

基于引导图像和自适应支持域的立体匹配

孔令寅, 朱江平, 应三丛*

四川大学计算机学院, 四川 成都 610065

摘要 提出一种基于引导图像和自适应支持域的局部立体匹配算法。首先对校正后的输入图像进行预处理得到引导图像;在匹配代价计算阶段,提出一种梯度计算方法,结合引导图像和输入图像的梯度信息,分别计算 x 和 y 方向的梯度,再与 AD(absolute difference)和 Census 变换融合构建匹配代价计算函数;在代价聚合阶段,使用基于自适应支持域的导向滤波;在视差细化阶段,提出一套基于自适应支持域的多步细化方法,通过该方法得到最终的视差图。实验结果表明,视差细化后全部区域的平均误差和方均根误差平均减少 43.7% 和 38%,非遮挡区域平均减少 33.7% 和 30.9%,所提算法具有较好的鲁棒性并能获得精度较高的视差结果。

关键词 机器视觉;局部立体匹配算法;引导图像;自适应支持域;导向滤波

中图分类号 TP301.6

文献标志码 A

doi: 10.3788/AOS202040.0915001

Stereo Matching Based on Guidance Image and Adaptive Support Region

Kong Lingyin, Zhu Jiangping, Ying Sancong*

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

Abstract In this study, we propose a local stereo matching algorithm based on guidance images and an adaptive support region. First, the guidance images can be obtained by preprocessing the rectified input images. During the matching cost calculation stage, we propose a gradient calculation method, which combines the gradient information of the guidance and input images to calculate the gradients along the x and y directions, respectively, and subsequently integrates the absolute difference (AD) and the Census transform to develop a matching cost calculation function. Further, we use a guided filter based on the adaptive support region during the cost aggregation stage. During the disparity refinement stage, a multi-step refinement method is proposed based on the adaptive support region and then the final disparity map is obtained. The experimental results prove that after disparity refinement, the average error (Avgerr) and root-mean-square error (RMSE) are reduced by 43.7% and 38% respectively for all the regions and by 33.7% and 30.9% for the non-occluded regions. The proposed algorithm exhibits improved robustness and can be used to obtain high precision disparity results.

Key words machine vision; local stereo matching algorithm; guidance image; adaptive support region; guided filter

OCIS codes 150.1135; 330.1400; 100.6890

1 引 言

双目立体匹配是一个利用同一场景的左、右两幅视点图像,基于双目视觉原理,运用立体匹配算法找出每个像素的匹配点,从而得到视差信息的过程,在三维测量^[1]、三维重建^[2]、机器人导航^[3]、无人驾驶^[4]等领域得到广泛应用,一直是计算机视觉中的研究热点。近几年基于神经网络的立体匹配算法发展迅猛,其中基于卷积神经网络(CNN)的立体匹配算法是目前研究热点之一。Žbontar 等^[5]提出一

种利用孪生深度网络计算匹配代价的算法,该算法使用一对图像块,通过训练让网络找出图像块之间的相似点。Pang 等^[6]提出级联残差学习(CRL)网络,该网络分为两段,每段独立计算视差图和多尺度残差。Chang 等^[7]提出金字塔立体匹配网络(PSMNet),该网络由空间金字塔池化模块和三维 CNN 模块组成。这类方法虽然匹配精度高、效果好,但训练时需要大量的数据,因此在实际应用中受到限制。

传统的立体匹配算法无需大量数据进行训练,

收稿日期: 2019-11-25; 修回日期: 2019-12-24; 录用日期: 2020-01-17

基金项目: 国家自然科学基金青年基金(61901287)、四川省重大科技专项(2018GZDZX0024)

* E-mail: yingsancong@scu.edu.cn

匹配准确率也能满足实际应用需要,所以应用更为广泛。传统算法通常包括全局算法、半全局算法、局部算法三类。全局算法一般通过求取全局能量函数的最小值来获得视差,常见的有图割算法^[8]、置信传播算法^[9]等,其特点是匹配准确率高,但计算效率低。半全局算法由 Hirschmüller^[10]提出,主要利用动态规划的思想,匹配准确率和计算效率介于全局算法和局部算法之间。局部算法在指定的支持域内通过代价聚合的方式进行匹配,匹配准确率一般低于前两种算法,但计算效率更高,并且更容易实现。局部算法一般分为 4 个步骤^[11]:匹配代价计算、代价聚合、视差计算和视差细化。

匹配代价计算是对左、右图像中每个像素的相似度进行量化,得到与视差范围对应的匹配代价体。计算方法主要分为 3 类,即基于像素的方法、基于窗口的方法、非参数方法。基于像素的方法有绝对差(AD)、平方差(SD)、梯度等;基于窗口的方法有 SAD(绝对差之和)、平方差之和(SSD)、归一化互相关(NCC)等;非参数方法有 Census 变换、Rank 变换、互信息等。每种方法都有不足,因而目前普遍的做法是将多种方法相融合以构造匹配代价计算函数^[12-14]。

代价聚合是在每个像素的支持域内对匹配代价求和或取平均值。目前常见的方法主要分为 3 类,即基于支持域的方法、基于滤波的方法、基于最小生成树(MST)的方法。基于支持域的方法以自适应支持域为主。其中 Zhang 等^[15]提出了基于十字交叉的自适应支持域构建方法。在此基础上,Mei 等^[12]对支持臂扩展的判断标准进行改进,提出了颜色和距离的双重阈值。闫利等^[16]提出一种双约束线性变化阈值的判定准则。基于滤波的方法主要采用导向滤波^[17]。如 Zhu 等^[13]采用基于自适应矩形支持域的导向滤波进行代价聚合,取得了不错的效果。闫利等^[16]将基于十字交叉的自适应支持域与导向滤波相结合,在提高匹配精度的同时,增强了鲁棒性。基于 MST 的方法由 Yang^[18]提出,采用基于全局的 MST 进行代价聚合。Li 等^[19]将 MST 支持域滤波与 PatchMatch(区域匹配)三维标签结合,算法获得了较高的匹配准确率。

视差计算是在代价聚合的基础上根据某种计算方法得到每个像素的视差。目前主要采用 WTA(赢者通吃)策略^[11],该策略选择最小的聚合代价值所对应的视差。

视差细化是对视差计算后生成的视差图中存在

的异常点进行检测和修正,从而得到更为准确的视差结果。左右一致性检测是目前广泛使用的异常点检测方法,区域投票^[12]、代价谱峰值分析与去除^[20]、亚像素求精^[12,21]、加权中值滤波^[17,22]等都是常用的异常点修正方法。每种方法都有局限性,因此将几种方法组合形成一套多步细化方法是目前比较常见的方式^[12,14,23-24]。

在现有局部立体匹配算法的基础上,本文提出一种基于引导图像和自适应支持域的立体匹配算法。在匹配代价计算阶段,提出一种梯度计算方法,并与 AD 和 Census 变换融合得到匹配代价计算函数;在代价聚合阶段,使用基于自适应支持域的导向滤波;在视差细化阶段,提出一套基于自适应支持域的多步细化方法。

