

# 一种改进 Census 变换与梯度融合的立体匹配算法

范海瑞<sup>1,2</sup>, 杨帆<sup>1,2</sup>, 潘旭冉<sup>1,2</sup>, 温洁<sup>1,2</sup>, 王晓宇<sup>1,2</sup>

<sup>1</sup>河北工业大学电子信息工程学院, 天津 300401;

<sup>2</sup>天津市电子材料与器件重点实验室, 天津 300401

**摘要** 针对现有立体匹配算法对噪声敏感、易失真、在视差不连续区域与弱纹理区域误匹配率高的问题, 提出一种改进 Census 变换与梯度融合的多尺度立体匹配算法。采用支持窗口内所有像素的加权平均灰度值作为 Census 变换的参考值, 将 Census 代价与由水平和垂直方向归一化结合的梯度代价进行加权融合, 通过设置噪声容限获得稳定的代价, 提高了单像素匹配代价的可靠性; 在分辨率尺度下, 采用改进引导滤波算法完成对匹配代价的聚合; 通过视差提取获得视差图。实验结果表明, 该算法在 Middlebury 测试平台上对标准立体图像对的平均误匹配率为 4.74%, 对 27 组扩展立体图像对的平均误匹配率为 8.67%。该算法使得视差不连续区域与弱纹理区域的误匹配率进一步降低, 且对噪声和光照等干扰表现出较好的稳健性。

**关键词** 机器视觉; 立体匹配; Census 变换; 梯度变换; 引导滤波

中图分类号 TN911.73

文献标识码 A

doi: 10.3788/AOS201838.0215006

## Stereo Matching Algorithm for Improved Census Transform and Gradient Fusion

Fan Hairui<sup>1,2</sup>, Yang Fan<sup>1,2</sup>, Pan Xuran<sup>1,2</sup>, Wen Jie<sup>1,2</sup>, Wang Xiaoyu<sup>1,2</sup>

<sup>1</sup>School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China;

<sup>2</sup>Tianjin Key Laboratory of Electronic Materials and Devices, Tianjin 300401, China

**Abstract** Aiming at the problems of noise-sensitive, easy distortion and with high false matching ratio in the disparity discontinuity region and weak texture region of the existing local matching algorithm, a multi-scale stereo matching algorithm for improved Census transform and gradient fusion is proposed. The weighted average gray value of all the pixels in the support window is used as the reference value of the Census transform. The Census cost is weighted combined with the gradient cost normalized by the horizontal and vertical directions, and a stable cost is obtained when the noise margin is set. Therefore, the reliability of the single pixel matching cost is obtained. Under the multi-scale, the improved guided filtering algorithm is used to complete the aggregation of the matching cost. The disparity map is obtained by parallax extraction. The experimental results demonstrate that the average false matching ratio of standard stereo image pairs obtained by the proposed algorithm is 4.74% on the Middlebury testing benchmark, and the average false matching ratio of the 27 extended stereo image pairs is 8.67%. In the parallax discontinuity region and the weak texture region, the false matching ratio is further reduced by the proposed algorithm, and it shows better robustness for noise and light.

**Key words** machine vision; stereo matching; Census transform; gradient transform; guided filter

**OCIS codes** 150.0155; 100.6890; 110.2960; 330.1400

## 1 引 言

立体匹配是机器视觉领域的研究热点之一, 通过查找基于不同视角拍摄的两幅或者多幅同一场景

不同图像间的匹配像素点, 计算匹配像素点间的二维视差, 然后将二维视差转换为三维深度信息, 从而构建出相应场景的三维环境模型, 在机器人导航、无人驾驶、虚拟现实和三维重建等领域有着广泛的应

收稿日期: 2017-08-24; 收到修改稿日期: 2017-10-02

基金项目: 河北省自然科学基金(E2016202341)、河北省高等学校科学技术研究项目(BJ2014013)。

作者简介: 范海瑞(1992—), 男, 硕士研究生, 主要从事机器视觉与模式识别方面的研究。E-mail: HairuiJY@126.com

导师简介: 杨帆(1966—), 男, 博士, 教授, 博士生导师, 主要从事智能信息、图像处理与机器视觉方面的研究。

E-mail: commanderjy@163.com(通信联系人)

用<sup>[1]</sup>。Scharstein 等<sup>[2]</sup>全面概述了一些代表性的立体匹配算法,建立了较全面的测评体系,提供了相应的测试数据集及比较方法。根据最优化方法的不同,立体匹配可分为全局算法和局部算法两类。常用的全局算法包括动态规划(DP)<sup>[3]</sup>、置信传播(BP)<sup>[4]</sup>、图割法(GC)<sup>[5]</sup>、最小生成树(MST<sup>[6]</sup>、MST-DC<sup>[7]</sup>)和分割树(ST)<sup>[8]</sup>等,全局算法精度较高,但是复杂度高、实时性差。局部算法可计算局部窗口像素的最优值,常见的局部算法有绝对误差和算法(SAD)、误差平方和算法(SSD)、归一化积相关算法(NCC)和自适应窗口<sup>[9]</sup>等,具有复杂度低、效率高和便于实现的优点,但在噪声和光照等干扰的条件下,局部算法的误匹配率( $R$ )显著提高。

Pollard 等<sup>[10]</sup>证明通过将图像梯度信息引入到匹配代价计算中,可以增强匹配代价对幅度等失真的抗干扰能力;Zabih 等<sup>[11]</sup>提出将非参数变换引入到图像匹配中,可以显著减小光照和噪声等干扰的影响;Hirschmuller 等<sup>[12]</sup>得出 Census 变换在光照失真的条件下具有很好的稳健性。非参数空间变换可以抑制噪声等干扰,将局部算法与非参数变换相结合,不仅可以降低算法复杂度,还可以增强匹配的抗干扰能力。

代价聚合可看作对匹配代价滤波的过程,以往用局部线性滤波器(如盒式滤波器和高斯滤波器等)在聚合阶段进行滤波处理,但由于各项同性的特性,不能很好地保护视差图中物体的边缘信息。Yoon 等<sup>[13]</sup>提出了基于双边滤波器(BF)的匹配算法,具有保护图像边缘的特性,但该算法的计算量与采用的滤波核窗口大小有关,计算复杂度较高。Hosni 等<sup>[14]</sup>提出了基于引导滤波器(GF)的匹配算法,相比 BF 具有较好的边缘保护特性,而且计算复杂度与支持窗口大小无关,计算量小,常被用来对匹配代价进行处理。以往在单分辨率尺度下进行代价聚合时,不能充分利用图像内容信息,Zhang 等<sup>[15]</sup>建立了跨尺度代价聚合框架,在不同分辨率尺度下用聚合算法进行代价聚合,可获得较好的视差图。

综合考虑上述算法的优缺点,提出一种改进 Census 变换与梯度融合的多尺度立体匹配算法;在 Census 代价计算阶段,用改进 Census 变换计算得到 Census 代价,将带有噪声容限的归一化结合梯度代价与 Census 代价进行加权融合,得到初始匹配代价;在代价聚合阶段,分别在不同分辨率尺度下对代价进行改进引导滤波,得到最终的匹配代价,很好地抑制了像素失真对代价计算的影响,增强了立体匹

