

# 基于改进 Census 变换和动态规划的立体匹配算法

针对什么问题, 提出什么算法(题目)。

祝世平 闫利那 李 政

按立体匹配的流程, 从操作层面描述算法

北京航空航天大学仪器科学与光电工程学院, 北京 100191

**摘要** 为降低双目立体匹配算法在视差不连续区域和噪声干扰情况下的误匹配率, 提出了一种基于改进 Census 变换和动态规划的立体匹配算法。采用支持区域为十字交叉形状窗口且设有噪声容限的改进 Census 变换进行代价计算, 提高了单像素匹配代价的可靠性; 利用引导图滤波器快速有效地完成代价聚合; 在视差选择阶段, 设计了一种改进的动态规划算法, 消除了扫描线效应, 提高了匹配速度和正确率; 经过视差后处理得到最终视差图。实验结果表明, 该算法在 Middlebury 测试平台上的平均误匹配率为 5.31%, 在低纹理区域和视差不连续区域均能得到准确的视差, 运算复杂度低且具有较好的稳健性。

**关键词** 机器视觉; 立体匹配; Census 变换; 动态规划; 引导图滤波; 视差

中图分类号 TN911.73

文献标识码 A

doi: 10.3788/AOS201636.0415001

用事实表明该算法的有效性

## Stereo Matching Algorithm Based on Improved Census Transform and Dynamic Programming

Zhu Shiping Yan Lina Li Zheng

School of Instrumentation Science and Optoelectronics Engineering, Beihang University, Beijing 100191, China

**Abstract** In order to reduce the mismatching rate of binocular stereo matching algorithm in the disparity discontinuity region and under noise disturbance, a stereo matching algorithm based on improved Census transform and dynamic programming is proposed. An improved Census transform with a noise margin is applied to compute the cost based on a cross shape support region. The reliability of single pixel matching cost is enhanced. The guided image filter is used to aggregate the cost volume fast and efficiently. In the disparity selecting step, an improved dynamic programming algorithm is designed to eliminate the scan-line effect and improve the matching speed and accuracy. The final disparity maps are gained after post-processing. The experimental results demonstrate that the proposed algorithm evaluated on the Middlebury benchmark achieves an average error rate of 5.31%, and the accurate disparity can be obtained in both low texture and disparity discontinuity regions with low computing complexity and strong robustness.

**Key words** machine vision; stereo matching; Census transform; dynamic programming; guided image filtering; disparity

**OCIS codes** 150.1135; 100.6890; 330.1400

## 1 引言

立体匹配是什么, 重要性, 应用

立体匹配是寻找双目或多目图像中的对应点进而估计场景深度信息的过程, 是计算机视觉的基础和热点问题之一。立体匹配技术可以获得较为准确的物体三维深度信息, 在机器人导航、虚拟现实及三维重构等领域得到广泛应用<sup>[1-2]</sup>。现有的立体匹配算法, 大致可分为全局算法和局部算法。全局立体匹配算法将立体匹配看作一个能量最小化问题, 主要有动态规划(DP)<sup>[3]</sup>、图割法<sup>[4]</sup>、置信传播<sup>[5]</sup>等算法。局部算法利用基于窗口的代价聚合计算得到单像素的视差, 虽然精度不如全局算法, 但结构简单、计算复杂度低、效率高, 在实际

立体匹配算法分类(局部和全局), 简单介绍全局算法的分类, 并指出局部算法相对于全局的优点, 为下文着重介绍作铺垫

收稿日期: 2015-11-02; 收到修改稿日期: 2015-12-01

基金项目: 国家自然科学基金(61375025, 61075011, 60675018)

作者简介: 祝世平(1970—), 男, 博士, 副教授, 主要从事图像处理和视频压缩等方面的研究。E-mail: spzhu@163.com

系统中易于实现。

传统算法代价算法的分类

其一个缺陷(光照的敏感)

