

融合灰色尺度的代价聚合的立体匹配*

韩先君, 杨红雨

(四川大学 计算机学院, 四川 成都 610065)

通讯作者: 杨红雨, E-mail: bchxjbc@163.com



摘要: 由粗略到精细, 分层策略和跨尺度的代价聚合在一定程度上有效地扩展了代价聚集并且能够生成高精度的视差图. 这类方法致力于在弱纹理区域找到正确的匹配点从而提高匹配率. 然而, 这类方法必须以多尺度为前提, 通常需要借助图像金字塔. 另外, 误差的传播以及薄壁结构的复原不理想限制了它们的应用. 针对弱纹理匹配的问题, 提出了一种通用的融合灰色尺度的代价聚合的立体匹配框架. 鉴于高斯滤波后的灰度图像能够更好地表示匹配图像对中的弱纹理区域, 该代价聚合融合了灰度图像的代价聚合. 同时, 算法不需要降采样以及建立图像金字塔, 这加快了聚合速度. 此外, 还引入了引导图像滤波和快速加权中值滤波, 用于代价聚合和视差求精. 同时, 在进行视差选择时, 为了避免 WTA (winner-take-all) 带来的歧义, 利用代价聚合后最小值和次小值之间的相互关系来确定最后的视差值. 最终, 在 Middlebury 测试平台上的实验结果表明: 融合灰色尺度的代价聚合的立体匹配能够有效地提高视差的精度.

关键词: 引导图像滤波; 局部立体匹配; 快速加权中值滤波; 融合灰色尺度

中文引用格式: 韩先君, 杨红雨. 融合灰色尺度的代价聚合的立体匹配. 软件学报, 2018, 29(Suppl. (2)): 44-53. <http://www.jos.org.cn/1000-9825/18016.htm>

英文引用格式: Han XJ, Yang HY. Fusion of gray scale cost aggregation for stereo matching. Ruan Jian Xue Bao/Journal of Software, 2018, 29(Suppl. (2)): 44-53 (in Chinese). <http://www.jos.org.cn/1000-9825/18016.htm>

Fusion of Gray Scale Cost Aggregation for Stereo Matching

HAN Xian-Jun, YANG Hong-Yu

(College of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: Coarse-to-Fine (CTF), hierarchical strategy, and cross-scale cost aggregation have efficiently expanded cost aggregation methods and yield a highly accurate disparity map to some extent. They are committed to providing a good trick to find the correct matching points in the weak texture region. However, these methods must be multi-scale as the prerequisite and usually need the assistance of image pyramid. They are limited to the propagation of errors from coarse to fine levels and poor recovery of thin structures. In this study, a generic fusion of gray scale cost aggregation framework is proposed which encourages the initial cost aggregation to integrate the cost aggregation of gray image. The main purpose of the gray image after Gaussian filter is to match the corresponding pixels in the weak texture region of the image better. Meanwhile, it does not need to scale down to build the image pyramid and aggregate cost at each scale and thus accelerate the step of cost aggregation. Furthermore, guided image filtering and fast weighted median filtering are introduced in this study for cost aggregation and disparity refine. In addition, to avoid choosing ambiguity that WTA (winner-take-all) brings, the interrelationship between the minimum value of cost aggregation and the second smallest value is utilized to determine the final disparity. It is shown that the fusion of gray scale cost aggregation framework is important as it effectively leads to significant improvement evaluated on Middlebury.

Key words: guided image filtering; local stereo matching; fast weighted median filtering; fusion of gray scale

* 基金项目: 国家自然科学基金(61572333)

Foundation item: National Natural Science Foundation of China (61572333)

收稿时间: 2017-06-01; 采用时间: 2018-11-08

从立体图像对中生成了稠密的视差图是计算机视觉中最热门的课题之一^[1].它广泛应用于三维重建、机器人导航和虚拟现实等方面.大量的立体匹配算法希望在生成高精度视差图的同时降低计算量.根据视差估计的策略不同,立体匹配算法主要分为两种类型(局部和全局方法)^[2].一般来说,全局方法集中于建立能量函数,同时优化所有像素的视差.全局的立体方法主要包括 TSGO^[3]、置信度传播^[4]、动态规划和图割^[5].图像亮度、颜色或局部特征则被广泛用于局部立体匹配方法,以确定每个像素的视差.

大多数立体匹配都遵循 4 步:1) 初始代价计算;2) 代价聚合;3) 视差优化;4) 视差求精.除此之外,预处理步骤是可选的.全局方法通常省略代价聚合这一步骤,认为视差的求取是一个全局优化的过程,并假定整个视差图是光滑的.与局部立体匹配方法相比,全局立体匹配算法可以获得更高的精度.然而,局部立体匹配方法更快,更符合实际操作.因此,在研究了目前主要的代价聚合方法之后,我们提出了一种基于融合灰色尺度的代价聚合的局部立体匹配方法.

本文的局部立体匹配算法既能保持其快速运行的特性,又能生成高精度的视差图.本文的主要贡献如下.