2 算法描述

所提算法主要包含 5 个步骤:以校正后的图像作为输入,对图像进行预处理得到引导图像;结合引导图像和输入图像的梯度信息,应用所提方法分别计算 x 和 y 方向的梯度,并与 AD 和 Census 变换融合构建计算函数,计算匹配代价体;建立自适应支持域(ASR),基于该支持域对匹配代价体进行导向滤波(GIF);用 WTA 进行视差计算,获得初始视差图;使用提出的多步细化方法进行视差细化,输出最终视差结果。算法流程如图 1 所示。

2.1 图像预处理

为得到准确的引导图像,首先对输入图像进行预处理。导向滤波具有边缘保持特性,平滑效果好且计算时间与支持域大小无关,因此采用导向滤波得到引导图像。

2.2 匹配代价计算

为增强算法的鲁棒性,采用 AD、Census 变换及 x 和 y 方向梯度相融合的方法进行匹配代价计算。

AD 使用图像 RGB 三个通道的信息计算,表达式为

$$C_{AD}(p, d) = \frac{1}{3} \sum_{i=R,G,B} |I_{Li}(p) - I_{Ri}(q)|, \quad (1)$$

式中: $I_{Li}(p)$ 为左图 i 通道上点 p 的值; $I_{Ri}(q)$ 为右图 i 通道上视差等于 d 时对应点 q 的值。

Census 变换通过比较中心像素与窗口内其他像素的大小,得到对应的二进制编码,然后计算两个编码的 Hamming 距离:

$$C_{Cen}(p, d) = \text{cen}(p) \oplus \text{cen}(q), \quad (2)$$

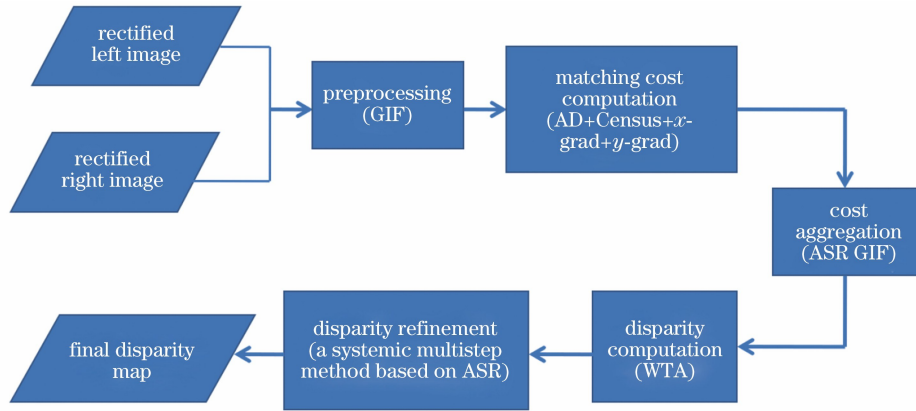


图 1 本文算法流程图

Fig. 1 Flowchart of proposed algorithm

式中： $\text{cen}(p)$ 和 $\text{cen}(q)$ 分别为左图中任意一点 p 和对对应点 q 的二进制编码。

图像梯度包含大量的结构信息，所以在匹配代价计算中广泛使用^[13-14, 25-26]。由于输入图像包含更丰富的局部细节信息，而引导图像主要包含边缘信

息，将两者梯度相结合，既包含了图像的局部特征，又包含了边缘特征^[14]，因而采用输入图像和引导图像相结合的方法。在此基础上，提出一种梯度计算方法分别计算 x 和 y 两个方向的梯度，表达式分别为

$$Cg_x(p, d) = \frac{1}{3} \sum_{i=R,G,B} [|g_{x_i}^{(IL)}(p) - g_{x_i}^{(IR)}(q)| + |g_{x_i}^{(GIL)}(p) - g_{x_i}^{(GIR)}(q)|], \quad (3)$$

$$Cg_y(p, d) = \frac{1}{3} \sum_{i=R,G,B} [|g_{y_i}^{(IL)}(p) - g_{y_i}^{(IR)}(q)| + |g_{y_i}^{(GIL)}(p) - g_{y_i}^{(GIR)}(q)|], \quad (4)$$

式中：IL和IR分别为输入图像对的左图和右图；GIL和GIR分别为引导图像对的左图和右图； g_{x_i} 和 g_{y_i} 分别为 i 通道上 x 和 y 方向的梯度。

将以上计算方法加权融合，得到匹配代价计算函数为

$$C(p, d) = 4 - \exp\left[-\frac{C_{AD}(p, d)}{\lambda_{AD}}\right] - \exp\left[-\frac{C_{Cen}(p, d)}{\lambda_{Cen}}\right] - \exp\left[-\frac{C_{g_x}(p, d)}{\lambda_{g_x}}\right] - \exp\left[-\frac{C_{g_y}(p, d)}{\lambda_{g_y}}\right], \quad (5)$$

式中： λ_{AD} 、 λ_{Cen} 、 λ_{g_x} 、 λ_{g_y} 分别为对应方法的权重； $C_{AD}(p, d)$ 、 $C_{Cen}(p, d)$ 、 $C_{g_x}(p, d)$ 、 $C_{g_y}(p, d)$ 分别为对应方法的匹配代价。

2.3 代价聚合

2.3.1 自适应支持域的构建

由于一个支持域内颜色相近的点可能来自相同的图像结构，因此这些点具有相似的视差^[15]。为确保支持域内只包含颜色相近的点，采用十字交叉法^[15]构建自适应支持域，臂长的判定准则采用距离和颜色双阈值的方法^[12]，具体规则如下。

1) $D_c(p, p_e) < C_1$ 且 $D_c(p_e, p_n) < C_1$ ；

2) $D_d(p, p_e) < L_1$ ；

3) 当 $L_2 < D_d(p, p_e) < L_1$ 时， $D_c(p, p_e) < C_2$ 。

其中： p_e 为点 p 支持臂上的一点； p_n 为 p_e 在支持臂方向上的前一个点； $D_c(p, p_e) = \max_{i=R,G,B} |I_i(p) - I_i(p_e)|$ ； $D_d(p, p_e) = |p - p_e|$ ； C_1 、 C_2 、 L_1 、 L_2 均为阈值。

按照以上规则，以点 p 为中心，分别向上、下、左、右4个方向扩展，当不满足规则时停止。点 p 的十字交叉区域由垂直方向支持臂 $V(p)$ 和水平方向支持臂 $H(p)$ 组成，自适应支持域 $R(p)$ 可表示为 $V(p)$ 上所有点 q 水平方向支持臂 $H(q)$ 的并集，如图2所示。