配对噪声和光照等干扰的稳健性,提高了视差图的质量。

## 2 算法描述

所提算法由匹配代价计算、匹配代价聚合、视差计算和视差精化等步骤组成。

### 2.1 匹配代价计算

传统 Census 变换是将中心像素灰度值与支持窗口邻域内像素进行比较生成匹配模板,然后进行非参数变换获得比特串,最后用 Hamming 距离计算匹配代价,算法对中心像素的依赖性强、抗干扰能力弱。Chang 等<sup>[16]</sup>提出用支持窗口内所有像素的均值作为参考值的算法(MCT),进一步提高单像素匹配代价的可靠性,但不能很好地利用中心像素与邻域像素间的相互关系;Zhu 等<sup>[17-18]</sup>在 MCT 中增加了噪声容限参考值(R-MCT),提高了单像素匹配代价的可靠性,但是仍然不能充分利用中心像素与邻域像素间的相互关系。

针对上述变换存在的问题,提出一种改进的 Census 变换来计算匹配代价(SWCT),每一个像素点的值都由其本身和邻域内其他像素经过加权平均后得到。算法具体操作步骤为:用支持窗口去扫描图像中的每一个像素,用支持窗口内邻域像素的加权平均灰度值作为 Census 变换参考值,即越靠近中心像素权重越高,越远离像素权重越低,权重分布如图 1 所示。改进算法参考值计算公式与权重计算公式为

$$I_{wm} = \frac{1}{D} \sum_{p \in N} W_{pq} I_p, \quad (1)$$

$$W_{pq} = \exp[-(\omega_x + \omega_y)^2 / \sigma^2], \quad (2)$$

式中  $I_{wm}$  为参考值,  $I_p$  为像素为  $p$  的灰度值,  $W_{pq}$  为像素  $q$  距中心像素  $p$  的权重,  $D = \sum W_{pq}$  为所有权重的和,  $N$  为支持窗口内的像素,  $p, q$  为像素索引,  $\omega_x, \omega_y$  为像素在支持窗口内的坐标,  $\sigma$  为标准差。通过 Census 变换过程计算得到该窗口的比特串为

$$c_{cen}(p) = \bigotimes_{p \in N} \xi(I_{wm}, I_q), \quad (3)$$

其中

$$\xi(x, y) = \begin{cases} 1, & x < y \\ 0, & \text{else} \end{cases}, \quad (4)$$

式中  $\bigotimes$  表示按位连接,  $N'$  为支持窗口内的邻域像素,  $I_q$  为邻域像素值。基于改进 Census 变换的匹配代价计算模型为

$$C_{cen}(p, d) = \min\{\text{Hamming}[c_{cen}(p), lc_{cen}(p-d)], T_{cen}\}, \quad (5)$$

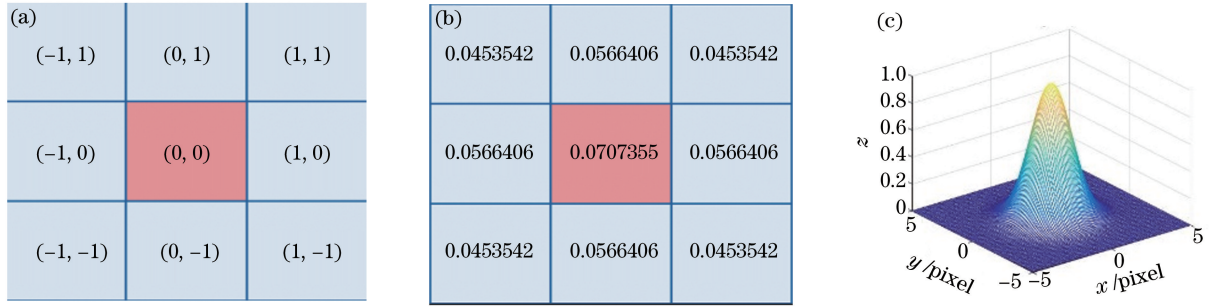


图 1 大小为  $3 \times 3$ 、标准差为 1.5 的支持窗口坐标与权重分布图。(a)坐标分布图;(b)权重坐标分布图;(c)权重分布图

Fig. 1 Support window coordinates and weight distribution map with size of  $3 \times 3$  and standard deviation of 1.5.

(a) coordinate distribution map; (b) weight coordinate distribution map; (c) weight distribution map

式中  $C_{\text{cen}}(p, d)$  为视差值为  $d$  时像素  $p$  的匹配代价值,  $T_{\text{cen}}$  为基于改进 Census 变换的截断值, Hamming 为计算汉明距离,  $c_{\text{cen}}(p)$  为像素  $p$  对应的二进制比特串,  $c_{\text{cen}}(p-d)$  为像素  $p-d$  对应的二进制比特串。

为进一步增强算法对于干扰的稳健性,同时防止边缘区域在加权后过于平滑,所提算法引入梯度变换,将 Census 变换计算的匹配代价与梯度变换计算的代价进行加权融合,这样既有利于提高算法对噪声的抗干扰性,也有利于视差不连续区域匹配代价的计算。传统基于梯度的匹配代价计算仅考虑  $x$  方向的梯度<sup>[15]</sup>,或者将  $x$ 、 $y$  方向的梯度进行平方<sup>[19]</sup>。为了充分体现像素灰度值在水平和垂直方向梯度所占的权重信息,所提算法对  $x$ 、 $y$  方向的梯度信息进行归一化结合,更好地体现了主导梯度方向的作用。定义  $x$ 、 $y$  方向的梯度向量为

$$\begin{cases} g_x = (\nabla_x G_R, \nabla_x G_G, \nabla_x G_B) \\ g_y = (\nabla_y G_R, \nabla_y G_G, \nabla_y G_B) \end{cases}, \quad (6)$$

式中 R、G、B 分别代表红、绿、蓝 3 个颜色通道分量,  $\nabla_x$ 、 $\nabla_y$  分别为  $x$ 、 $y$  方向的梯度算子;  $G_R$ 、 $G_G$ 、 $G_B$  分别为在 3 个颜色通道的梯度;  $g_x$ 、 $g_y$  分别为  $x$ 、 $y$  方向的梯度。基于改进梯度变换的匹配代价计算公式为

$$C_g(p, d) = \min [\omega_1 \|g_x^L(p) - g_x^R(p-d)\| + \omega_2 \|g_y^L(p) - g_y^R(p-d)\|, T_g], \quad (7)$$

式中  $C_g(p, d)$  为当像素  $p$  的视差值为  $d$  时的匹配代价值;  $\omega_1$ 、 $\omega_2$  为  $x$ 、 $y$  方向的权值,其中  $\omega_1 + \omega_2 = 1$ ; 上标 L、下标 R 分别表示左和右;  $T_g$  为基于梯度代价的截断值;  $g_x^L(p)$  为左图像在像素  $p$  处的梯度;  $g_x^R(p-d)$  为右图像在像素  $p-d$  处的梯度。

将改进的 Census 变换和梯度的匹配代价进行加权融合,具体表示公式为

$$C(p, d) = 2 - \exp[-C_g(p, d)/T_g] - \exp[-C_{\text{cen}}(p, d)/T_{\text{cen}}]. \quad (8)$$