(1) 通过分析由粗略到精细(CTF),分层策略和跨尺度的代价聚合的不足,提出了一种类似的框架——基于融合灰度尺度的代价聚合.该方法致力于提高弱纹理的匹配精度.

(2) 引入了一种新的视差求精策略——基于引导滤波核的快速加权中值滤波.与以前的方法相比,该方法能够快速并且有效地保护视差图的边缘.

(3) 为了解决 WTA 带来的歧义选择问题,引入可信度来获取最终的视差.

本文第 1 节讨论算法的相关工作.融合灰度尺度的代价聚合方法将在第 2 节中详细描述.第 3 节描述在 Middlebury 立体匹配评价标准中最常用的 4 幅基准图像对本文算法的测试.

1 相关工作

局部立体匹配需要计算两幅图像对应对应的两两像素之间的相似程度以得到初始的匹配代价.基于窗口的代价函数主要包括绝对差之和^[6]、平方差和以及归一化后的互相关^[7]方法.随后,针对目标边界附近的离群点提出了一些鲁棒性的算法,如 MI(mutual information)^[8,9]算法、非参数变换 CENSUS^[10]、基于小波变换的算法^[11]和局部二进制模式^[12].另外,基于梯度的^[13]和基于 Gabor 特征的^[14]也通常被用来作为初始代价量.在上述方法的基础上,文献[15,16]对这些代价项进行了改进,并取得了很好的效果.所有这些方法都有各自的优缺点,而合并这些代价量是一个很好的策略,因为联合的代价项更能体现像素之间的相似程度.例如,文献[17]中的算法利用了 census 变换系数的汉明距离以及像素亮度值差的组合,提高了视差图的精度.文献[18]则利用了颜色和梯度值的组合,也取得了很好的效果.

初始的匹配代价通常含有大量的噪声,用聚合函数将所有像素的初始代价量聚合起来消除可能存在的噪声,这通常在一个窗口内进行.它假定视差在同一个相似区域有着相同的视差.在这一领域中最引人注目的方法是自适应权重^[19],它根据颜色相似性和邻近原则调整权重值.然而,权重要求对窗口中的每个像素进行独立计算,其主要缺点是计算成本高.因此,很多方法的提出旨在减少其执行时间.例如,采用积分图像的快速双边滤波^[20]是典型的加快聚合速度的立体匹配方法.此外,文献[21,22]提出了一种利用最小生成树对图像进行聚合的方法.文献[23]利用在一个计算复杂度较低的支持区域中采用指数步长的方法,也取得了很好的效果.

在代价聚合方面,引导图像滤波^[24]是在精度和速度上取得的另一个阶段性成果.引导图像滤波^[24]通过考虑引导图像的结构来计算最后滤波所得到的图像.其优点是具有保护边缘的特性和独立于滤波器大小的运行时间.大量的立体匹配算法^[25-31]将其作为代价聚合方法.该方法的性能优于传统聚合方法.其中最具有代表性的是文献[6],它在时间和精度方面取得了很好的效果.此外,文献[32]提出的快速引导滤波将执行时间从 $O(n)$ 减少到 $O(n/S^2)$.文献[33]使用这种方法进行立体匹配,虽然带来了更快的聚合速度,但以牺牲精确度为代价.虽然引导滤波器能够有效地保护边缘、提高匹配精度,而且比双边滤波器快,但它可能会受到光晕伪影的影响.随后,文献[34]提出了加权引导图像滤波,文献[35]提出了梯度主导的引导图像滤波来减少这种现象.然而,这些方法作为代价聚合函数时增加了额外的时间开销.

另一方面,CTF、分层策略、多尺度这些方法逐渐成为代价聚合的主要方法.它们以多尺度为前提,通常需要图像金字塔的辅助.文献[36]提出了一种针对稀疏结构的算法,用特殊的拉普拉斯金字塔对图像的能量进行编码.文献[37]提出了一种利用具有相应视差候选集的层次结构来减少匹配比较的方法.另外,文献[38]从统一的代价优化角度规划了代价聚合,提出了一种新的通用跨尺度代价聚合框架,允许代价聚合中融合多尺度.然而,这些方法受限于从粗到精的误差传播,特别是薄壁结构的恢复较差.另外,图像金字塔的生成要花费大量的时间以及通过尺度空间传播的不可恢复的错误的视差估计.因此,为了降低计算量,避免这些算法带来的问题,本文提出了一种基于融合灰度图像的代价聚合的立体匹配框架.