今年来提出的匹配代价,并故意把Census的放到后面,以引出后文

传统的局部立体匹配代价计算方法有基于绝对差(AD)、平方差(SD)<sup>[6]</sup>、自适应权重(AW)<sup>[7]</sup>以及采样不敏感的BT(Birchfield-Tomasi)<sup>[8]</sup>等几种主要算法。当光照及辐射度失真时,基于图像的颜色和几何信息算法<sup>[9-10]</sup>的匹配精度会急剧下降。近年来,研究者提出了许多新的匹配代价。Hirschmuller等<sup>[11]</sup>利用图像熵提出了一种基于互信息的匹配代价计算方法,De-Maeztu等<sup>[12]</sup>改进了基于梯度的匹配代价,Zabih等<sup>[13]</sup>提出基于非参数变换的方法:Rank和Census变换。Lee等<sup>[14]</sup>提出了三状态Census变换。非参数变换算法提高了匹配精度和稳健性,但传统Census变换依赖于中心像素灰度值大小,对噪声比较敏感,支持窗口是以待匹配像素为中心的一个固定的矩形区域,支持区域的选择过大或过小也会影响匹配精度。

非参数变换算法的缺点

局部算法的缺陷

在视差选择阶段,选择局部算法使匹配代价最小的候选点作为最终匹配点,通常用胜者为王算法(WTA)方法<sup>[15]</sup>,该方法快速有效,但其基于局部贪心策略并且只利用了当前像素的代价信息,没有考虑相邻像素的视差限制,容易受噪声等因素的影响产生视差突变,使获得的初始视差图存在较多异常值和误匹配。解决这个问题,研究引入DP算法进行视差选择<sup>[15-16]</sup>。传统的DP算法进行视差选择的时间复杂度为 $O(ND^2)$ ,其中 $N$ 为图像大小, $D$ 为视差搜索范围。当视差范围 $D$ 较大时,该算法运算复杂,且会产生扫描线效应。

针对局部缺陷,引入DP算法,但又指出其缺陷(慢、扫描线效应)

Census变换依赖于中心像素,对噪声敏感,受支持窗口大小的影响,DP算法复杂度较高且会产生扫描线效应。基于此,本文提出了一种新的局部立体匹配算法,主要贡献为:对Census变换进行了改进,采用一种设有噪声容限及采用交叉形状窗口的三状态Census变换计算代价,增强了稳健性;在视差选择阶段采用WTA-DP相结合的动态规划进行视差选择,消除了传统动态规划固有的扫描线效应,并修改了动态搜索半径,降低了运算复杂度。

总结Census和Dp算法的缺点,来引出针对此缺点本文的贡献:

## 2 算法描述

根据Scharstein等<sup>[6]</sup>提出的立体匹配算法的分类和评价,立体匹配算法通常包括以下4个步骤:1)匹配代价计算;2)代价聚合;3)视差计算与优化;4)视差提精。本文算法也遵循该流程。

### 2.1 基于改进Census变换计算匹配代价

传统Census变换的基本原理是以一个矩形窗口遍历图像,通常选取中心像素的灰度值作为参考值,将矩形窗口中各像素的灰度值与参考值进行比较,并用0和1来表示其大小关系。Census变换的实质是将图像灰度值编码成二进制码流,来表示周围邻域像素相对于中心像素的大小关系。变换过程可表达为

$$T(p) = I_{\text{census}}(p) = \bigotimes_{q \in N_p} \xi[I(p), I(q)], \quad (1)$$

式中 $N_p$ 表示 $p$ 的邻域, $I(p)$ 、 $I(q)$ 为 $p$ 、 $q$ 对应的灰度值, $\bigotimes$ 符号表示按位连接, $\xi[I(p), I(q)]$ 满足

$$\xi[I(p), I(q)] = \begin{cases} 0, & I(q) \leq I(p) \\ 1, & I(q) > I(p) \end{cases}. \quad (2)$$

由于传统Census变换比较依赖于中心像素的大小且对图像噪声较为敏感,为了减小由于图像噪声引起的误匹配,使算法具有更好的稳健性,提出一种改进的Census变换来计算匹配代价。其核心思想是,首先取窗口内各像素灰度值的平均值 $\bar{I}(p)$ 来代替中心像素的灰度值,并作为参考值,再将邻域窗口中各像素的灰度值 $I(q)$ 与参考值进行比较。为了增强稳健性,为参考值增加一个噪声容限 $\alpha$ ,并采用两个二进制位来表达三种状态,定义为

$$\xi[I(p), I(q)] = \begin{cases} 10, & I(q) > \bar{I}(p) + \alpha \\ 01, & I(q) < \bar{I}(p) - \alpha \\ 00, & \text{else} \end{cases}. \quad (3)$$

这个思路很厉害,噪音可能对真值造成一个上下的波动,所以给了一个缓冲区间。但另一方面,对于低纹理区域,此方法对于差异的检测是否会过于迟钝?