## 2.2 匹配代价聚合

为了更好地保护视差图的边缘信息,提出一种改进的引导滤波算法;同时为了更充分地利用图像细节信息并降低弱纹理区域的误匹配率,在多分辨率尺度下进行代价聚合。建立引导滤波模型为

$$C'(p, d) = aG(p) + b, \quad \forall p \in \omega_k, \quad (9)$$

式中  $k$  为窗口  $\omega_k$  内的邻域像素,  $C'(p, d)$  为输出代价图,  $G(p)$  为输入引导图,  $a$ 、 $b$  为线性函数的系数。滤波是一个求最优解的过程,传统引导滤波算法构建的优化函数模型为

$$E(a'_k, b'_k) = \sum_{k \in \omega_k} \{ [a'_k G(p) + b'_k - C(p)]^2 + \lambda a'^2_k \}, \quad (10)$$

式中  $C(p)$  为输入待滤波图,即匹配代价矩阵;  $a'_k$  为系数矩阵;  $b'_k$  为系数向量;  $\lambda$  为正则化系数,对系数矩阵  $a'_k$  进行约束,防止过大。

在视差不连续与高纹理区域中,边缘对获得视差图的好坏有非常大的影响。为了获得更好的边缘保持特性,引入基于权重的边缘保持项  $W(G, p)$ , 建立新的引导滤波模型为

$$E(a_k, b_k) = \sum_{k \in \omega_k} \{ [a_k G(p) + b_k - C(p)]^2 + \lambda a_k^2 W^{-1}(G, p) \}, \quad (11)$$

$$W(G, p) = \text{diag}(W_G^R, W_G^G, W_G^B), \quad (12)$$

$$W_G^C = N^{-1} \sum_{k=1}^N \frac{\delta_G^2(k) + \epsilon}{\delta_G^2(p) + \epsilon}, \quad (13)$$

式中  $a_k$  为系数矩阵;  $W(G, p)$  为引导图中以  $p$  为中心像素支持窗口的局部权重矩阵;  $W_G^C$  为局部权重的颜色分量,  $C$  代表 R、G、B;  $N$  为引导图像支持窗口内的像素总数;  $\delta_G(k)$ 、 $\delta_G(p)$  为对应支持窗口的标准差。通过最小二乘法求解可得

$$a_k = \left[ \sum_k \lambda W(G, p)^{-1} \right]^{-1} \times \left[ \frac{1}{\omega_k} \sum_{p \in \omega_k} G(p) C(p) - \mu_k \bar{C}_k \right], \quad (14)$$

$$\mathbf{b}_k = \overline{\mathbf{C}}_k - \mathbf{a}_k^T \boldsymbol{\mu}_k, \quad (15)$$

式中  $\boldsymbol{\mu}_k$  为引导图  $G$  在窗口  $\omega_k$  中的均值矩阵;  $|\omega_k|$  为窗口  $\omega_k$  中的像素数量;  $\overline{\mathbf{C}}_k$  为待滤波图像  $C$  在窗口  $\omega_k$  中的均值矩阵。

引导滤波算法获得的视差图如图 2 所示, 图中红色标注点为错误匹配像素点。从图中可以看

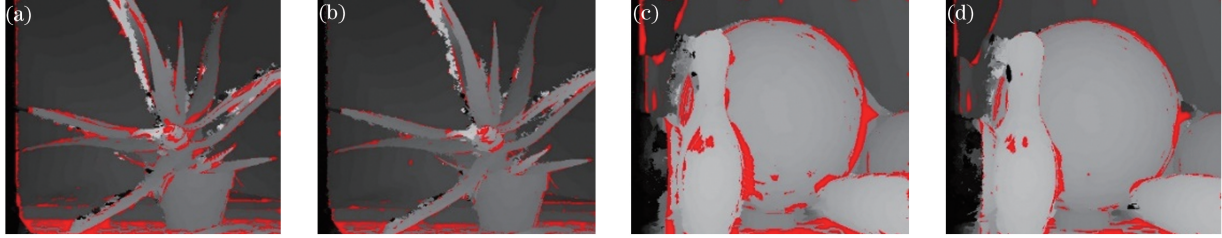


图 2 改进前后引导滤波算法获得的视差图。(a)(c)改进前;(b)(d)改进后

Fig. 2 Disparity maps before and after improvement by guidance filter algorithm.

(a)(c) before improvement; (b)(d) after improvement

在计算每个窗口的线性系数时,发现一个像素会被多个窗口包含,这时只需将所有包含该像素点的线性函数系数求均值即可,最终可得滤波模型为

$$\mathbf{C}(p, d) = \frac{1}{|\omega_k|} \sum_{p \in \omega_k} [\mathbf{a}_k \mathbf{C}(p) + \mathbf{b}_k] = \overline{\mathbf{a}}_k \mathbf{C}(p) + \overline{\mathbf{b}}_k, \quad (16)$$

式中  $\overline{\mathbf{a}}_k$ 、 $\overline{\mathbf{b}}_k$  分别为窗口  $\omega_k$  中  $\mathbf{a}_k$ 、 $\mathbf{b}_k$  的均值;  $\mathbf{C}(p, d)$  为经过引导滤波的代价。

在不同分辨率尺度下分别对匹配代价图进行改进引导滤波,具体过程参考文献[15],最终的聚合代价为

$$\hat{\mathbf{C}}(p, d) = \sum_{s=0}^S \mathbf{A}^{-1}(0, s) \tilde{\mathbf{C}}^0(p^0, d^0), \quad (17)$$

式中  $\mathbf{A}$  为求解过程中的系数矩阵,为  $(S+1) \times (S+1)$  大小的三对角常数矩阵;  $S$  为向下采样层数;  $\tilde{\mathbf{C}}^0(p^0, d^0)$  为第 0 层的匹配代价矩阵。

### 2.3 视差计算与精化

通过代价聚合得到最终匹配代价  $\hat{\mathbf{C}}(p, d)$ , 用 Winner-Takes-All(WTA) 策略进行视差选择即可获得初始视差图。每个像素最优视差获取公式为

$$d_{\text{opt}} = \underset{0 \leq d \leq d_{\text{max}}}{\operatorname{argmin}} \hat{\mathbf{C}}(p, d), \quad (18)$$

式中  $d_{\text{opt}}$  表示最小匹配代价对应的最优视差;  $d_{\text{max}}$  表示最大视差范围;  $p$  表示当前像素。

在视差精化阶段,首先通过左右一致性检测来检测误匹配异常点,具体做法是根据左右两幅输入图像,分别得到左右两幅视差图。对于左图中的一个像素点  $p$ ,求得的最优视差值是  $d_1$ ,那么像素点  $p$  在右图的对应点应该是像素点  $p - d_1$ ,像素点  $p - d_1$  的视差值记为  $d_2$ ,则