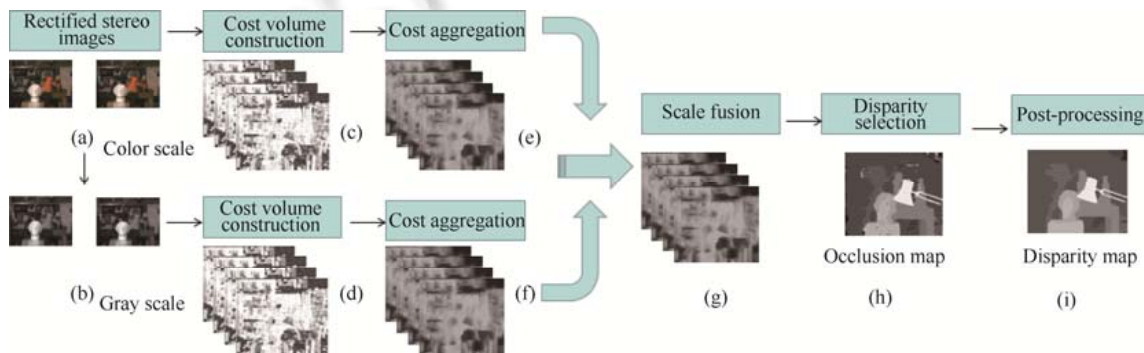
在视差选择时,几乎所有的方法都是基于 WTA 算法.该方法通过选择具有最小代价聚合值的视差来生成原始的视差图.它简单且易于实现,并且最后生成的视差是整数值.但在现实世界中,两个像素的视差不可能总是整数值.文献[39,40]中的工作对亚像素立体匹配进行了深入的讨论,结果表明这种亚像素会更加精确.

尽管代价聚合方法快速发展,但是人们对视差优化的关注却很少.许多算法使用加权中值滤波来平滑和填充区域,而权值选择的是双边滤波.文献[41]表明视差优化和代价聚合一样重要.与之前耗时的加权中值滤波相比,文献[41]是第 1 个以常数时间为基准的快速加权中值滤波.该方法释放了视差优化的潜力,不仅提高了匹配速度,而且提高了视差精度.

基于上述的讨论,本文选择了引导图像滤波和快速加权中值滤波器分别进行代价聚合和视差优化,并且快速加权中值滤波使用引导滤波核.此外,本文将灰度图像的代价聚集成到原始的代价聚合中.该方法旨在提高图像中弱纹理区域的匹配精度.同时,利用代价聚集中的最小值与次最小值之间的相互关系来确定最终的视差.

2 提出的算法

本文算法主要包括 4 个步骤:代价计算、代价聚合、视差选择和优化.算法流程如图 1 所示.



(a) 输入校正后的立体图像对;(b) 进行高斯滤波后的灰度图像;(c) 彩色尺度的代价量;(d) 灰色尺度的代价量;
(e) 彩色尺度代价聚合;(f) 灰色尺度代价聚合;(g) 进行两种尺度融合;(h) 进行视差选择;(i) 视差求精

Fig.1 Schematic flow diagram of the proposed algorithm

图 1 算法流程图

2.1 代价计算

在代价计算中,通过截断的绝对值之差(TAD)的方法,并结合梯度信息和颜色信息实现了在像素 p 处的视差 d 的代价值 $C(p,d)$.彩色尺度中的 TAD 通过式(1)来表示.

$$C_{\text{color}} = \min \left(\frac{1}{3} \sum_{i=1}^{i=3} |I_{\text{left}}^i(p) - I_{\text{right}}^i(p-d)|, T_c \right) \quad (1)$$

其中, $I_{\text{left}}^i(p)$ 和 $I_{\text{right}}^i(p-d)$ 分别是像素 p 在立体匹配图像对中的 RGB 颜色空间中的第 i 个通道的像素值. T_c 是一个颜色截断值.梯度部分的 TAD 值由式(2)给出.

$$C_{grad_x} = \min\left(\left|\nabla_x(I'_{left}(p)) - \nabla_x(I'_{right}(p-d))\right|, T_{g_x}\right) \quad (2)$$

其中, $\nabla_x(I'_{left}(p))$ 和 $\nabla_x(I'_{right}(p))$ 是像素 p 在水平方向上的梯度. 同时, T_{g_x} 是一个截断值. 最终, 联合的彩色尺度的匹配代价如式(3)所示.

$$C(p, d) = \alpha_{color} C_{color} + \alpha_{grad_x} C_{grad_x} \quad (3)$$

α_{color} 和 α_{grad_x} 是超参数, 控制像素值和梯度部分的比例.

为了提高图像中弱纹理区域的匹配精度, 本文融合了灰度图像的匹配代价. 首先对校正后的立体图像对进行灰度处理, 然后用高斯滤波平滑. 因此, 建立灰色尺度的代价量是一个类似的过程. 由于灰度图像只有 1 个通道, TAD 灰度可以表示为

$$\hat{C}_{gray} = \min\left(\left|\hat{I}_{left}(p) - \hat{I}_{right}(p-d)\right|, T_g\right) \quad (4)$$

$\hat{I}_{left}(p)$ 和 $\hat{I}_{right}(p)$ 是灰度图像在像素 p 处的灰度值. 高斯核的参数设置在 0.3, 滤波半径设置为 $5.T_g$ 同样是一个截断值. 另外, 已经计算好的 TAD 梯度值直接和式(4)进行组合. 最终灰色尺度的匹配代价如式(5)所示.

$$\hat{C}(p, d) = \alpha_{gray} \hat{C}_{gray} + \alpha_{grad_x} C_{grad_x} \quad (5)$$

为简单起见, 将式(5)的参数设置成和式(3)一样.