由摄像机畸变引起的图像噪声与像素灰度值相关,其大小正比于灰度值,故噪声阈值为

$$\alpha = \left\lceil \frac{I(p)}{\beta} \right\rceil, \quad (4)$$

式中 $\lceil \cdot \rceil$ 表示就近取整运算符, $\beta$ 为常数,在合成图像和真实图像中 $\beta$ 取值分别为500和50。

影响 Census 变换的另一个重要因素是窗口大小。窗口过小, 包含的空间信息太少; 窗口越大, 包含的遮挡像素越多, 误差也越大。为了解决这一矛盾, 采用 Zhang 等<sup>[17]</sup>提出的十字交叉形状的变换窗口, 在尽量获取更多空间信息的同时减少遮挡区域的影响。对于当前待匹配像素  $p$ , 在水平方向和垂直方向分别扩展其邻域像素构成一个十字交叉区域, 作为支持窗口的初始骨架, 分别用  $H(p)$  和  $V(p)$  表示。区域由 4 个方向的臂长  $\{h_p^-, h_p^+, v_p^-, v_p^+\}$  确定, 并随着图像的局部色彩信息自适应地改变。以  $h_p^-$  为例, 臂长的判别准则为

$$D_c(p_i, p) < \tau, \quad (5)$$

$$D_s(p_i, p) < L, \quad (6)$$

式 (5) 中  $\tau$  和  $L$  为预设的色彩阈值和距离阈值, 设  $\tau = 20$ ,  $L = 17$ 。  $D_s(p_i, p)$  为像素  $p_i$  和  $p$  的空间距离差,  $D_c(p_i, p)$  为色彩差, 定义为

$$D_c(p_i, p) = \max_{c \in \{R, G, B\}} |I_c(p_i) - I_c(p)|. \quad (7)$$

利用上述判定准则可得到像素  $p$  的臂长  $\{h_p^-, h_p^+, v_p^-, v_p^+\}$  大小, 最坏的情况下, 该步骤需要  $4L_{\max}N$  次比较操作,  $L_{\max}$  为臂长  $\{h_p^-, h_p^+, v_p^-, v_p^+\}$  中的最大值, 因此十字交叉窗口的时间复杂度为  $O(L_{\max}N)$ 。

由 Census 变换定义的匹配代价  $C_{\text{census}}(p, d)$  表示参考图像中点  $p(x, y)$  和匹配图像中对应于视差  $d$  的点  $p_d = (x - d, y)$  之间的相似性度量, 通过 Hamming 距离计算得到, 即两个二进制数之间不相同的位数个数为

$$C_{\text{census}}(p, d) = \text{Hamming}[T(p), T(p_d)] = T(p) \oplus T(p_d). \quad (8)$$

图 1 为传统 Census 变换和改进 Census 变换在交叉窗口下的工作过程。如图 1(a) 所示, 在平坦区域, 相邻像素的灰度值近似, 传统 Census 变换得到的二进制序列波动较大。如图 1(b) 所示, 改进 Census 变换能有效地抵抗噪声的影响, 其变换结果保持一致, 表现出很好的稳健性, 噪声阈值  $\alpha = 2$ 。

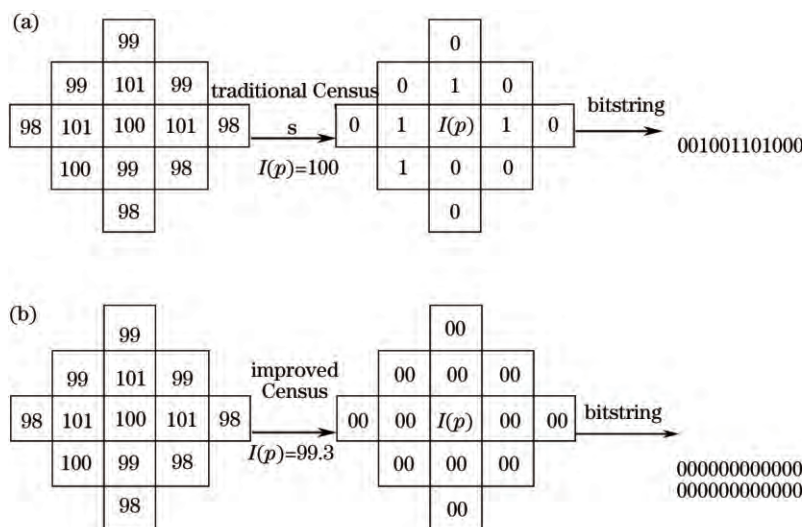


图 1 (a) 传统 Census 变换与 (b) 改进 Census 变换的比较

Fig.1 Comparison between (a) traditional Census transform and (b) improved Census transform