出,图 2(b)、(d)对应的视差图中红色误匹配标记点明显少于图 2(a)、(c),而且图 2(b)、(d)在图中边缘处视差不连续区域的误匹配明显优于图 2(a)、(c),这是由于改进引导滤波算法的约束条件并加入基于权重的边缘保持项后,算法误匹配率明显降低。

$$|d_1 - d_2| \leq T_h. \quad (19)$$

将不满足 (19) 式条件的像素点  $p$  视为异常点,  $T_h$  表示视差阈值。

对检测到的异常点进行修正,即对于一个异常点  $p$ ,分别水平往左、右找到第一个非异常点,记作  $p_l$ 、 $p_r$ 。将点  $p$  的视差值修正为  $p_l$  和  $p_r$  视差值  $d(p_l)$ 、 $d(p_r)$  中较小的那一个,计算公式为

$$d(p) = \min[d(p_l), d(p_r)]. \quad (20)$$

## 3 实验结果与分析

所提算法使用 Middlebury2.0 立体匹配评估测试平台提供的数据集进行实验,包括 2001, 2003, 2005, 2006 datasets 4 个数据集中 31 组立体图像对<sup>[20]</sup>。实验环境为 Visual studio 2013 下 C/C++ 编程环境, Windows 10, X64 系统, Intel Core i5-3210 主频为 2.5 GHz, 内存为 4 GB。

为了验证算法的实际性能,除特别说明外,文中涉及的视差图都未经过任何视差精化和后处理,为初始视差图;对比实验中误差限为 1 个像素,即当匹配结果与真实视差图相差一个像素以上时,认为该点是误匹配点。实验中涉及的参数设置如表 1 所示。

### 3.1 匹配代价计算验证

为了验证匹配代价算法的有效性,选取基于 SAD、Census 变换(CT)、MCT<sup>[16]</sup>、改进 Census 变换(MCT'<sup>[17-18]</sup>)、梯度变换(GRD<sup>[15]</sup>) 5 种不同匹配代价计算方法,对 Middlebury2.0 中 4 组标准立体图像对进行实验,与所提算法进行对比,通过非遮挡区域和所有区域的误匹配率来比较算法的性能,如表 2、3 所示。

表 1 所提立体匹配算法涉及的参数

Table 1 Parameters involved in proposed stereo matching algorithm

Parameter	$\omega_k$	$T_g$	$\epsilon$	$\sigma$	$T_h$	$\omega_1$	$T_{cen}$	$\lambda_G$	$\omega_2$
Value	9	4.335	$1 \times 10^{-5}$	1.5	1	0.9	45	0.065025	0.1

表 2 非遮挡区域不同匹配代价计算方法误匹配率

Table 2 Percentage of false match in no-occluded region of different matching cost algorithms %

Algorithm	SAD	CT	MCT	MCT'	GRD	Proposed algorithm
Tsukuba	4.65	3.54	4.58	4.50	2.72	2.53
Venus	3.45	2.00	3.63	3.55	1.68	1.60
Teddy	14.21	8.56	13.05	11.32	7.45	7.57
Cones	7.70	4.87	7.49	6.78	4.45	4.04
Avg	7.50	4.74	7.19	6.54	4.08	3.93

表 3 所有区域不同匹配代价计算方法误匹配率

Table 3 Percentage of false match in all regions of different matching cost algorithms %

Algorithm	SAD	CT	MCT	MCT'	GRD	Proposed algorithm
Tsukuba	4.97	3.12	3.36	3.59	3.59	2.86
Venus	8.34	5.25	7.17	7.15	4.12	4.13
Teddy	25.54	19.28	23.48	22.76	17.56	17.27
Cones	23.40	17.04	19.71	18.88	16.13	16.06
Avg	15.56	11.17	13.43	13.10	10.35	10.08

表中结果表明,所提匹配代价计算方法未经过视差精化时在非遮挡区域的平均误匹配率为 3.93%,所有区域的平均误匹配率为 10.08%,相比

于其他代价算法,所提算法的误匹配率进一步降低。

CT、MCT、GRD 与所提算法的视差图如图 3 所示,可以看出,所提算法的视差图中红色标记点明显少于其他算法,在边缘交界处、网格孔等处尤为明显;图 3(a)、(b)中弱纹理区域的误匹配点明显少于图 3(c),说明在弱纹理区域 Census 变换比梯度变换表现得更好;在边缘处视差不连续区域,图 3(c)的红色误匹配标记点明显比图 3(a)、(b)中少,说明梯度变换在视差不连续区域表现得更好;图 3(d)为所提算法的视差图,可以看出在弱纹理区域的误匹配红色标记点明显减少,表明所提算法的误匹配率低、精度高。

### 3.2 代价聚合验证

为验证所提代价聚合方法的性能,对所提算法计算的匹配代价分别采用 4 种经典的聚合算法进行聚合,通过非遮挡区域误匹配率来对比聚合算法的性能。对比算法有盒式滤波(BoxF)、双边滤波(BF)<sup>[13]</sup>、引导滤波(GF)<sup>[14]</sup>和最小生成树(MST)<sup>[6]</sup>聚合算法,对 4 组标准立体图像对进行对比实验,实验中得到的 Teddy 视差图如图 4 所示,红色标记点为误匹配点。非遮挡区域的 R 对比如图 5 所示。

从图 4 可以看出,图 4(e)在视差不连续区域的

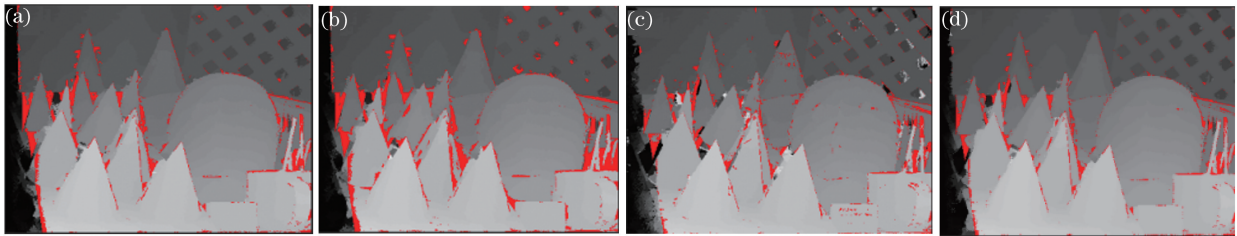


图 3 不同匹配代价算法获得的视差图。(a) CT; (b) MCT; (c) GRD; (d)所提算法

Fig. 3 Disparity maps obtained by different matching cost algorithms. (a) CT; (b) MCT; (c) GRD; (d) proposed algorithm

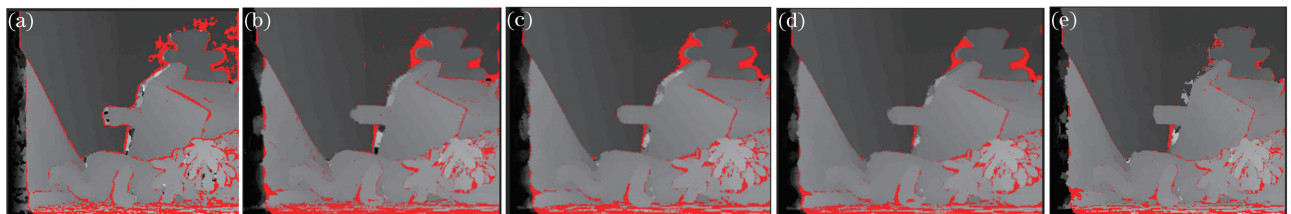


图 4 不同聚合算法的视差图。(a) BoxF,  $R=10.81\%$ ; (b) BF,  $R=8.12\%$ ; (c) GF,  $R=7.85\%$ ; (d) MST,  $R=8.31\%$ ; (e) 所提算法,  $R=7.57\%$