2.3.2 基于自适应支持域的导向滤波

与传统的导向滤波相比，基于自适应支持域的导向滤波取得了更好的效果^[13, 16]，因此采用该方法进行代价聚合。为提高计算效率，利用正交积分图技术^[15]对自适应支持域进行求和。由于RGB图像对边缘的保护作用更加明显^[17]，因此直接使用输入图像作为导向图。设导向图为 I ，待滤波图像为 m ，

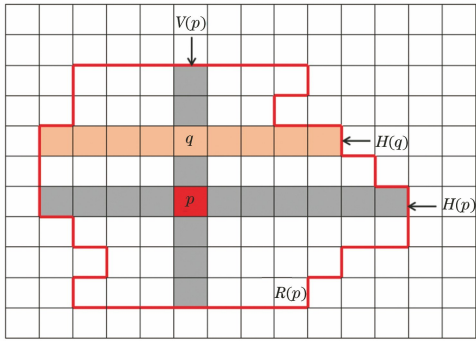


图 2 十字交叉法构建自适应支持域示意图

Fig. 2 Schematic of constructing adaptive support region based on cross method

滤波后的结果为 \mathbf{n} , 基于自适应支持域的导向滤波的线性系数 \mathbf{a}_k 和 \mathbf{b}_k 的计算公式分别为

$$\mathbf{a}_k = (\boldsymbol{\Sigma}_k + \epsilon \mathbf{U})^{-1} \left(\frac{1}{|\mathbf{R}_k|} \sum_{j \in \mathbf{R}_k} \mathbf{I}_j \mathbf{m}_j - \boldsymbol{\mu}_k \bar{\mathbf{m}}_k \right), \quad (6)$$

$$\mathbf{b}_k = \bar{\mathbf{m}}_k - \mathbf{a}_k^T \boldsymbol{\mu}_k, \quad (7)$$

式中: ϵ 为正则化参数; \mathbf{U} 为 3×3 的单位矩阵; $|\mathbf{R}_k|$ 为自适应支持域 \mathbf{R}_k 内点的数量; $\boldsymbol{\mu}_k$ 和 $\bar{\mathbf{m}}_k$ 分别为导向图像 \mathbf{I} 和待滤波图像 \mathbf{m} 在 \mathbf{R}_k 中的均值; $\boldsymbol{\Sigma}_k$ 为导向图像 \mathbf{I} 在 \mathbf{R}_k 中的 3×3 协方差矩阵; k 为图像索引。

最后, 利用线性模型计算滤波后的结果, 表达式为

$$\mathbf{n}_j = \bar{\mathbf{a}}_j^T \mathbf{I}_j + \bar{\mathbf{b}}_j, \quad (8)$$

式中: $\bar{\mathbf{a}}_j$ 和 $\bar{\mathbf{b}}_j$ 分别为 \mathbf{a}_k 和 \mathbf{b}_k 在 \mathbf{R}_k 中的均值。

2.4 视差计算

采用 WTA 方法, 选取最小匹配代价所对应的视差作为初始视差, 表达式为

$$d_{\text{ini}}(p) = \underset{0 \leq d \leq d_{\text{max}}}{\operatorname{argmin}} [C_{\text{CA}}(p, d)], \quad (9)$$

式中: $C_{\text{CA}}(p, d)$ 为经过代价聚合后的匹配代价值; $d_{\text{ini}}(p)$ 为点 p 的初始视差。

2.5 视差细化

初始视差图中一般都包含很多异常点, 需通过视差细化进行检测和修正, 因此视差细化和代价聚合同等重要^[22]。

针对这一步骤, 对现有方法进行重新组合, 提出一套基于自适应支持域的多步细化方法。该方法包括左右一致性检测和异常点分类、自适应支持域投票、自适应支持域四向传播插值、两向传播插值、无对应异常点插值、亚像素求精等环节。

2.5.1 左右一致性检测和异常点分类

左右一致性检测通过左图的视差找到右图对应的匹配点, 并判断两点视差是否相等, 如果相等则为

可靠点, 否则为异常点。判断方法为

$$|d_{\text{L}}(x_0, y_0) - d_{\text{R}}[x_0 - d_{\text{L}}(x_0, y_0), y_0]| \leq 1, \quad (10)$$

式中: d_{L} 和 d_{R} 分别为左图和右图的初始视差。

将检测到的异常点分为两类, 一类在右图中有对应匹配点; 另一类在右图中不存在对应点。将第一类称为有对应异常点, 第二类称为无对应异常点。后续步骤会对两类异常点分别进行修正。

2.5.2 自适应支持域投票

对所有异常点进行自适应支持域内投票, 统计可靠点的总票数和最高票数, 再进行判断。判断的表达式为

$$N_{\text{T}} > N, \quad (11)$$

$$\frac{N_{\text{max}}}{N_{\text{T}}} > P, \quad (12)$$

式中: N_{T} 为总票数; N_{max} 为最高票数; N 和 P 均为阈值。若同时满足, 用票数最高点的视差替换异常点视差, 并将其标记为可靠点。为了处理尽可能多的异常点, 该过程需迭代 5 次。

2.5.3 自适应支持域四向传播插值

针对有对应异常点, 在自适应支持域内沿着该点四个支持臂的方向分别找到距离最近的可靠点, 对应视差分别记为 $d_{\text{rl}}, d_{\text{rr}}, d_{\text{ru}}, d_{\text{rd}}$ 。如果 d_{rl} 和 d_{rr} 均存在, 则用 d_{hmin} 替换异常点的视差; 如果 d_{ru} 和 d_{rd} 均存在, 则用 d_{vmin} 替换异常点的视差; 如果 $|d_{\text{hmin}} - d_{\text{vmin}}| \leq 2$, 则用 $(d_{\text{hmin}} + d_{\text{vmin}})/2$ 替换异常点的视差; 否则, 不进行替换。其中, $d_{\text{hmin}} = \min(d_{\text{rl}}, d_{\text{rr}})$, $d_{\text{vmin}} = \min(d_{\text{ru}}, d_{\text{rd}})$ 。为了处理尽可能多的异常点, 该过程需迭代 3 次。采用该方法进行替换, 并标记为可靠点。

2.5.4 两向传播插值

对于剩余的有对应异常点, 沿着点的左右两边分别找到距离最近的可靠点, 对应视差记为 d_{rl} 和 d_{rr} 。如果 d_{rl} 和 d_{rr} 均存在, 则用 d_{hmin} 替换异常点的视差; 否则, 不进行替换。采用该方法进行替换, 并标记为可靠点。

2.5.5 无对应异常点插值

经过上述步骤, 剩余的异常点主要为无对应异常点, 由于该类异常点一般出现在图中的最左边区域, 所以采用单向传播插值方式, 即在点的右边找到距离最近的可靠点, 用其视差进行替换。