2.2 融合灰色尺度的代价聚合

生成的代价项含有大量的噪声, 因此必须进行聚合操作. 引导滤波具有保护图像边缘的特性并且具有线性的运行时间, 且仅仅取决于图像像素的个数. 因此, 两种尺度的聚合方法都使用了引导图像滤波.

彩色匹配代价 $C(p, d)$ 和灰度匹配代价 $\hat{C}(p, d)$ 分别由式(3)和式(5)计算得出, 代价聚合表示如下.

$$C'(p, d) = \sum_q W_{p,q} C(p, d) \quad (6)$$

$$C''(p, d) = \sum_q \hat{W}_{p,q} \hat{C}(p, d) \quad (7)$$

其中, 滤波权重 $W_{p,q}$ 和 $\hat{W}_{p,q}$ 分别依赖于彩色尺度的引导图像(原始的立体匹配图像)和灰色尺度的引导图像(灰度化后的立体匹配图像), 因为引导图像的内容影响着匹配代价. $C'(p, d)$ 和 $C''(p, d)$ 表示像素 p 在视差 d 的代价值滤波后的结果. 彩色尺度的滤波权重表示如下.

$$w(p, q) = \frac{1}{|w_k|^2} \sum_{(p,q) \in w_k} \left(1 + (I(p) - \mu_k)^T (\Sigma_k + \varepsilon U)^{-1} (I(q) - \mu_k)\right) \quad (8)$$

Σ_k 和 μ_k 分别是图像 I 中的像素 k 在区域窗口 w_k 中的协方差和均值. $|w_k|$ 表示该窗口内的总的像素值. ε 是一个平滑因子. $I(p), I(q)$ 和 μ_k 是 3×1 的向量, 协方差矩阵 Σ_k 和本征矩阵 U 是 3×3 . 然而, 引导滤波权重的巨大的优势就是计算简便并且独立于窗口的大小. 线性的计算方法如下所示.

$$a_k = (\Sigma_k + \varepsilon U)^{-1} \left(\frac{1}{|w_k|} \sum_{i \in w_k} I_i f_i - \mu_k \bar{f}_k \right) \quad (9)$$

$$b_k = \bar{f}_k - a_k^T \mu_k \quad (10)$$

$$q_i = \bar{a}_i^T I_i + \bar{b}_i \quad (11)$$

其中 f_i 是引导图像 f 中像素 i 的值, \bar{f}_k 是 f 在窗口 w_k 中的均值. q_i 是通过 \bar{a} 和 \bar{b}_i 计算出的最终结果. 此外, 所有的求和操作可以使用一个固定窗口大小的盒滤波器进行, 并且时间复杂度是 $O(N)$. 对于灰色尺度的引导滤波, 基于同样的原则, a_k 和 b_k 的计算方法如下所示.

$$a_k = \frac{\frac{1}{|w_k|} \sum_{i \in w_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon} \quad (12)$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (13)$$

μ_k 和 σ_k^2 是图像 I 在 w_k 范围内的均值和方差. \bar{p}_k 是 p 在 w_k 范围内的均值.

以上两种尺度的代价进行聚合之后,最后将它们的聚合值进行加权融合.这基于一种假设,即彩色尺度代表图像细节和边缘,而高斯滤波后的灰度图像描述了弱纹理区域.将它们的代价聚合值按一定的比例组合起来,将能更好地表示像素的匹配代价值,如图 2 所示.因此,最终代价聚合是

$$Z(p, d) = \beta C'(p, d) + (1 - \beta) C''(p, d) \quad (14)$$

β 控制着两种尺度的比例.

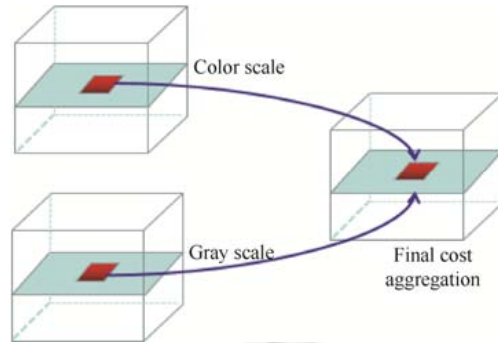


Fig.2 Fusion of gray scale cost aggregation

图 2 融合灰色尺度的代价聚合

2.3 基于置信度的视差选择

传统的 WTA 因易于实现而被广泛地用于视差选择,具体表达如式(15)所示.

$$d_p = \arg \min_{d \in D} \hat{C}(p, d) \quad (15)$$

其中, D 表示所有可能的视差集合.然而,不同的视差层可能同时具有几个相近的代价聚集值,但是 WTA 只选择其中的一个作为视差.另外,现实世界中不可能所有的视差都是整数.为了避免这种选择歧义问题,提出了一种置信度的概念,可以提高视差选择的可靠性,通过代价聚合最小值和次小值来进行计算.

$$\hat{Z}_{(p)} = \frac{\min_{d \in D} Z(p, d)}{\min_{d \in D, d \neq d_p} Z(p, d)} \quad (16)$$

当代价聚合最小值 $\min_{d \in D} Z(p, d)$ 和次小值 $\min_{d \in D, d \neq d_p} Z(p, d)$ 之间非常接近时,最后得出的比例就会很高.当该比例非常接近 1 时,不能随意地选择最小代价聚合值所在的视差层作为最终的视差值.比较简单的方法是取最小值和次小值的均值.最后的视差值计算如下.

$$\hat{d}_p = \begin{cases} \frac{d_p + d_{p'}}{2}, & \hat{Z}_{(p)} \geq T_d \\ d_p, & \hat{Z}_{(p)} < T_d \end{cases} \quad (17)$$

其中, d_p 表示代价聚合后的最小值所在的视差层, $d_{p'}$ 是次小值所指向的视差层. T_d 是一个阈值,通常设置为 0.85 或者更高,可以根据具体情况进行调整.

