜, 沈三明, 等. 一种基于随机抽样一致性的车道线快速识别算法 [J]. 上海交通大学学报, 2014, 48 (12): 1721 - 1726.
Peng Hong, Xiao Jinsheng, Shen Sanming, et al. A fast algorithm based on RANSAC for vision lane detection [J]. Journal of Shanghai Jiaotong University, 2014, 48 (12): 1721 - 1726. (in Chinese)
- [13] Scharstein D, Szeliski R. <http://vision.middlebury.edu/stereo/> [EB/OL].
- [14] Geiger, Andreas, Roser, et al. Efficient large-scale stereo matching [A]. Computer Vision-ACCV 2010 [C]. Berlin Heidelberg: Springer 2010. 25 - 38.

作者简介



肖进胜 男, 1975 年 7 月出生于湖北武汉, 博士, 武汉大学电子信息学院副教授, 主要研究方向为视频图像处理, 计算机视觉.

E-mail: xiaojsh@whu.edu.cn



田红 女, 1994 年 8 月出生于湖北荆州, 武汉大学电子信息学院硕士生, 主要研究方向为视频图像处理.

E-mail: tianh2016@whu.edu.cn

邹文涛 男, 1994 年 9 月生, 武汉大学电子信息学院硕士生, 主要研究方向为视频图像处理.

匡雨丽 女, 1992 年生, 武汉大学电子信息学院硕士生, 主要研究方向为视频图像处理.

谢红刚 男, 1974 年生, 工学博士, 湖北工业大学电气与电子学院讲师, 主要从事图像处理、图像复原等领域的研究.