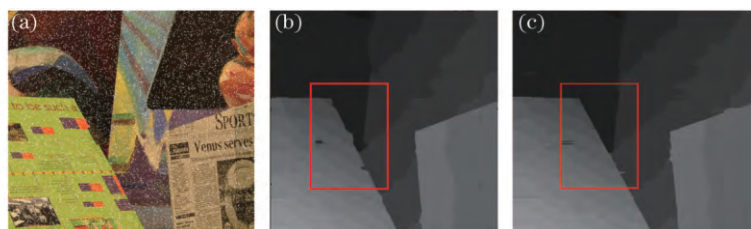


图 2 传统 Census 变换和改进 Census 变换得到的视差图。(a) 加入椒盐噪声的 Venus 图像; (b) 传统 Census 变换得到的视差图; (c) 改进 Census 变换得到的视差图

Fig.2 Disparity maps computed by traditional Census algorithm and improved Census algorithm. (a) Venus image added with salt and pepper noise; (b) disparity map computed by traditional Census algorithm; (c) disparity map computed by the improved Census algorithm



如图 2(b)所示,在视差不连续和平坦区域均存在较多误匹配点,得到的视差图比较粗糙。图 2(c)所示为改进 Census 匹配代价得到的视差图,比图 2(b)平滑了很多(红色方框内),且在视差不连续区域的效果要明显好于图 2(b)。因此,改进后的 Census 匹配代价计算方法具有更好的稳健性,匹配结果更加精确。

## 2.2 自适应代价聚合

2.1 节定义了基于单个像素的原始匹配代价,但由于单个像素的匹配代价鉴别性不高,易受噪声影响,需要利用周围像素的信息进行代价聚合,以提高视差区分性。采用引导图滤波技术对代价空间进行滤波,即用引导图滤波器产生的 AW 完成代价聚合<sup>[18-19]</sup>,滤波器核函数可表示为

$$W_{ij}(I) = \frac{1}{|w|^2} \sum_{k \in w_i \cap w_j} \left\{ 1 + \frac{[I(i) - \mu_k][I(j) - \mu_k]}{\sigma_k^2 + \varepsilon} \right\}, \quad (9)$$

式中  $i, j$  为输入图像  $I$  的像素索引,  $\mu_k$  和  $\sigma_k^2$  分别为  $I$  在  $w_k$  内的均值和方差,  $|w|$  为窗口内像素个数,  $\varepsilon$  为控制参数。

相比于双边滤波器的计算复杂度  $O(NW)$ ,其中  $W$  为滤波窗口的大小,  $N$  为图像的大小,引导图滤波器的计算复杂度仅为  $O(N)$ 。计算复杂度仅与图像大小相关,与滤波窗口大小无关,能满足实时性要求。利用引导图滤波进行代价聚合的详细过程可见文献[18]。引导图像  $I$  为待匹配的原始参考图像,滤波器输入  $p$  为每一层视差空间图像  $DSI(:, :, d)$ ,滤波后的输出  $q$  即为聚合后的匹配代价。图 3 所示为代价图滤波后的结果。图 3 为  $d=5$  和  $d=10$  的视差代价图实验结果。图 3(a)、(c)为聚合前的代价图,图 3(b)、(d)为经过引导图滤波之后的代价图。可以看出,受噪声干扰的原始代价图经过滤波之后,更加平滑可靠,可用于下一步的视差选择。

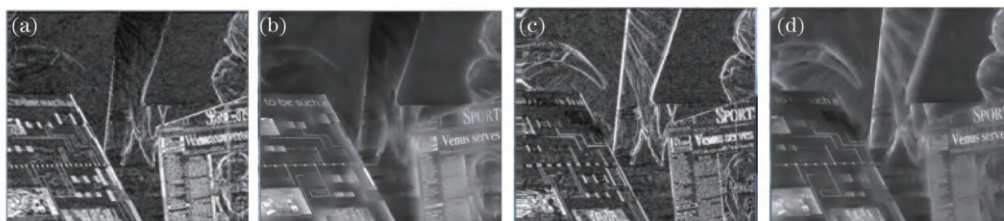


图 3 代价滤波的结果。(a) 视差值  $d=5$  时的代价图;(b)  $d=5$  时经过引导图滤波后的代价图;(c) 视差值  $d=10$  时的代价图;(d)  $d=10$  时经过引导图滤波后的代价图