Fig. 4 Disparity maps obtained by different aggregation algorithms. (a) BoxF,  $R=10.81\%$ ; (b) BF,  $R=8.12\%$ ; (c) GF,  $R=7.85\%$ ; (d) MST,  $R=8.31\%$ ; (e) proposed algorithm,  $R=7.57\%$

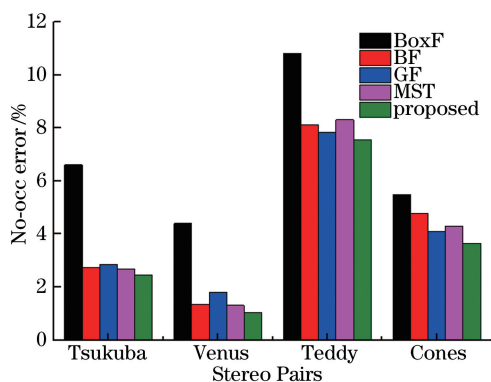


图 5 非遮挡区域(no-occ)不同代价聚合算法误匹配率对比图

Fig. 5 Comparison of false matching rates of different cost aggregation algorithms in no-occluded regions

红色误匹配标记点明显少于其他图,这是由于在改进引导滤波算法中加入边缘保持项,可使效果得到很大提升;与其他图相比,图 4(e)在平滑的弱纹理区域的效果也得到提升,这是由于融合多尺度充分利用了图像细节信息;图 5 可以直观地反映出所提算法在非遮挡区域的误匹配率最低、精度最高。实验结果表明,所提代价聚合算法可以进一步提高视差不连续区域和弱纹理区域的匹配

表 4 噪声条件下不同匹配代价计算方法在非遮挡区域的平均误匹配率

Table 4 Average false matching rate of different computing methods for matching costs in no-occluded region under noise conditions

Method	Noiseless	R under salt and pepper noise					R under Gaussian noise				
		1%	5%	10%	15%	Avg	0.1	0.5	1	2	Avg
SAD	7.50	15.54	33.44	51.32	76.51	36.86	14.36	32.67	53.37	88.97	47.34
CT	4.74	5.30	7.46	10.29	13.70	8.30	7.04	21.99	32.82	50.75	28.15
MCT	7.19	10.89	20.78	26.26	29.83	18.99	9.15	21.90	32.13	47.86	27.76
MCT'	6.54	9.88	21.45	28.87	33.23	19.99	8.77	20.78	33.32	47.95	27.71
GRD	4.08	4.63	6.73	8.29	9.65	6.68	8.30	42.24	64.61	79.78	48.73
Proposed algorithm	3.93	3.93	5.47	7.31	9.67	6.06	6.16	20.84	32.01	49.10	27.03

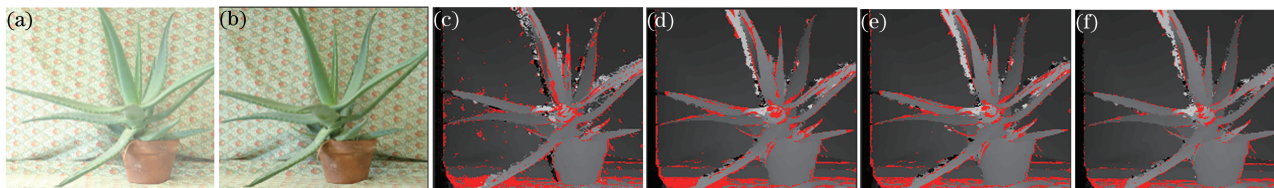


图 6 不同立体匹配算法在 Aloe 图像对上的测试结果。(a) Aloe 左原图;(b) Aloe 右原图;(c) GRD,  $R=10.19\%$ ;  
(d) MCT,  $R=10.07\%$ ;(e) MCT',  $R=9.74\%$ ;(f) 所提算法,  $R=6.74\%$

Fig. 6 Test results of different stereo matching algorithms on Aloe image pairs. (a) Aloe left image;(b) Aloe right image;  
(c) GRD,  $R=10.19\%$ ;(d) MCT,  $R=10.07\%$ ;(e) MCT',  $R=9.74\%$ ;(f) proposed algorithm,  $R=6.74\%$

精度,降低误匹配率。

### 3.3 抗干扰性验证

为测试所提算法的稳健性,选取 5 种算法(SAD、CT、MCT、MCT'和 GRD)与所提算法进行比较,分别对噪声和光照两种干扰进行对比测试。

1) 噪声测试。分别给 4 组标准立体图像加入噪声密度为 1%、5%、10%和 15%的椒盐噪声和标准差为 0.1,0.5,1.0,2.0 的高斯噪声,针对两种干扰来分析算法对噪声的稳健性;通过非遮挡区域的平均误匹配百分比来对比 6 种立体匹配算法的性能,实验结果如表 4 所示。

噪声实验结果表明,在增加噪声后,所提算法均优于其他几种匹配算法,而且随着噪声比例的增加,所提算法的误匹配率变化幅度最小,则所提算法对噪声变换具有更好的稳健性。

2) 光照测试。对 Middlebury2.0 数据集上的 Aloe、Baby1、Bowling2 和 Dolls 4 组立体图像在光照不一致的条件下进行对比实验,对比算法包括 GRD、MCT、MCT'和所提算法,同样通过非遮挡区域的误匹配百分比来比较算法的性能,实验结果如图 6~9 所示。

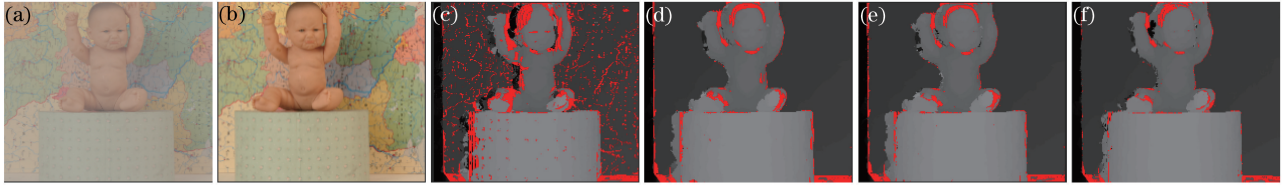


图 7 不同立体匹配算法在 Baby1 图像对上的测试结果。(a) Baby1 左原图；(b) Baby1 右原图；

(c) GRD,  $R=12.82\%$ ；(d) MCT,  $R=4.91\%$ ；(e) MCT',  $R=4.53\%$ ；(f) 所提算法,  $R=3.75\%$

Fig. 7 Test results of different stereo matching algorithms on Baby1 image pairs. (a) Baby1 left image; (b) Baby1 right image; (c) GRD,  $R=12.82\%$ ; (d) MCT,  $R=4.91\%$ ; (e) MCT',  $R=4.53\%$ ; (f) proposed algorithm,  $R=3.75\%$