2.5.6 亚像素求精

为减小视差不连续带来的误差, 使用基于二次多项式插值的方法进行亚像素求精, 即

$$d_{\text{sub}} = d_n - \frac{C_{\text{CA}}(p, d_n + 1) - C_{\text{CA}}(p, d_n - 1)}{2[C_{\text{CA}}(p, d_n + 1) + C_{\text{CA}}(p, d_n - 1) - 2C_{\text{CA}}(p, d_n)]} \quad (13)$$

式中： d_n 为经过之前步骤后点 p 的视差； $C_{\text{CA}}(p, d_n + 1)$ 和 $C_{\text{CA}}(p, d_n - 1)$ 分别为 p 点在视差为 $d_n + 1$ 和 $d_n - 1$ 时经过代价聚合后的匹配代价。最后使用 3×3 中值滤波进行平滑处理，得到最终视差结果。

3 实验结果与分析

采用 Middleberry V3 测试平台 training set 和 test set 中的图像进行实验。实验参数为 $\lambda_{\text{AD}} = 30/255$, $\lambda_{\text{Cen}} = 45/255$, $\lambda_{g_x} = 5/255$, $\lambda_{g_y} = 15/255$, $\epsilon = 0.01^2$, $C_1 = 15/255$, $C_2 = 12/255$, $L_1 = \max(h, w)/20$, $L_2 = \max(h, w)/40$, $N = 50$, $P = 0.5$, 其中 h 和 w 分别表示图像的高和宽。

3.1 梯度计算方法验证

为验证所提梯度计算方法的效果，将所提梯度计算方法与另外两种梯度计算方法进行对比。第一种（简称 C1）为 x 方向梯度（仅包含输入图像）+ y 方向梯度（仅包含输入图像），计算时采用不同权重；第二种（简称 C2）为 x 方向梯度（仅包含输入图

像）+ y 方向梯度（仅包含输入图像），计算时采用相同权重。三种方法均与 AD 和 Census 变换融合，得到三种代价计算方法，它们采用相同的代价聚合策略，且均不进行视差细化处理，得到 Adirondack 图像的视差图，如图 3 所示。

从图 3 可以看出，所提梯度计算方法的视差图中误匹配区域面积（方框）更小，整体效果明显优于其他两种方法。

为了更加全面地进行对比，按照上述方式计算 training set 中所有图像的视差图，由 Middleberry V3 测试平台得到平均误差 (Avgerr) 和方均根误差 (RMSE) 的加权平均值，结果如图 4 所示。

图 4 的结果表明，在图像的全部 (all) 区域和非遮挡 (nonocc) 区域，所提梯度计算方法的 Avgerr 和 RMSE 的加权平均值均小于其他两种方法，整体匹配精度更高。

3.2 视差细化效果验证

为验证所提多步细化方法的有效性，对 3.1 中得到的 training set 所有图像的视差图均采用所提多

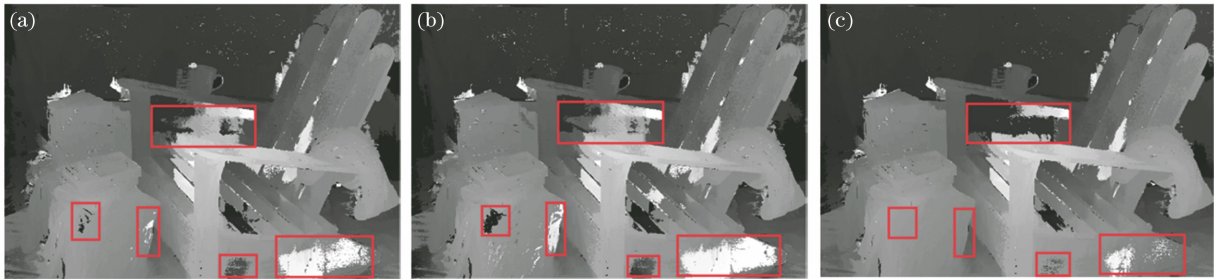


图 3 三种代价计算方法的视差图。(a) C1; (b) C2; (c) 所提梯度计算方法

Fig. 3 Disparity maps of three cost computation methods. (a) C1; (b) C2; (c) proposed gradient calculation method

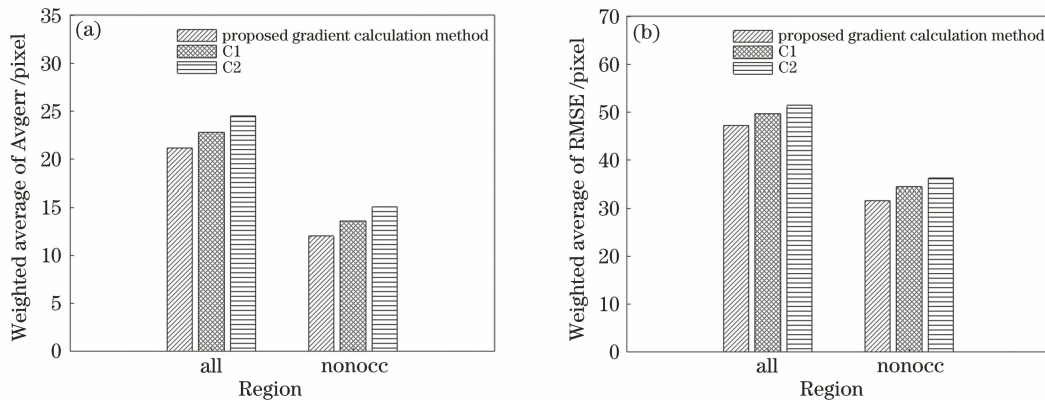


图 4 全部区域和非遮挡区域的加权平均值。(a) Avgerr; (b) RMSE

Fig. 4 Weighted averages for all regions and non-occluded regions. (a) Avgerr; (b) RMSE