Fig.3 Results of cost filtering. (a) Cost map aggregation when  $d=5$ ; (b) cost map after guided image filtering when  $d=5$ ;

(c) cost map aggregation when  $d=10$ ; (d) cost map after guided image filtering when  $d=10$

## 2.3 视差选择

传统局部算法大多简单地使用 WTA 法进行视差选择,获得初始视差值为

$$d_0(p) = \arg \min_d [C(p, d)], \quad (10)$$

式中  $C(p, d)$  为视差值为  $d$  时的匹配代价值。WTA 方法虽然快速有效,但没有考虑相邻像素的视差限制,容易受噪声等影响而产生视差突变,因此引入动态规划的全局思想进行视差选择。对于参考图像中的一条扫描线  $y$ ,从左至右遍历每一个像素点  $p(x, y)$ ,计算其对应于所有视差值  $d$  的能量  $M(x, y, d)$ ,并保存最小能量值及对应的视差位置,可表达为

$$M(x, y, d) = C(x, y, d) + \min_{d' \in [0, d_{\max}]} [M(x-1, y, d') + \lambda |d - d'|], \quad (11)$$

式中  $C(x, y, d)$  为像素点  $(x, y)$  在视差值为  $d$  时的代价值,为能量方程的数据项,  $d'$  为  $p$  的相邻像素  $p'(x-1, y)$  的视差取值,  $\lambda$  为参数常量。

最后一列的能量计算完之后,可通过反向追踪法确定最优路径,得到该行扫描线上所有像素的最优视差配置。对每条扫描线重复以上过程,可确定整幅图像的全部视差值。该方法的运算复杂度为  $O(WD^2)$ ,其中  $W$  为图像宽,  $D$  为视差范围,不适用于实时应用。为降低复杂度,对传统动态规划算法进行了改进,设计了一种基于 WTA 引导的 DP 算法,在运算效率和匹配精度上均能取得较好效果。考虑视差平滑性假设,认为  $p$  的相邻像素视差值与其相近,将  $d$  的取值限制为  $\{d-1, d, d+1\}$ ,得到修改后的公式为

$$M(x, y, d) = C(x, y, d) + \min_{d' \in [d-1, d, d+1]} [M(x-1, y, d') + \lambda |d - d'|]. \quad (12)$$

修改后的算法复杂度降为  $O(WD)$ , 但简化后的算法计算得到的视差在深度不连续区域变化过慢, 会模糊深度边缘, 产生扫描线效应。为了避免出现过度平滑现象, 利用 WTA 法得到的初始视差图来引导上述动态规划过程, 核心是利用 WTA 法为动态规划过程提供一个额外的视差候选值。利用(10)式计算得到初始视差图  $d_0$ , 对于像素  $p(x, y)$ , 用  $d_0(x-1, y)$  作为  $p'$  的第 4 个视差候选值, 即有

$$M(x, y, d) = C(x, y, d) + \min_{d' \in [d-1, d, d+1, d_0(x-1, y)]} [M(x-1, y, d') + \lambda |d - d'|], \quad (13)$$

当  $d' \in [d-1, d, d+1]$  时,  $\lambda = 1$ ,  $d' = d_0$  时,  $\lambda = 0.02$ 。由于初始视差值  $d_0$  在开始时已计算好, 因而改进后的动态规划算法运算复杂度并未增加, 仍为  $O(WD)$ 。图 4 为标准的 Venus 图像利用改进后的 WTA-DP 算法获得的视差图, 并与 WTA 和原始 DP 方法相比较。改进方法获得的视差图质量优于单独使用 WTA 或 DP 方法。改进方法将 WTA 和 DP 的优点相结合, 匹配结果较 WTA 法更平滑, 同时又能很好地保存视差边缘和估计遮挡点视差值。

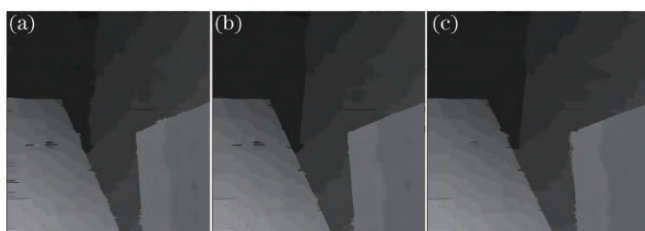


图 4 不同的方法进行视差选择得到的视差图。(a) WTA 法; (b) DP 法; (c) 改进的 WTA-DP 法

Fig.4 Disparity maps selected by different methods. (a) WTA method; (b) DP method; (c) improved WTA-DP method