2.4 后处理

在经过以上处理步骤后,视差图可能包含遮挡点和大量不连续区域中的离群值.在后处理中,视差求精通常包括遮挡处理和中值滤波两个部分.

在遮挡检测中,首先按照之前的步骤再次计算从左图到右图的视差值(之前是右图到左图).然后检查两个视差图中对应的像素视差值是否相同或相似,如式(18)所示.

$$|\widehat{d}_{\text{left}}(p) - \widehat{d}_{\text{right}}(p-d)| \leq \delta \quad (18)$$

其中, $\widehat{d}_{\text{left}}(p)$ 是从左图到右图的视差; $\widehat{d}_{\text{right}}(p)$ 是右图到左图的视差. δ 的值一般设为 1. 当式(18)不成立时, 将对应的像素标记为遮挡点. 同时, 分别水平往左和往右找到第 1 个非遮挡点, 即满足该不等式, 然后选择两者之间的最小值作为填充值.

在中值滤波时, 本文使用最近提出的快速加权中值滤波^[41]来减少视差图中的条纹状伪影. 另外, 文献[41]中使用了盒滤波作为核函数. 为了更好地保护边缘和加快算法的执行时间, 本文使用引导滤波核替换原来的盒滤波, 在实验中取得了很好的效果.

3 实验结果

本文通过大量的实验来调整参数, 参数 $\{r, \varepsilon, \alpha_{\text{color}}, \alpha_{\text{grad}_x}, T_c, T_{g_x}, \alpha_{\text{gray}}, \alpha_{\text{grad}_x}\}$ 的推荐设置为(9, 0.0001, 0.1, 0.9, 0.1, 0.028, 0.1, 0.9). 对于 β 的值, 建议范围在 0.55~0.95 之间.

通过在 Middlebury 测试平台上测试本文的算法, 使用 4 幅标准测试图像集(Tsukuba、Venus、Cones、Teddy). 图 3 展示了融合灰色尺度后最后生成的视差图和没有融合灰色尺度生成的视差图的对比. 可以发现, 融合后的比没有融合生成的视差图更好. 特别是在一些弱纹理较多的图像中, 如果在代价聚合中融合了灰色尺度的代价值, 效果会更好. 这是由于高斯滤波后的灰度图像进行代价聚合, 其代价值相当于是原图像中弱纹理区域的代价聚合值. 因此, 在生成最终的视差图中考虑到纹理区域, 那么其结果将更加精确.

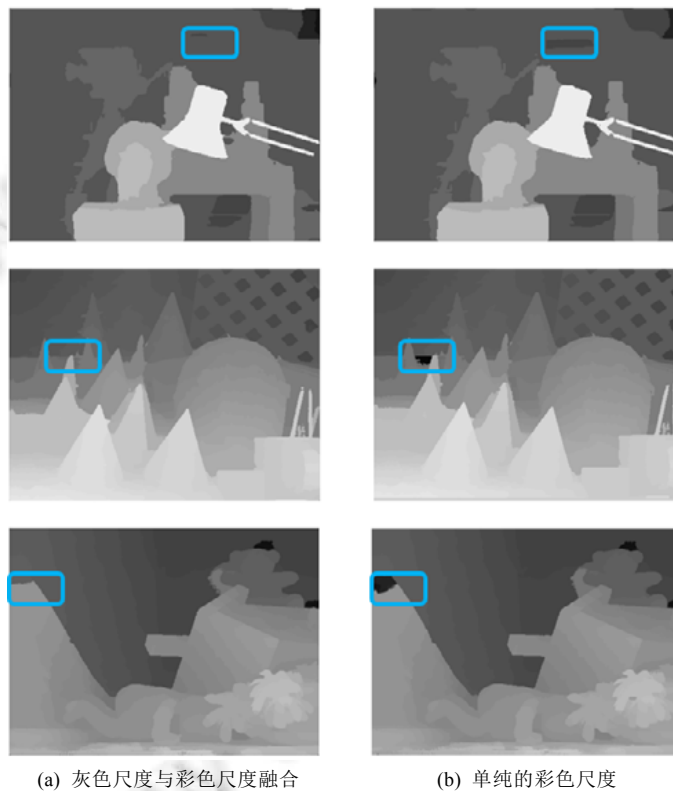


Fig.3 Fusion of gray scale and without fusion comparison
图 3 融合灰度与未融合的比较

图 4 显示了本文算法和目前排名靠前的算法 CostFilter^[6]和 P-LinearS^[26]之间的实验对比. 算法 CostFilter^[6]

和 P-LinearS^[26]都使用了引导滤波.在视差求精步骤中,文献[6]使用双边滤波,取得了很好的效果.文献[6]在 Middlebury 评价体系中是之前最好的局部立体方法.文献[26]通过一系列的后处理方法也取得了很好的效果.在图 5 中显示了相应的匹配误差图(若视差超过 1 个像素,则将作为黑色像素点在图中展示).从该图可以发现,与文献[6,26]相比,本文的算法可以获得更精确的视差图.