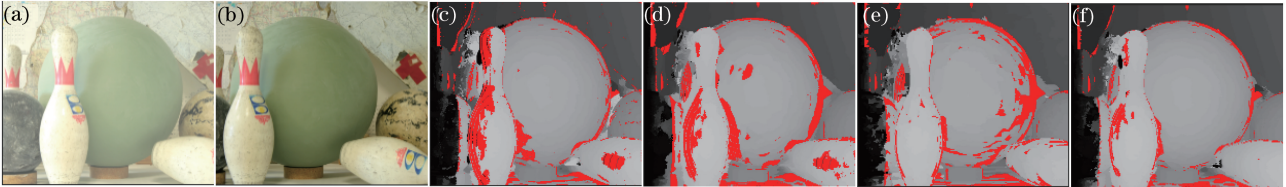


图 8 不同立体匹配算法在 Bowling2 图像对上的测试结果。(a) Bowling2 左原图；(b) Bowling2 右原图；

(c) GRD,  $R=12.91\%$ ；(d) MCT,  $R=15.77\%$ ；(e) MCT',  $R=14.21\%$ ；(f) 所提算法,  $R=9.04\%$

Fig. 8 Test results of different stereo matching algorithms on Bowling2 image pairs. (a) Bowling2 left image; (b) Bowling2 right image; (c) GRD,  $R=12.91\%$ ; (d) MCT,  $R=15.77\%$ ; (e) MCT',  $R=14.21\%$ ; (f) proposed algorithm,  $R=9.04\%$

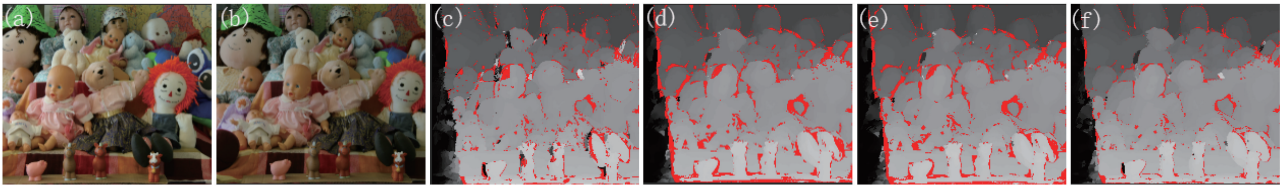


图 9 不同立体匹配算法在 Dolls 图像对上的测试结果。(a) Dolls 左原图；(b) Dolls 右原图；(c) GRD,  $R=7.66\%$ ；

(d) MCT,  $R=11.55\%$ ；(e) MCT',  $R=11.87\%$ ；(f) 所提算法,  $R=5.77\%$

Fig. 9 Test results of different stereo matching algorithms on Dolls image pairs. (a) Dolls left image; (b) Dolls right image; (c) GRD,  $R=7.66\%$ ; (d) MCT,  $R=11.55\%$ ; (e) MCT',  $R=11.87\%$ ; (f) proposed algorithm,  $R=5.77\%$

从图 6~9 可以看出,所提算法在不同光照条件下均能取得明显优于其他几种算法的匹配效果;实验中在 Aloe 图像对中的最大误匹配率达 10.19%,而所提算法仅为 6.74%;在 Baby1 立体图像对中,所提算法比误匹配率最高的 GRD 算法低了 9.07%;在 Bowling2 图像对中的最大误匹配率达 15.77%,所提算法仅为 9.04%;在 Dolls 图像对中的最大误匹配率达 11.87%,所提算法仅为 5.77%。实验结果表明,所提算法对光照变化具有一定稳健性,而且从图中还可以看出,所提算法在视差不连续区域和弱纹理区域的红色误匹配标记点明显少于其他算法。

对不同噪声和光照的对比实验分析表明,所提算法对噪声和光照等变换具有很好的稳健性,而且还可以进一步降低图像在视差不连续区域和弱纹理区域的误匹配率。

### 3.4 算法对比

为测试所提算法的总体性能,对 Tsukuba、Venus、Teddy、Cones 4 组立体图像对进行测试,结

果如图 10 所示,其中图 10(d)、(f)中红色标记区域为误匹配点。

为便于更直观地分析算法性能,选取 7 种经典算法进行比较,包括 BPcompressed<sup>[21]</sup>、GC-occ<sup>[5]</sup>、AdaptAggrDP<sup>[22]</sup>、RTCensus<sup>[23]</sup>、FastAggreg<sup>[24]</sup>、SemiGlob<sup>[25]</sup>和 GradAdaptWgt<sup>[26]</sup>,计算在非遮挡区域和所有区域的误匹配率,对比结果如表 5 所示,视差图经视差精细化处理,误差限为 1 pixel。可以看出,所提算法的平均误匹配率最低,为 4.74%,明显低于其他算法,在 Venus 上非遮挡区域误匹配率达到 0.23%,所有区域的误匹配率为 1.05%,在几种算法中最优。

为进一步测试所提算法的总体性能,对 Middlebury2.0 平台另外 27 组立体图像对进行测试,与另外 7 种立体匹配算法进行比较,通过非遮挡区域误匹配率百分比来对比算法的性能。对比算法包括 MCT<sup>[16]</sup>、MCT'<sup>[17]</sup>、MST<sup>[6]</sup>、Adaptweight (AW)<sup>[13]</sup>、GF<sup>[14]</sup>、CT-GF<sup>[15]</sup>和CT-MST<sup>[15]</sup>,实验结果如表 6 所