## 2.4 视差后处理

初始视差图中还存在一些误匹配, 需要进行后处理, 如图 5 所示。首先采用一种区域投票的策略, 在一个局部支持区域内统计视差直方图, 并选择出现频率最大的视差值作为该点的最终视差, 提高了匹配精度。支持区域将沿用代价计算阶段的十字交叉窗口自适应区域。采用左右一致性检验来检测遮挡点。设  $d_L(p)$ 、 $d_R(p)$  分别为左视差值和右视差值, 当  $p$  点和其对应点的视差不一致时, 即  $d_L(p) \neq d_R[p - d_L(p)]$ , 则认为  $p$  是遮挡点, 并将其视差值标记为 0。在  $p$  点所属行扫描线上分别寻找其左、右第一个非遮挡点, 选择较小的视差值作为遮挡点的视差, 从而完成遮挡点的填补。最后, 采用加权中值滤波器对视差图进行平滑滤波, 获得最终的视差图。

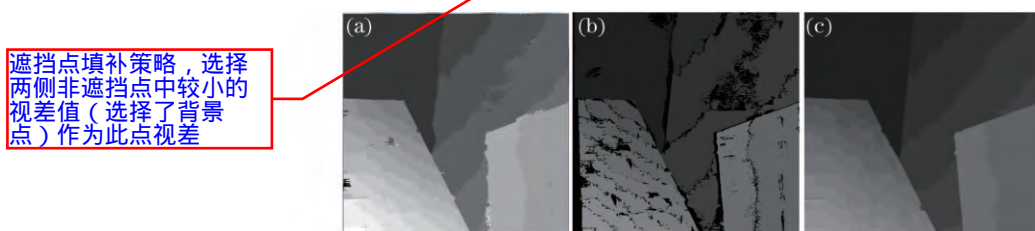


图 5 视差后处理的结果。(a) 区域投票后的视差图; (b) 左右一致性检测; (c) 得到的最终视差图

Fig.5 Effect of disparity post-processing. (a) Disparity map after region voting; (b) left-right consistence check; (c) obtained final disparity map

## 3 实验结果

### 3.1 参数变化的稳健性

为了验证改变参数设置时算法的稳健性, 设置主要的参数为代价计算和视差后处理阶段自适应区域的颜色阈值  $\tau$ 、距离阈值  $L$ 。将 Venus 作为测试图像, 用所得视差图与真实视差图的峰值信噪比(PSNR)

$$R_{\text{PSN}} = 10 \lg \left\{ \frac{255^2 MN}{\sum_{x,y \in (M,N)} [f(x,y) - g(x,y)]^2} \right\}, \quad (14)$$

作为匹配结果的评价标准,式中  $M, N$  为图像的大小,  $f(x, y)$  为改进算法得到的  $(x, y)$  点的像素值,  $g(x, y)$  为对应的真实视差图中的像素值。如图 6(a)所示,令  $\tau = 20$ , 改变  $L$  值, 当距离阈值  $L > 17$  时, 视差图像的 PSNR 值基本上不再变化, 趋于稳定。如图 6(b)所示, 当  $L = 17$ ,  $\tau$  改变时, 视差图的 PSNR 基本保持不变, 具有较好的稳健性。

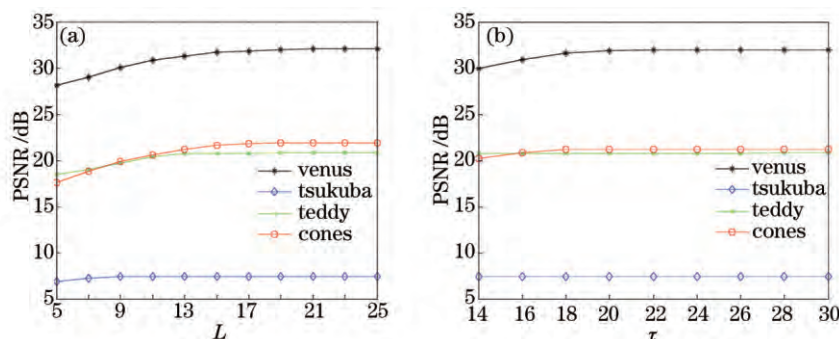


图 6 算法参数的稳健性实验。(a)  $\tau = 20$ ,  $5 \leq L \leq 25$ ; (b)  $L = 17$ ,  $14 \leq \tau \leq 30$