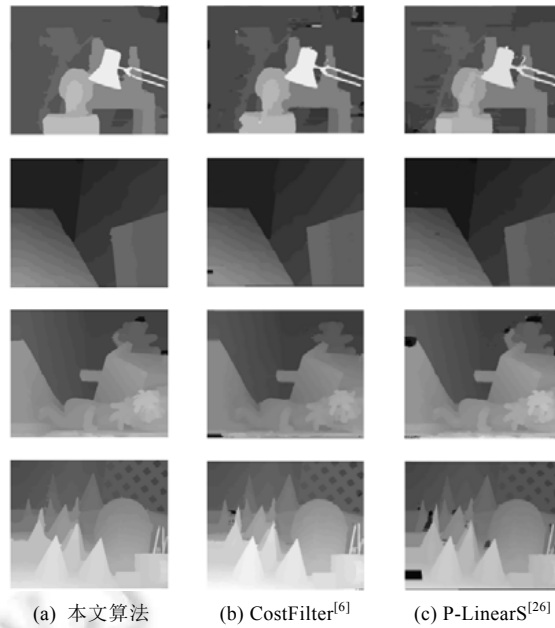


Fig.4 Experimental results of the proposed algorithm and compared with other algorithms

图 4 算法的实验结果及与其他算法的比较

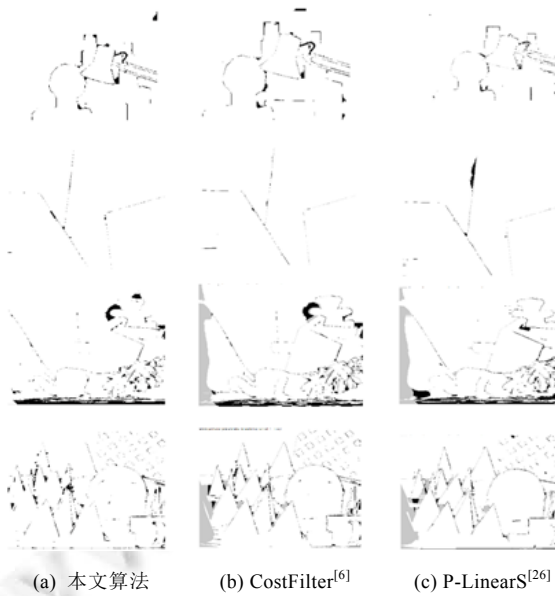


Fig.5 Mismatch map comparison

图 5 误匹配图对比

图 6 解释了 β 的作用,灰色尺度的代价聚合部分在最后的代价聚合值中随着 β 的增加而降低.正如所假定的,高斯滤波后的灰度图像代表了弱纹理.因此,如果灰色尺度的代价聚合部分的比例上升,则弱纹理中的误差匹配将减少.但同时,这会将灰色尺度的代价聚合值引入到另一个非弱纹理区域,最终会影响整个视差图的效果.实验结果表明,如果该值设置为 0.5,则错误率反而更高.所以 β 的值必须根据具体情况进行调整.

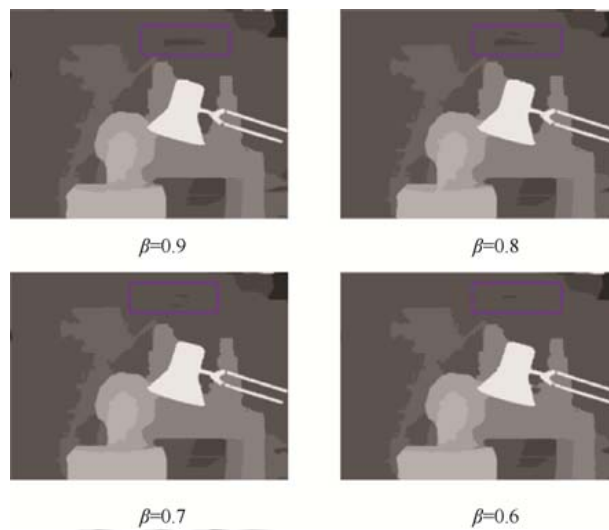


Fig.6 Effect of β

图 6 β 的作用

表 1 显示了本文提出的算法和 CostFilter^[6]以及 P-LinearS^[26]之间错误率的比较.其中,“n-occ”表示非遮挡区域的错误匹配率,“all”表示所有区域的错误匹配率,“disc”代表视差不连续区域的错误率,“Avg”表示平均错误率.4 张图片中的 β 值分别设置为 0.75,0.65,0.75,0.90.可以看出,本文提出的算法在一些非遮挡区域和某些视差不连续区域中的错误匹配率都有所降低.注意到,“disc”这一项的错误率极大地降低了.此外,即使在一些小地方我们的错误率稍微高一点,但是从整体错误率来看,我们的错误率也比其他两种算法都要低.通过这些比较可以发现,本文算法可以在图像边缘和平坦区域获得更加精确的视差图.