步细化方法进行视差细化,每个环节执行后,全部区域和非遮挡区域的 Avgerr 和 RMSE 的加权平均值

变化情况如图 5 所示。

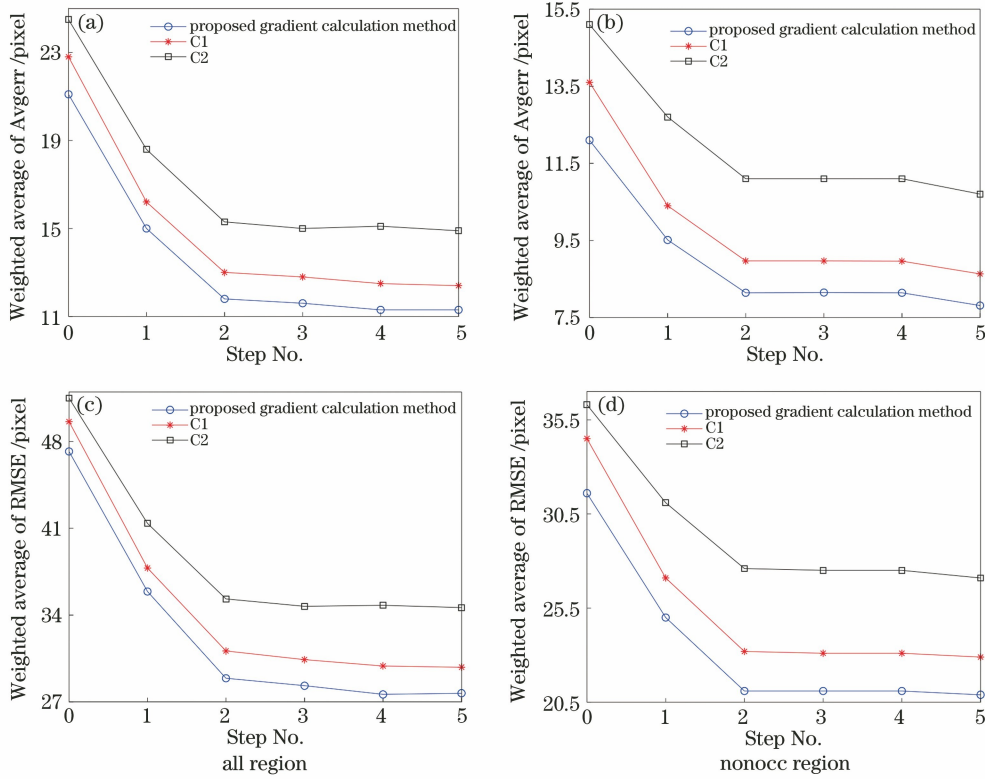


图 5 每个环节进行视差细化后的加权平均值。(a)(b) Avgerr; (c)(d) RMSE

Fig. 5 Weighted average after disparity refinement on each step. (a)(b) Avgerr; (c)(d) RMSE

由图 5 可知,对于三种代价计算方法得到的视差图,所提多步细化方法均能有效减少误差。其中,自适应支持域投票(step 1)和自适应支持域四向传播插值(step 2)起到了主要作用;两向传播插值(step 3)和无对应异常点插值(step 4)在降低全部区域的误差方面有一定的效果,尤其是对于 C1 和所提梯度计算方法的视差图;亚像素求精(step 5)在减少非遮挡区域的平均误差方面有一定的作用。

为了更直观展示所提多步细化方法的效果,表 1~3 分别列出了所提多步细化方法对三种代价计算方法的视差图进行视差细化前后的误差的加权平均值及降低的百分比。其中,在全部区域,Avgerr 降低百分比的平均值为 43.7%,RMSE 降低百分比的平均值为 38%;在非遮挡区域,Avgerr 降低百分比的平均值为 33.7%,RMSE 降低百分比的平均值为 30.9%。说明所提多步细化方法能有效修正视差图中的异常点,使视差结果更加准确。

3.3 算法对比分析

为进一步验证所提算法的鲁棒性和综合性能,选择其他 6 种算法进行对比实验,包括 LE-ELAS^[27]、IEBIMst^[28]、SPS^[29]、SM-AWP^[30]、

DSGCA^[31]、DoGGuided^[32],部分图像的实验结果如图 6 所示。

表 1 视差细化前后误差的加权平均值和降低的百分比 (所提梯度计算方法的视差图)

Table 1 Weighted average of errors before and after disparity refinement and reduced percentage (disparity map of proposed gradient calculation method)

Parameter	Weighted average /pixel		Reduced percentage / %
	Before disparity refinement	After disparity refinement	
Avgerr(all)	21.1	11.30	46.4
Avgerr(nonocc)	12.1	7.81	35.5
RMSE(all)	47.2	27.70	41.3
RMSE(nonocc)	31.6	20.90	33.9

从图 6 可以看出,所提算法的边缘保持效果更好,噪声更小,物体表面也更加平滑。

为了进行定量对比,将 training set 和 test set 所有图像在全部区域和非遮挡区域中的 Avgerr 和 RMSE 进行比较,结果如表 4~7 所示,其中加粗数据为最低值。

表 2 视差细化前后误差的加权平均值和降低的百分比
(C1 的视差图)

Table 2 Weighted average of errors before and after disparity refinement and reduced percentage (disparity map of C1)

Parameter	Weighted average /pixel		Reduced percentage / %
	Before disparity refinement	After disparity refinement	
Avgerr(all)	22.8	12.40	45.6
Avgerr(nonocc)	13.6	8.63	36.5
RMSE(all)	49.6	29.80	39.9
RMSE(nonocc)	34.5	22.90	33.6

表 3 视差细化前后误差的加权平均值和降低的百分比
(C2 的视差图)

Table 3 Weighted average of errors before and after disparity refinement and reduced percentage (disparity map of C2)

Parameter	Weighted average /pixel		Reduced percentage / %
	Before disparity refinement	After disparity refinement	
Avgerr(all)	24.5	14.9	39.2
Avgerr(nonocc)	15.1	10.7	29.1
RMSE(all)	51.5	34.6	32.8
RMSE(nonocc)	36.3	27.1	25.3