Fig.6 Test of parameter robustness. (a)  $\tau = 20$ ,  $5 \leq L \leq 25$ ; (b)  $L = 17$ ,  $14 \leq \tau \leq 30$

### 3.2 Middlebury 标准图像测试

为验证算法的有效性,用 C 语言编程实现算法,硬件平台中央处理器(CPU)为 Intel(R) Core(TM) i7-4790,主频为 3.60 GHz。采用由 Middlebury 立体匹配算法测试平台提供的标准立体图像对对算法进行了实验和评测<sup>[20]</sup>。该网站提供了 4 组基准彩色图像对 Tsukuba、Venus、Teddy 和 Cones,其分辨率分别为  $384 \text{ pixel} \times 288 \text{ pixel}$ 、 $434 \text{ pixel} \times 383 \text{ pixel}$ 、 $450 \text{ pixel} \times 375 \text{ pixel}$  和  $450 \text{ pixel} \times 375 \text{ pixel}$ 。实验时,4 组图像的视差搜索范围依次分别为  $0 \sim 15 \text{ pixel}$ 、 $0 \sim 19 \text{ pixel}$ 、 $0 \sim 59 \text{ pixel}$  和  $0 \sim 59 \text{ pixel}$ 。根据 Middlebury 平台对测试图像的要求,算法在对所有图像对的测试过程中,除了视差搜索范围不相同外,所有参数均保持相同,具体为  $\{\beta, \tau, L, \varepsilon, W\} = \{50, 20, 17, 0.0001, 9 \times 9\}$ 。将实验结果与真实视差图比较可得到量化的匹配误差,从而客观地评价算法精度。

图 7 为改进算法的测试结果精度,从上到下依次为图像 Tsukuba、Venus、Teddy 和 Cones。图 7(d)为改进算法的误匹配像素图,图中大片白色区域是匹配正确的点,灰色区域和黑色区域分别代表遮挡区域和无遮挡区域的误匹配点。

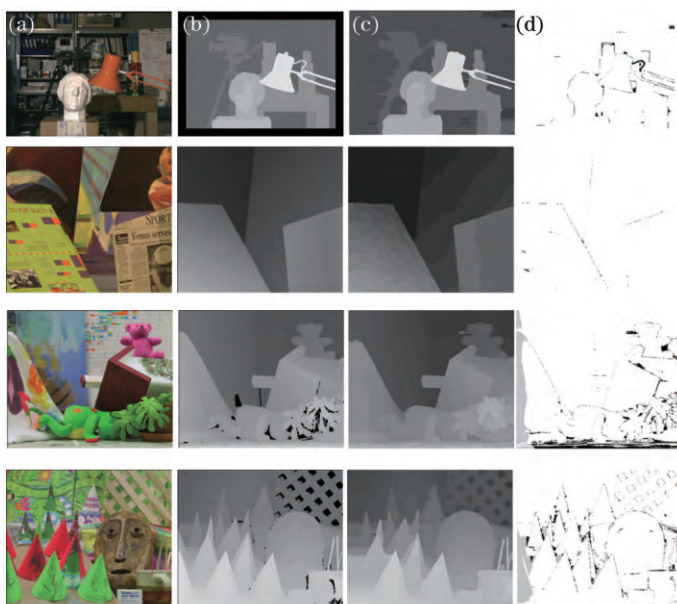


图 7 Middlebury 测试数据库实验结果。(a) 参考图像; (b) 真实视差图; (c) 改进算法得到的视差图; (d) 误匹配像素图

Fig.7 Experimental results of the Middlebury benchmark images. (a) Reference images; (b) real disparity; (c) disparity maps obtained by the improved algorithm; (d) error matched pixel map



为了更加客观地评价改进算法性能,计算了不同区域的误匹配像素百分比  $B$  为

$$B = \frac{1}{N} \sum_{(x,y)} [ |d_c(x,y) - d_r(x,y)| > \delta_d ], \quad (15)$$

式中  $d_c(x,y)$  是点  $(x,y)$  在改进算法得到视差图中的视差值,  $d_r(x,y)$  是其在真实视差图中的视差值,  $N$  代表图像的大小,误差限  $\delta_d = 1$ , 即当匹配结果与真实视差图相差 1 pixel 以上时,认为该点是误匹配点。