Table 1 Middlebury test result of the proposed algorithm and other algorithms

表 1 本文算法与其他算法的 Middlebury 测试结果

算法	Tsukuba			Vensu			Teddy			Cones			Avg
	n-occ	all	disc	n-occ	all	disc	n-occ	all	disc	n-occ	all	disc	
本文	1.35	1.57	7.21	0.18	0.26	2.41	5.67	11.68	16.02	2.37	8.12	8.17	5.417
CostFilter ^[6]	1.51	1.85	7.61	0.20	0.39	2.42	6.16	11.8	16.0	2.71	8.24	7.66	5.546
P-LinearS ^[26]	1.10	1.67	5.92	0.53	0.89	5.71	6.69	12.0	15.9	2.60	8.44	6.71	5.681

4 总 结

虽然由粗到细(CTF)、分层策略和跨尺度的代价聚合在一定程度上能够生成高精度的视差图.但是这些方法通常需要图像金字塔的辅助来构建多尺度,这增加了时间开销.因此,同样地针对弱纹理区域,本文提出了一种通用的融合灰色尺度的代价聚合框架.初始的代价聚合融合了灰度图像的代价值,这能够更好地表示图像之间的匹配代价.因为颜色尺度代表图像细节和边缘,而高斯滤波后的灰度图像描述了弱纹理区域.同时,利用代价聚合的最小值与次小值之间的相互关系来确定最终的视差,目的是避免选择 WTA 带来的歧义.最后,引入引导图像滤波和快速加权中值滤波方法进行代价聚合和视差求精.引导滤波器是一个常数执行时间,有效地保护了图像边缘.实验结果表明,该算法质量高,耗时短.

References:

- [1] Zhan Y, Gu Y, Huang K, *et al.* Accurate image-guided stereo matching with efficient matching cost and disparity refinement. *IEEE Trans. on Circuits & Systems for Video Technology*, 2016,26(9):1632–1645.
- [2] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int'l Journal of Computer Vision*, 2002,47(1-3):7–42.
- [3] Mozerov MG, Weijer JV. Accurate stereo matching by two-step energy minimization. *IEEE Trans. on Image Processing*, 2015, 24(3):1153–1163.
- [4] Klaus A, Sormann M, Karner K. Segment-Based stereo matching using belief propagation and a self-adapting dissimilarity measure. In: *Proc. of the Int'l Conf. on Pattern Recognition*. IEEE Computer Society, 2006. 15–18.
- [5] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*, Vol.2. 2001. 508–515.
- [6] Hosni A, Rhemann C, Bleyer MC, *et al.* Fast cost-volume filtering for visual correspondence and beyond. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2011. 3017–3024.
- [7] Baek ET, Ho YS. Cost aggregation with guided image filter and superpixel for stereo matching. In: *Proc. of the Signal and Information Processing Association Summit and Conf.* 2017. 1–4.
- [8] Hirschmuller H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, 2007,30(2):328–341.
- [9] Hirschmuller H. Accurate and efficient stereo processing by semi-global matching and mutual information. In: *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. IEEE, 2005. 807–814.
- [10] Fife WS, Archibald JK. Improved census transforms for resource-optimized stereo vision. *IEEE Trans. on Circuits & Systems for Video Technology*, 2013,23(1):60–73.
- [11] Sarkar I, Bansal M. A wavelet-based multiresolution approach to solve the stereo correspondence problem using mutual information. *IEEE Trans. on Systems Man & Cybernetics Part B Cybernetics*, 2007,37(4):1009–1014.
- [12] Nguyen VD, Nguyen DD, Sang JL, *et al.* Local density encoding for robust stereo matching. *IEEE Trans. on Circuits & Systems for Video Technology*, 2014,24(12):2049–2062.
- [13] Hermann S, Vaudrey T. The gradient—A powerful and robust cost function for stereo matching. In: *Proc. of the Image and Vision Computing*. IEEE, 2012. 1–8.
- [14] Liu H, Liu Y, Ouyang S, *et al.* A novel method for stereo matching using Gabor feature image and confidence mask. In: *Proc. of the Visual Communications and Image Processing*. IEEE, 2013. 1–6.
- [15] Fife WS, Archibald JK. Improved census transforms for resource-optimized stereo vision. *IEEE Trans. on Circuits & Systems for Video Technology*, 2013,23(1):60–73.
- [16] Jiao J, Wang R, Wang W, *et al.* Local stereo matching with improved matching cost and disparity refinement. *IEEE Multimedia*, 2014,21(4):16–27.
- [17] Cigla C, Alatan AA. Information permeability for stereo matching. *Signal Processing Image Communication*, 2013,28(9):1072–1088.
- [18] Klaus A, Sormann M, Karner K. Segment-Based stereo matching using belief propagation and a self-adapting dissimilarity measure. In: *Proc. of the Int'l Conf. on Pattern Recognition*. IEEE Computer Society, 2006. 15–18.
- [19] Yoon KJ, Kweon IS. Adaptive support-weight approach for correspondence search. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2006,28(4):650–656.
- [20] Mattoccia S, Giardino S, Gambini A. Accurate and efficient cost aggregation strategy for stereo correspondence based on approximated joint bilateral filtering. In: Zha HB, Taniguchi RI, Maybank S, eds. *Proc. of the Asian Conf. on Computer Vision*. LNCS 5995, Berlin, Heidelberg: Springer-Verlag, 2010. 371–380.
- [21] Yang Q. A non-local cost aggregation method for stereo matching. In: *Proc. of the Computer Vision and Pattern Recognition*. IEEE, 2012. 1402–1409.
- [22] Mei X, Sun X, Dong WM, *et al.* Segment-Tree based cost aggregation for stereo matching. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2013. 313–320.