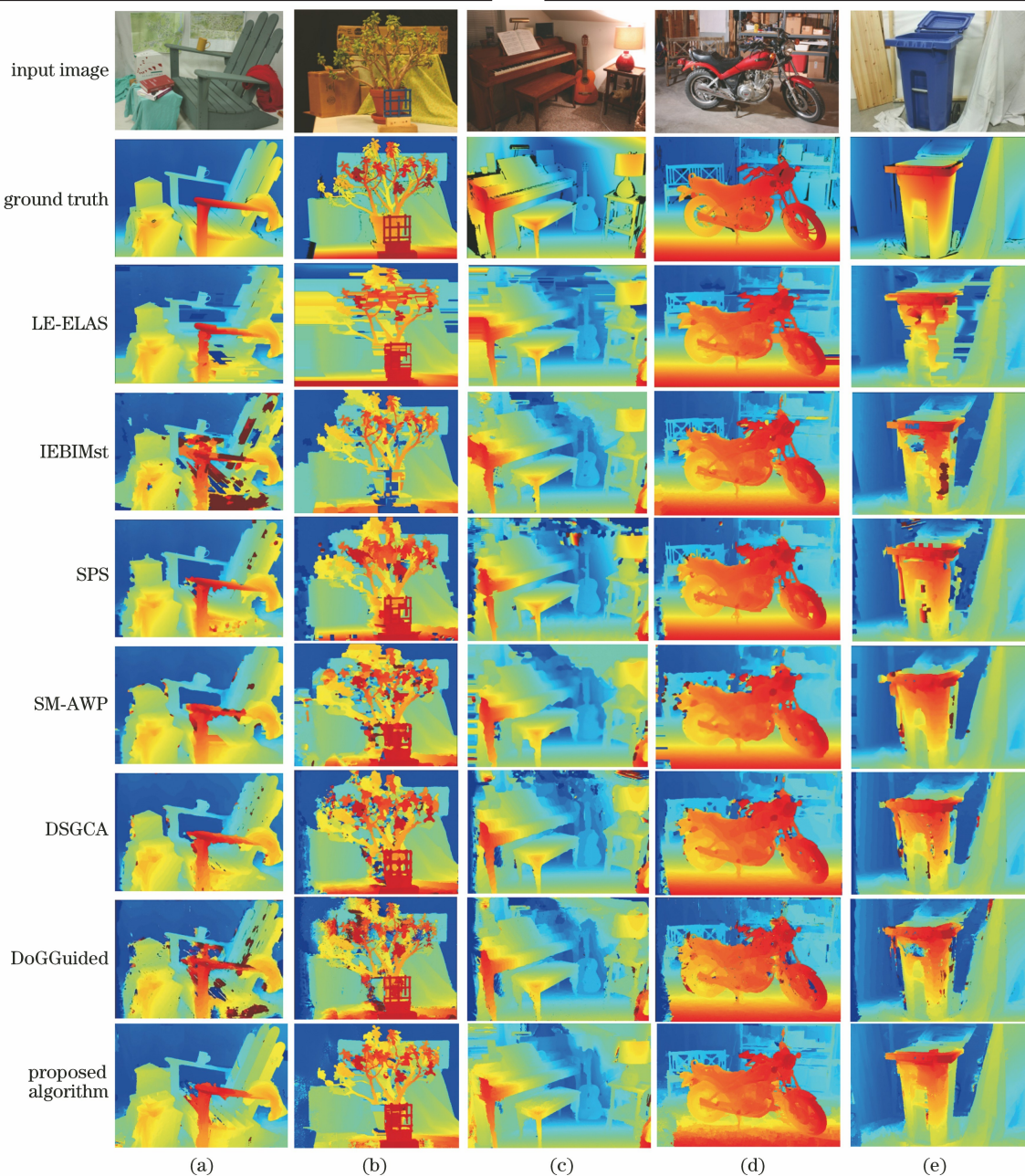


图 6 视差结果对比。(a) Adirondack; (b) Jadeplant; (c) Piano; (d) Motorcycle; (e) Recycle

Fig. 6 Comparison of disparity results. (a) Adirondack; (b) Jadeplant; (c) Piano; (d) Motorcycle; (e) Recycle

表 4 全部区域的 Avgerr 对比
Table 4 Comparison of Avgerr in all regions

pixel

Image name	LE-ELAS	IEBIMst	SPS	SM-AWP	DSGCA	DoGGuided	Proposed algorithm
Adirondack	9.31	27.30	6.51	10.50	7.68	20.10	6.40
ArtL	5.90	15.10	15.20	19.90	21.70	28.00	9.00
Jadeplant	64.50	55.60	40.00	62.70	45.00	56.50	26.10
Motorcycle	7.24	5.54	8.35	11.00	10.60	13.80	8.11
MotorcycleE	7.65	8.21	8.45	12.50	10.40	16.80	11.40
Piano	6.25	6.40	12.00	9.08	11.50	13.40	6.15
PianoL	9.69	18.90	25.00	29.70	24.50	37.30	34.00
Pipes	12.80	11.80	16.10	21.10	19.90	23.80	14.90
Playroom	10.10	18.00	25.20	20.70	24.60	30.30	10.50
Playtable	23.90	17.90	15.70	9.50	34.50	30.80	16.70
PlaytableP	4.27	4.95	12.40	9.75	14.80	13.00	10.00
Recycle	7.39	5.29	8.81	7.18	7.56	9.13	4.20
Shelves	8.48	17.10	23.70	11.40	17.30	19.00	9.97
Teddy	2.98	5.31	8.01	9.44	12.20	13.40	3.35
Vintage	14.00	10.90	53.70	16.80	43.80	13.60	10.90
Australia	15.20	9.17	8.64	19.10	16.60	18.20	12.00
AustraliaP	6.94	5.54	8.77	18.20	12.40	12.60	8.31
Bicycle2	6.68	7.54	11.40	16.00	12.90	17.60	13.70
Classroom2	24.60	27.90	20.20	29.30	32.60	34.90	9.09
Classroom2E	69.60	55.00	27.00	51.10	39.30	76.30	67.10
Computer	12.40	13.80	22.20	22.50	20.60	22.10	13.20
Crusade	21.70	74.30	50.80	91.80	49.50	73.40	36.30
CrusadeP	21.00	74.60	50.20	94.90	50.50	71.30	35.60
Djembe	2.73	2.10	3.65	7.33	5.71	6.64	2.98
DjembeL	13.80	29.10	17.20	31.80	24.50	39.00	19.50
Hoops	22.80	45.00	38.70	37.70	36.30	56.60	23.00
Livingroom	10.30	9.49	30.40	16.80	22.90	25.90	7.18
Newkuba	16.20	13.30	20.30	28.50	23.20	28.70	11.30
Plants	43.30	23.30	26.20	32.20	27.70	33.90	25.80
Staircase	21.30	30.90	39.40	36.40	39.80	57.50	29.80
Weighted average	15.95	20.95	20.70	26.50	22.80	29.20	15.20

表 5 非遮挡区域的 Avgerr 对比
Table 5 Comparison of Avgerr in non-occluded regions

Image name	LE-ELAS	IEBIMst	SPS	SM-AWP	DSGCA	DoGGuided	Proposed algorithm	pixel
Adirondack	8.46	26.10	3.57	6.31	3.25	15.20	4.84	
ArtL	3.83	4.67	5.34	9.65	5.95	9.57	4.62	
Jadeplant	41.10	41.90	22.80	31.80	18.90	27.10	16.10	
Motorcycle	5.12	2.72	3.11	4.71	3.60	5.64	4.58	
MotorcycleE	5.80	4.99	3.15	6.39	3.41	8.31	7.72	
Piano	5.54	5.69	9.34	6.68	7.17	8.09	5.20	
PianoL	8.97	17.50	22.90	28.40	21.10	32.40	34.40	
Pipes	7.44	5.47	6.78	10.60	7.23	9.67	7.53	
Playroom	8.76	12.90	12.50	9.08	9.36	14.00	5.05	
Playtable	22.40	14.80	9.70	5.09	29.40	24.50	13.00	
PlaytableP	3.47	3.26	7.64	5.18	7.94	5.32	5.67	
Recycle	6.93	4.99	6.27	3.86	3.80	5.56	3.37	
Shelves	8.26	16.40	22.30	9.73	14.70	16.20	9.49	
Teddy	2.29	2.64	1.52	3.64	3.51	4.15	2.15	
Vintage	13.10	10.40	52.60	10.70	39.70	15.00	9.64	
Australia	13.40	6.53	5.32	13.50	11.00	12.30	8.48	
AustraliaP	5.27	3.36	5.48	12.70	6.75	6.62	5.70	
Bicycle2	4.88	5.04	7.70	11.00	7.01	11.20	10.50	
Classroom2	19.30	19.30	5.60	17.50	13.70	16.30	5.35	
Classroom2E	66.50	45.70	12.50	41.80	21.50	62.60	64.80	
Computer	6.06	3.41	8.05	11.00	5.90	6.83	3.92	
Crusade	15.60	51.30	15.10	70.10	6.72	34.00	19.20	
CrusadeP	13.70	46.40	13.10	72.30	5.85	30.60	15.30	
Djembe	1.94	1.52	1.84	4.01	2.78	3.65	2.23	
DjembeL	13.30	29.20	16.10	28.30	22.20	37.00	19.20	
Hoops	18.20	39.20	22.80	25.80	17.20	35.00	16.90	
Livingroom	9.62	8.77	19.40	7.88	11.90	13.40	6.19	
Newkuba	12.80	8.19	12.60	21.50	11.10	14.20	7.60	
Plants	35.40	16.90	14.50	19.70	14.10	19.10	17.90	
Staircase	19.10	27.60	27.50	21.70	23.80	34.40	24.40	
Weighted average	12.43	15.10	11.00	17.36	10.23	15.85	10.36	