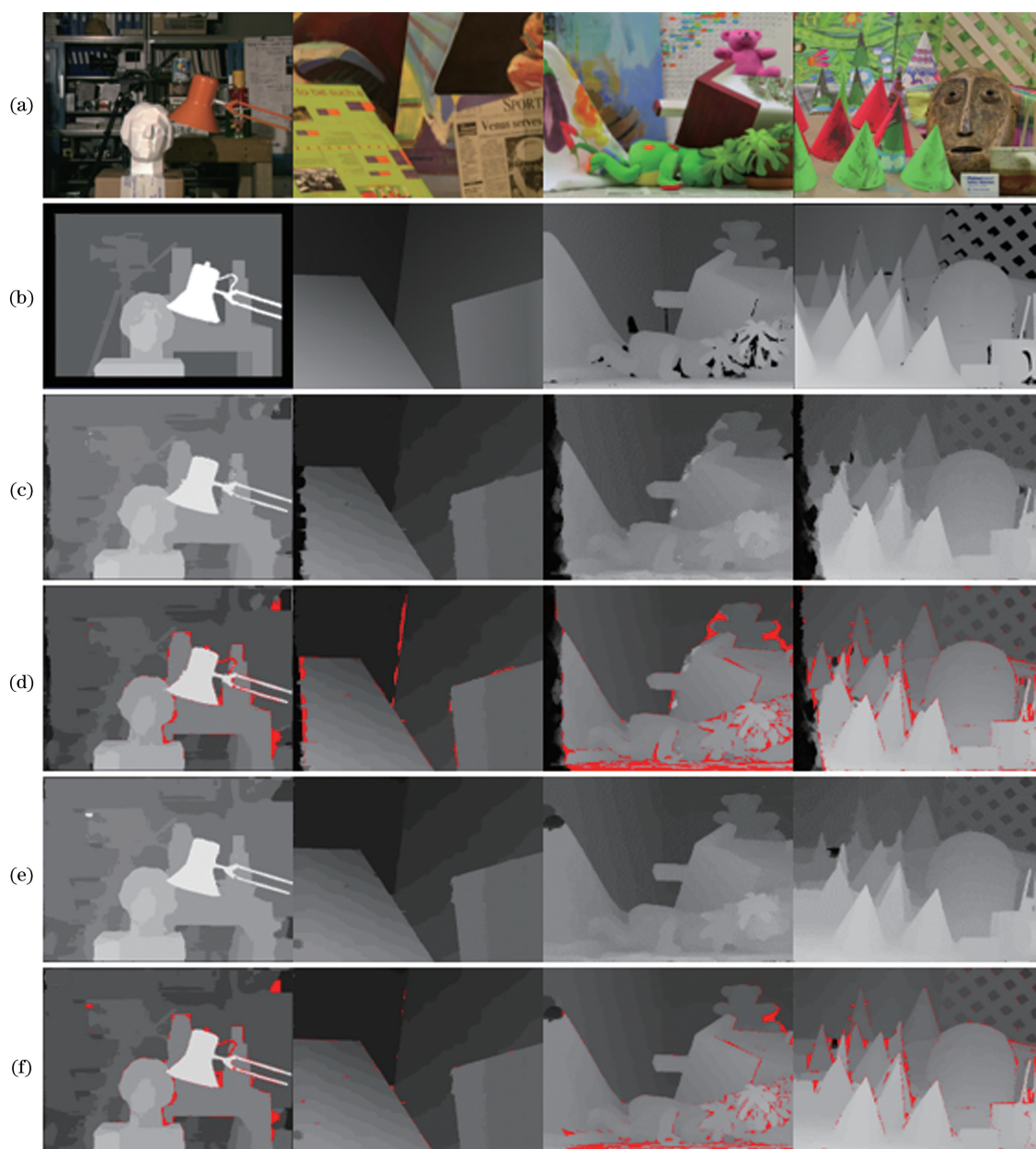


图 10 所提算法在 Middlebury2.0 图像对上的实验结果。(a)测试图左原图;(b)标准视差图;(c)所提算法计算的视差图(未精化处理);(d)误匹配标记图(未精化处理);(e)所提算法计算的视差图(精化处理);(f)误匹配标记图(精化处理)

Fig. 10 Experimental results of proposed algorithm on Middlebury2.0 image pairs. (a) Testing left image; (b) standard disparity map; (c) disparity map of proposed algorithm(without disparity refinement); (d) mismatched map (without disparity refinement); (e) disparity maps obtained by proposed algorithm (disparity refinement); (f) mismatched map (disparity refinement)

示,视差图未经视差精化,误差限为 1 pixel。

表 6 表明,所提算法在 27 组立体图像对的平均误匹配率最低,仅为 8.67%,比其他算法中最高误匹配率降低一半左右且精度最高;在 Bowling1 中最高误匹配率为 18.22%,所提算法仅为 6.07%,降低了 12.15%;在 Lampshade1 中,最高误匹配率为 25.43%,而所提算法仅为 10.04%,降低了 15.39%。实验结果表明,所提算法进一步降低了误匹配率,提高了匹配精度。

8 种算法在本文实验环境中对 Middlebury2.0 中 31 组立体图像对的测试运行平均时间如表 7 所示,其中 MST 算法的运行时间最少,AW 算法的运行时间最长;所提算法的运行时间为 6.63 s,时间较长,这是因为对图像进行多层采样处理增加了运行时间;同时所提算法在计算匹配代价与代价聚合过程中分别采用高斯权重模型,在计算复杂度上有所提升,导致处理时间增加,但是所提算法在误匹配率与稳健性上有很大的提升。



表 5 不同立体匹配算法在非遮挡区域和所有区域误匹配率

Table 5 False matching rates of different stereo matching algorithms in no-occlusion region and all regions

Algorithm	Tsukuba		Venus		Teddy		Cones		Average
	No-occluded	All	No-occluded	All	No-occluded	All	No-occluded	All	
BPcompressed <sup>[21]</sup>	2.68	3.63	1.33	1.89	8.36	13.9	3.71	9.85	5.67
GC-occ <sup>[5]</sup>	1.19	2.01	1.64	2.19	11.20	17.40	5.36	12.4	6.67
AdaptAggrDP <sup>[22]</sup>	1.57	3.50	1.53	2.69	6.79	14.3	5.53	13.2	6.14
RTCensus <sup>[23]</sup>	5.08	6.25	1.58	2.42	7.96	13.8	4.10	9.54	6.34
FastAggreg <sup>[24]</sup>	1.16	2.11	4.03	4.75	9.04	15.2	5.37	12.6	6.78
SemiGlob <sup>[25]</sup>	3.26	3.96	1.00	1.57	6.02	12.2	3.06	9.75	5.10
GradAdaptWgt <sup>[26]</sup>	2.26	2.63	6.99	1.39	8.00	13.10	2.61	7.67	5.58
Proposed algorithm	2.52	3.23	0.23	1.05	5.70	12.35	3.29	9.75	4.74

表 6 不同立体匹配算法在非遮挡区域的误匹配率

Table 6 False matching rate of different stereo matching algorithms in no-occluded region

Stereo pairs	False matching rate / %							
	MCT <sup>[16]</sup>	MCT' <sup>[17]</sup>	MST <sup>[6]</sup>	AW <sup>[13]</sup>	GF <sup>[14]</sup>	CT-GF <sup>[15]</sup>	CT-MST <sup>[15]</sup>	Proposed algorithm
Aloe	9.49	9.31	6.61	6.50	7.06	7.04	6.57	5.93
Art	18.61	19.21	13.74	12.88	12.03	12.16	13.86	10.16
Baby1	4.29	3.91	7.65	5.45	3.16	3.31	7.91	2.53
Baby2	6.02	5.98	14.12	13.64	3.97	4.03	13.67	2.66
Baby3	6.99	6.72	9.53	8.76	4.71	4.97	9.47	3.91
Books	12.26	11.34	12.97	12.65	10.21	9.96	12.47	8.01
Bowling1	18.22	17.23	19.14	16.87	6.21	6.18	19.58	6.07
Bowling2	11.15	11.16	12.63	10.43	8.18	8.47	12.58	5.88
Cloth1	1.98	1.87	1.18	1.25	1.88	1.96	1.15	1.24
Cloth2	5.79	5.21	5.40	5.56	4.20	4.83	5.36	3.43
Cloth3	4.05	4.14	3.20	3.81	2.79	2.80	3.16	2.29
Cloth4	3.58	3.64	2.53	3.21	2.25	2.38	2.40	1.64
Dolls	11.43	12.15	7.86	6.43	7.12	7.27	7.62	5.26
Flowerpots	14.27	15.21	17.27	15.44	9.63	9.67	17.05	7.53
Lampshade1	24.73	25.43	11.29	12.36	12.12	10.93	11.37	10.04
Lampshade2	27.60	27.77	14.20	17.44	12.48	10.54	14.08	11.16
Laundry	24.77	25.21	19.39	16.89	20.64	18.54	17.86	15.51
Midd1	48.75	40.14	18.96	36.47	36.95	33.45	19.21	31.76
Midd2	47.70	40.01	20.21	35.56	34.02	32.96	19.32	31.90
Moebius	14.16	13.86	11.33	13.26	10.64	10.92	11.66	9.57
Monopoly	28.49	27.97	16.85	15.43	19.74	18.99	16.64	19.50
Plastic	44.40	39.43	30.30	32.44	27.06	22.22	31.06	25.80
Reindeer	13.83	12.44	11.96	11.73	7.84	8.15	11.92	6.89
Rocks1	5.38	5.19	4.14	5.43	3.97	4.03	3.97	2.86
Rocks2	4.13	3.24	3.33	5.02	2.67	2.74	3.98	1.94
Wood1	6.49	7.17	11.17	8.99	4.93	4.93	10.98	3.72
Wood2	4.41	4.37	4.20	6.15	2.71	2.95	4.12	2.18
Average	15.67	14.79	11.52	12.59	10.34	9.87	11.45	8.67