- [23] Yu W, Chen T, Hoe JC. Real time stereo vision using exponential step cost aggregation on GPU. In: Proc. of the IEEE Int'l Conf. on Image Processing. IEEE, 2010. 4225–4228.
- [24] He K, Sun J, Tang X. Guided image filtering. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013,35(6):1397–1409.
- [25] Kordelas GA, Alexiadis DS, Daras P, *et al.* Content-Based guided image filtering, weighted semi-global optimization, and efficient disparity refinement for fast and accurate disparity estimation. IEEE Trans. on Multimedia, 2016,18(2):155–170.
- [26] De-Maeztu L, Mattoccia S, Villanueva A, *et al.* Linear stereo matching. In: Proc. of the Int'l Conf. on Computer Vision. IEEE, 2011. 1708–1715.
- [27] Hosni A, Bleyer M, Rhemann C, *et al.* Real-Time local stereo matching using guided image filtering. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo. IEEE, 2011. 1–6.
- [28] Ttofis C, Kyrkou C, Theoharides T. A low cost real-time embedded stereo vision system for accurate disparity estimation based on guided image filtering. IEEE Trans. on Computers, 2016,65(9):2678–2693.
- [29] Baek ET, Ho YS. Cost aggregation with guided image filter and superpixel for stereo matching. Signal and Information Processing Association Summit and Conf. IEEE, 2017. 1–4.
- [30] Hong GS, Koo MS, Saha A, *et al.* Efficient local stereo matching technique using weighted guided image filtering (WGIF). In: Proc. of the IEEE Int'l Conf. on Consumer Electronics. IEEE, 2016.
- [31] Çiğla C, Alatan AA. Efficient edge-preserving stereo matching. In: Proc. of the IEEE Int'l Conf. on Computer Vision Workshops. IEEE Computer Society, 2011. 696–699.
- [32] He KM, Sun J. Fast guided filter. <http://arXiv:1505.00996>
- [33] Hong GS, Park JK, Kim BG. Near real-time local stereo matching algorithm based on fast guided image filtering. In: Proc. of the European Workshop on Visual Information Processing. IEEE, 2016. 1–5.
- [34] Li ZG, Zheng JH, Zhu ZJ, *et al.* Weighted guided image filtering. IEEE Trans. on Image Processing, 2014,24(1):120–129.
- [35] Kou F, Chen WH, Wen CY, *et al.* Gradient domain guided image filtering. IEEE Trans. on Image Processing, 2015,24(11):4528–4539.
- [36] Sizintsev M. Hierarchical stereo with thin structures and transparency. In: Proc. of the Canadian Conf. on Computer and Robot Vision. IEEE Computer Society, 2008. 97–104.
- [37] Hu W, Zhang K, Sun L, *et al.* Comparisons reducing for local stereo matching using hierarchical structure. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo. IEEE, 2013. 1–6.
- [38] Zhang K, Fang Y, Min D, *et al.* Cross-Scale cost aggregation for stereo matching. In: Proc. of the Computer Vision and Pattern Recognition. 2014. 1590–1597.
- [39] Gehrig SK, Badino H, Franke U. Improving sub-pixel accuracy for long range stereo. Computer Vision & Image Understanding, 2007,116(1):16–24.
- [40] Nehab D, Rusinkiewicz S, Davis J. Improved sub-pixel stereo correspondences through symmetric refinement. In: Proc. of the 10th IEEE Int'l Conf. on Computer Vision. IEEE Computer Society, 2005. 557–563.
- [41] Ma Z, He K, Wei Y, *et al.* Constant time weighted median filtering for stereo matching and beyond. In: Proc. of the IEEE Int'l Conf. on Computer Vision. IEEE, 2014. 49–56.



韩先君(1988—),男,四川成都人,博士生,工程师,CCF 学生会员,主要研究领域为计算机视觉,三维重建,数字图像处理,计算机图形学.



杨红雨(1967—),女,博士,教授,博士生导师,主要研究领域为计算机视觉,视觉合成,数字图像处理与计算机图形学,实时软件工程,空管信息智能处理技术,视觉导航,飞行模拟技术.