表 6 全部区域的 RMSE 对比
Table 6 Comparison of RMSE in all regions

pixel

Image name	LE-ELAS	IEBIMst	SPS	SM-AWP	DSGCA	DoGGuided	Proposed algorithm
Adirondack	25.40	58.90	23.00	30.30	25.10	48.90	19.20
ArtL	16.80	38.80	39.90	38.70	48.60	59.50	22.80
Jadeplant	120.00	128.00	95.80	119.00	102.00	118.00	62.60
Motorcycle	23.30	20.10	29.10	33.80	32.50	39.40	23.20
MotorcycleE	24.00	27.40	29.80	36.90	32.40	44.60	28.90
Piano	13.90	13.90	32.10	19.30	29.20	31.60	12.40
PianoL	18.70	39.00	53.80	56.30	50.80	65.80	65.30
Pipes	30.70	30.20	41.00	47.30	47.40	52.60	35.10
Playroom	25.90	38.00	59.50	47.80	58.30	66.70	27.60
Playtable	52.30	38.60	43.30	25.00	68.30	59.50	40.10
PlaytableP	12.60	14.20	37.20	25.50	38.40	34.40	28.00
Recycle	17.60	17.50	26.50	22.50	23.30	25.70	10.60
Shelves	15.30	30.40	44.40	23.80	33.50	33.90	19.40
Teddy	8.32	18.60	27.80	25.90	34.90	36.70	13.60
Vintage	27.40	28.80	131.00	45.80	105.00	66.10	27.70
Australia	34.90	27.90	30.00	46.60	40.80	45.20	32.60
AustraliaP	26.60	22.90	30.90	45.80	35.90	38.30	27.50
Bicycle2	21.00	23.50	32.10	37.70	32.60	40.20	32.30
Classroom2	55.80	68.70	60.90	64.10	82.60	85.40	27.30
Classroom2E	112.00	106.00	76.40	93.10	88.30	148.00	127.00
Computer	28.20	34.10	53.70	42.30	46.70	48.00	31.70
Crusade	58.70	156.00	141.00	152.00	131.00	151.00	79.40
CrusadeP	59.00	160.00	140.00	156.00	134.00	150.00	80.10
Djembe	8.10	7.30	15.00	27.80	19.70	21.50	9.28
DjembeL	31.00	58.60	42.80	61.30	51.40	66.40	42.70
Hoops	51.70	77.90	78.20	72.60	73.70	99.90	50.10
Livingroom	23.00	24.90	67.90	39.80	52.80	58.60	17.60
Newkuba	53.30	38.10	65.20	78.20	68.00	81.90	33.70
Plants	72.60	54.60	59.30	63.90	62.90	69.40	55.70
Staircase	46.00	48.20	85.80	73.60	78.70	102.00	51.20
Weighted average	36.25	48.15	54.85	55.20	56.15	64.50	35.70

表 7 非遮挡区域的 RMSE 对比

Table 7 Comparison of RMSE in non-occluded regions

pixel

Image name	LE-ELAS	IEBIMst	SPS	SM-AWP	DSGCA	DoGGuided	Proposed algorithm
Adirondack	24.60	58.00	16.40	21.30	13.20	42.20	15.80
ArtL	13.80	15.60	17.20	22.70	17.90	28.30	14.10
Jadeplant	91.90	121.00	75.70	81.50	63.60	75.70	46.80
Motorcycle	18.30	11.40	15.50	19.10	14.90	21.40	14.80
MotorcycleE	20.00	20.90	15.50	24.00	14.50	28.80	22.30
Piano	13.10	12.70	27.50	14.50	20.20	19.40	10.50
PianoL	17.90	37.40	52.00	56.20	47.30	60.90	66.80
Pipes	23.30	19.20	24.50	32.20	24.60	29.60	24.50
Playroom	25.50	31.50	35.80	27.20	25.80	37.50	13.90
Playtable	51.20	32.40	33.10	14.30	63.90	52.20	33.70
PlaytableP	11.30	9.20	29.10	15.30	25.60	15.50	15.30
Recycle	17.10	16.90	21.90	12.90	13.60	18.10	8.36
Shelves	15.10	29.40	43.50	20.20	29.80	29.70	18.60
Teddy	6.86	10.70	6.63	12.10	14.40	14.20	9.72
Vintage	26.30	29.00	134.00	26.20	104.00	50.70	25.00
Australia	31.50	21.30	22.40	37.80	31.00	35.40	25.20
AustraliaP	22.80	16.40	23.50	36.70	24.90	25.90	22.40
Bicycle2	17.10	17.40	26.00	29.90	22.00	30.70	27.30
Classroom2	49.80	57.50	26.10	48.50	47.10	52.60	19.50
Classroom2E	112.00	99.20	53.30	86.30	59.10	136.00	126.00
Computer	16.60	9.47	30.00	22.50	17.50	20.50	11.30
Crusade	49.80	131.00	76.70	127.00	32.90	84.20	51.00
CrusadeP	48.60	123.00	71.20	131.00	31.10	79.20	37.90
Djembe	5.73	5.54	9.13	17.90	11.80	14.00	7.17
DjembeL	30.80	59.20	41.70	55.00	49.20	65.10	43.00
Hoops	47.50	72.70	59.70	60.10	46.30	73.20	43.40
Livingroom	22.70	24.30	51.80	21.30	32.60	35.30	16.00
Newkuba	52.00	26.50	54.00	73.80	42.50	50.50	27.90
Plants	62.40	45.00	42.40	46.60	42.90	47.90	44.20
Staircase	42.00	41.80	75.40	56.20	60.60	66.20	43.50
Weighted average	31.45	39.15	37.45	41.20	31.45	41.50	26.65