表 7 不同算法在 31 组立体图像对中的平均运行时间

Table 7 Average running time of 31 Middlebury stereo pairs with different algorithms

s

Method	MCT <sup>[16]</sup>	MCT' <sup>[17]</sup>	MST <sup>[6]</sup>	AW <sup>[13]</sup>	GF <sup>[14]</sup>	CT-GF <sup>[15]</sup>	CT-MST <sup>[15]</sup>	Proposed
Average running time /s	3.97	4.35	1.56	15.87	4.76	6.75	4.12	6.63

## 4 结 论

提出了一种改进的局部立体匹配算法,用支持窗口的加权平均灰度值作为 Census 变换参考值,引入对不同方向归一化加权融合的梯度变换,采用 Census 代价和梯度代价加权融合的方法求得初始匹配代价,有效抑制了失真对代价计算的影响,增强了立体匹配的稳健性;利用改进引导滤波算法分别在不同分辨率尺度下进行代价聚合,充分利用了图像的细节信息,降低了图像在视差不连续区域和弱纹理区域的误匹配率。将所提算法与相关算法进行对比实验,结果表明所提算法对噪声和光照等干扰的抵抗能力优于其他非参数变换算法,具有很好的稳健性,在视差不连续区域和弱纹理区域的匹配精度进一步提高。但是,从误差标记的视差图中可以看出,所提算法在细节区域中精度还有待提高,不同区域的匹配结果相差较大。今后的研究中将针对细节区域进行进一步优化,同时需要对处理速度进行进一步优化。

### 参 考 文 献

- [1] Zhang Z Z, Wang H, Wu Q, *et al.* 3D shape reconstruction based on digital holography of stereomatching technology[J]. *Laser & Optoelectronics Progress*, 2014, 51(11): 110901.  
张赞赞, 王辉, 吴琼, 等. 基于立体匹配技术的数字全息三维形貌重构[J]. *激光与光电子学进展*, 2014, 51(11): 110901.
- [2] Scharstein D, Szeliski R, Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[C]//IEEE Workshop on Stereo and Multi-Baseline Vision, 2001: 131-140.
- [3] Veksler O. Stereo correspondence by dynamic programming on a tree[C]//Conference on Computer Vision and Pattern Recognition, 2005: 384-390.
- [4] Felzenszwalb P F, Huttenlocher D P. Efficient belief propagation for early vision[J]. *International Journal of Computer Vision*, 2006, 70(1): 41-54.
- [5] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts[C]//8<sup>th</sup> IEEE International Conference on Computer Vision, 2001: 508-515.
- [6] Yang Q. A non-local cost aggregation method for stereo matching[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2012: 1402-1409.
- [7] Yao P, Zhang H, Xue Y B, *et al.* Iterative color-depth MST cost aggregation for stereo matching[C]//IEEE International Conference on Multimedia and Expo, 2016: 16266824 .
- [8] Mei X, Sun X, Dong W, *et al.* Segment-tree based cost aggregation for stereo matching[C]//26<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition, 2013: 313-320.
- [9] Tombari F, Mattoccia S, Stefano L D, *et al.* Classification and evaluation of cost aggregation methods for stereo correspondence[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2008: 2609-2616.
- [10] Pollard S B, Mayhew J E, Frisby J P. PMF: A stereo correspondence algorithm using a disparity gradient limit[J]. *Perception*, 1985, 14(4): 449-470.
- [11] Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence[C]//European Conference on Computer Vision, 1994: 151-158.
- [12] Hirschmuller H, Scharstein D. Evaluation of stereo matching costs on images with radiometric differences[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(9): 1582-1599.
- [13] Yoon K J, Kweon I S. Adaptive support-weight approach for correspondence search[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2006, 28(4): 650-656.
- [14] 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.
- [15] Zhang K, Fang Y Q, Min D B, *et al.* Cross-scale cost aggregation for stereo matching[C]. 27<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition, 2014:1590-1597.
- [16] Chang Y C, Tsai T H, Hsu B H, *et al.* Algorithm and architecture of disparity estimation with minicensus adaptive support weight[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(6): 792-805.

- [17] Zhu S P, Yan L N, Li Z. Stereo matching algorithm based on improved census transform and dynamic programming[J]. *Acta Optica Sinica*, 2016, 36(4): 0415001.  
祝世平, 闫利那, 李政. 基于改进 Census 变换和动态规划的立体匹配算法[J]. *光学学报*, 2016, 36(4): 0415001.
- [18] Kwak J C, Park T R, Koo Y S, *et al.* Implementation of improved census transform stereo matching on a multicore processor[C] // *Multimedia and Ubiquitous Engineering*, 2013: 989-995.
- [19] Zhu S P, Li Z. A stereo matching algorithm using improved gradient and adaptive window[J]. *Acta Optica Sinica*, 2015, 35(1): 0110003.  
祝世平, 李政. 基于改进梯度和自适应窗口的立体匹配算法[J]. *光学学报*, 2015, 35(1): 0110003.
- [20] Scharstein D, Szeliski R. The middle stereo visionpage [OE]. (2017-08-24). <http://vision.middlebury.edu>.
- [21] Yu T, Lin R S, Super B, *et al.* Efficient message representations for belief propagation[C] // 11<sup>th</sup> IEEE International Conference on Computer Vision, 2007: 604-611.
- [22] Wang L, Yang R, Gong M, *et al.* Real-time stereo using approximated joint bilateral filtering and dynamic programming[J]. *Journal of Real-Time Image Processing*, 2014, 9(3): 447-461.
- [23] Humenberger M, Zinner C, Weber M, *et al.* A fast stereo matching algorithm suitable for embedded real-time systems[J]. *Computer Vision & Image Understanding*, 2010, 114(11): 1180-1202.
- [24] Tombari F, Mattoccia S, Stefano L D, *et al.* Near real-time stereo based on effective cost aggregation[C] // 19<sup>th</sup> International Conference on Pattern Recognition, 2009: 1044-1051.
- [25] Hirschmuller H. Accurate and efficient stereo processing by semi-global matching and mutual information[C] // *Conference on Computer Vision and Pattern Recognition*, 2005: 807-814.
- [26] De-maeztu L, Villanueva A. Stereo matching using gradient similarity and locally adaptive support-weight[J]. *Pattern Recognition Letters*, 2011, 32(13): 1643-1651.