由表 4~7 可知:所提算法在 Jadeplant、Piano、Recycle、Classroom2、Livingroom 图像上取得了最低的误差;在 Adirondack、Playroom、Vintage、Newkuba、Hoops、Computer 图像上取得了较好的结果;对于鲁棒性要求较高的图像如 ArtL、MotorcycleE、PlaytableP、AustraliaP、CrusadeP、DjembeL、Avgerr 较低,说明所提算法在不同光照和曝光条件下具有较好的鲁棒性;另外,PianoL 和 Classroom2E 图像的误差较高,主要是图像对之间的亮度差异较大,对匹配精度产生了较大影响。总之,所提算法除了非遮挡区域 Avgerr 的加权平均值略高于 DSGCA 算法,其余的三个加权平均值均低于另外 6 种算法,因此所提算法在 training set 和 test set 图像上的综合性能更加稳定。

4 结 论

提出一种基于引导图像和自适应支持域的局部立体匹配算法。在代价计算阶段,结合输入图像和引导图像,提出了一种梯度计算方法用来计算 x 和 y 方向梯度,并与 AD 和 Census 变换融合构建匹配代价计算函数。该梯度计算方法提高了匹配精度,取得了较好的效果。在视差细化阶段,提出一套基于自适应支持域的多步细化方法,该方法无论是在全部区域还是在非遮挡区域均能有效减小误差。实验结果表明,所提算法具有较好的鲁棒性,能获得精度较高的视差结果,但对亮度差异较大的图像对,结果并不理想,后续工作将进一步研究该问题。

参 考 文 献

- [1] Li Y T. Research on three dimensional accurate measurement method based on stereo vision [D]. Wuhan: Huazhong University of Science and Technology, 2016: 2-10.
李云廷. 基于立体视觉的三维精确测量方法研究 [D]. 武汉: 华中科技大学, 2016: 2-10.
- [2] Hamzah R A, Ibrahim H, Hassan A H A. Stereo matching algorithm for 3D surface reconstruction based on triangulation principle [C]//2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), August 23-24, 2016, Yogyakarta, Indonesia. New York: IEEE, 2016: 119-124.
- [3] Ge Z, Zhu Y, Liang G H. A 3D terrain reconstruction method of stereo vision based quadruped robot navigation system[J]. Proceedings of SPIE, 2017, 10322: 1032242.
- [4] Xiao X W, Guo B X, Li D R, et al. Multi-view stereo matching based on self-adaptive patch and image grouping for multiple unmanned aerial vehicle imagery[J]. Remote Sensing, 2016, 8(2): 89-119.
- [5] Žbontar J, LeCun Y. Stereo matching by training a convolutional neural network to compare image patches [J]. The Journal of Machine Learning Research, 2016, 17(1):1-32.
- [6] Pang J H, Sun W X, Ren J S, et al. Cascade residual learning: a two-stage convolutional neural network for stereo matching [C]//2017 IEEE International Conference on Computer Vision Workshops (ICCVW), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 878-886.
- [7] Chang J R, Chen Y S. Pyramid stereo matching network [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT. New York: IEEE, 2018: 5410-5418.
- [8] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts [C]//Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE, 2001: 508-515.
- [9] Sun J, Zheng N N, Shum H Y. Stereo matching using belief propagation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7): 787-800.
- [10] Hirschmüller H. Stereo processing by semiglobal matching and mutual information [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 328-341.
- [11] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. International Journal of Computer Vision, 2002, 47: 7-42.
- [12] Mei X, Sun X, Zhou M C, et al. On building an accurate stereo matching system on graphics hardware [C]//2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), November 6-13, 2011, Barcelona, Spain. New York: IEEE, 2011: 467-474.
- [13] Zhu S P, Yan L N. Local stereo matching algorithm with efficient matching cost and adaptive guided image filter[J]. The Visual Computer, 2017, 33(9): 1087-1102.
- [14] Zhan Y L, Gu Y Z, Huang K, et al. Accurate image-guided stereo matching with efficient matching cost and disparity refinement [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26

- (9): 1632-1645.
- [15] Zhang K, Lu J B, Lafruit G. Cross-based local stereo matching using orthogonal integral images[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2009, 19(7): 1073-1079.
- [16] Yan L, Wang R, Liu H, et al. Stereo matching method based on improved cost computation and adaptive guided filter[J]. *Acta Optica Sinica*, 2018, 38(11): 1115007.
闫利, 王芮, 刘华, 等. 基于改进代价计算和自适应引导滤波的立体匹配[J]. *光学学报*, 2018, 38(11): 1115007.
- [17] 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.
- [18] Yang Q X. A non-local cost aggregation method for stereo matching [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI. New York: IEEE, 2012: 1402-1409.
- [19] 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.
- [20] Yang Q Q, Ji P, Li D X, et al. Fast stereo matching using adaptive guided filtering[J]. *Image and Vision Computing*, 2014, 32(3): 202-211.
- [21] Stentoumis C, Grammatikopoulos L, Kalisperakis I, et al. On accurate dense stereo-matching using a local adaptive multi-cost approach[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 91: 29-49.
- [22] Ma Z Y, He K M, Wei Y C, et al. Constant time weighted median filtering for stereo matching and beyond[C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, Australia. New York: IEEE, 2013: 49-56.
- [23] Jiao J B, Wang R G, Wang W M, et al. Local stereo matching with improved matching cost and disparity refinement[J]. *IEEE Multimedia*, 2014, 21(4): 16-27.
- [24] 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.
- [25] Hosni A, Rhemann C, Bleyer M, et al. Fast cost-volume filtering for visual correspondence and beyond [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(2): 504-511.
- [26] Mei X, Sun X, Dong W M, et al. Segment-tree based cost aggregation for stereo matching[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, OR, USA. New York: IEEE, 2013: 313-320.
- [27] Jellal R A, Lange M, Wassermann B, et al. LS-ELAS: line segment based efficient large scale stereo matching [C]//2017 IEEE International Conference on Robotics and Automation (ICRA), May 29-June 3, 2017, Singapore. New York: IEEE, 2017: 146-152.
- [28] He C, Zhang C X, Chen Z, et al. Minimum spanning tree based stereo matching using image edge and brightness information[C]//2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), October 14-16, 2017, Shanghai, China. New York: IEEE, 2017: 1-5.
- [29] Legendre C, Batsos K, Mordohai P. High-resolution stereo matching based on sampled photoconsistency computation [C]//British Machine Vision Conference, September 4-7, 2017, London, UK. Guildford: BMVA Press, 2017.
- [30] Razak S S A, M. Othman, et al. The effect of adaptive weighted bilateral filter on stereo matching algorithm[J]. *International Journal of Engineering and Advanced Technology*, 2019, 8(3): 284-287.
- [31] Williem, Park I K. Deep self-guided cost aggregation for stereo matching[J]. *Pattern Recognition Letters*, 2018, 112: 168-175.
- [32] Kitagawa M, Shimizu I, Sara R. High accuracy local stereo matching using DoG scale map [C]//2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), May 8-12, 2017, Nagoya, Japan. New York: IEEE, 2017: 258-261.