改进算法和其他较好算法的误匹配像素比如表 1 所示, Noocc 表示非遮挡区域误匹配像素比, All 表示总误匹配像素比, Disc 表示深度不连续区域误匹配像素比。改进算法在匹配精度上高于局部自适应权重算法, 例如 AsswGrad<sup>[21]</sup>, VariableCross<sup>[17]</sup>, CostFilter<sup>[19]</sup> 和 AdpDP<sup>[22]</sup>, 以及全局算法 GlobalGCP<sup>[23]</sup>, 在视差不连续区域和遮挡区域的像素平均误匹配百分比均低于其他几种方法。其中 CostFilter 是当前最好的局部匹配算法之一, 采用的也是基于引导图滤波的代价聚合方式。

表 1 阈值为 1 时的误匹配像素平均百分比(%)

Table 1 Average percentage of bad pixels with disparity error threshold is 1 (%)

Algorithm	Tsukuba			Venus			Teddy			Cones			Average error
	Noocc	All	Disc	Noocc	All	Disc	Noocc	All	Disc	Noocc	All	Disc	
Proposed	1.41	1.74	6.86	0.27	0.38	3.04	5.94	11.4	14.9	2.64	7.86	7.37	5.31
CostFilter	1.51	1.85	7.61	0.20	0.39	2.42	6.16	11.8	16.0	2.71	8.24	7.66	5.55
GlobalGCP	0.87	2.54	4.69	0.46	0.53	2.22	6.44	11.5	16.2	3.59	9.49	8.9	5.60
Assw-Grad	1.57	2.00	7.32	0.89	1.00	3.18	7.20	12.4	16.1	3.68	9.18	8.62	6.10
GradAdpWgt	2.26	2.63	8.99	0.99	1.39	4.92	8.00	13.1	18.6	2.61	7.67	7.43	6.55
VariableCross	1.99	2.65	6.77	0.62	0.96	3.20	9.75	15.1	18.2	6.28	12.7	12.9	7.60
AdpDP	2.05	4.22	10.6	1.92	2.98	20.3	7.32	14.4	17.6	6.41	13.7	16.5	9.82

图 8 为改进算法和其他算法在 4 幅测试图像对上的运行时间。由于采用引导图滤波代替了双边滤波, 且改进后的动态规划时间复杂度降为  $O(WD)$ , 缩小了视差搜索范围, 使得算法计算效率得到较大改善<sup>[7]</sup>, 改进算法在 4 幅测试图像上的平均运行时间分别为 5.1、10.5、29.8、29.2 s。与 Ass-Grad、自适应权重算法相比较, 改进算法能够有效减少匹配时间, 但还未能实时处理, 今后可进一步研究利用图形处理器(GPU)并行运算提高运行速度。

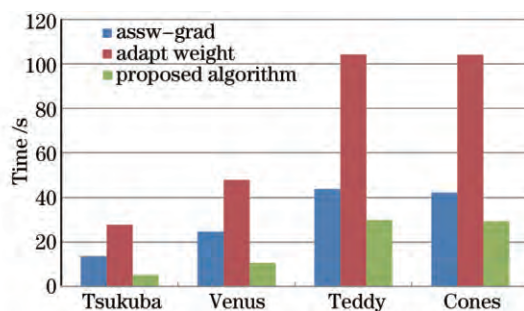


图 8 各算法运行时间

Fig.8 Running time of algorithms

为了进一步验证本文算法的优越性,对 Middlebury 平台提供的另外 5 幅标准图像进行了测试,5 组标准图像对依次为 Sawtooth、Book、Mobius、Wood、Dolls,其分辨率分别为 434 pixel × 380 pixel、463 pixel × 370 pixel、463 pixel × 370 pixel、457 pixel × 370 pixel 和 463 pixel × 370 pixel。实验时,5 组图像的视差搜索范围依次分别为 0~19 pixel、0~80 pixel、0~80 pixel、0~59 pixel 和 0~80 pixel,其他参数保持不变,实验结果如图 9 所示,从上到下依次为图像 Sawtooth、Book、Mobius、Wood、Dolls。通过实验结果与真实视差图的比较,可以看出改进算法得到的视差图整体比较光滑,在视差不连续区域和平坦区域可以取得较好的效果。但是视差图中仍然存在一些误匹配点,需进一步改进。

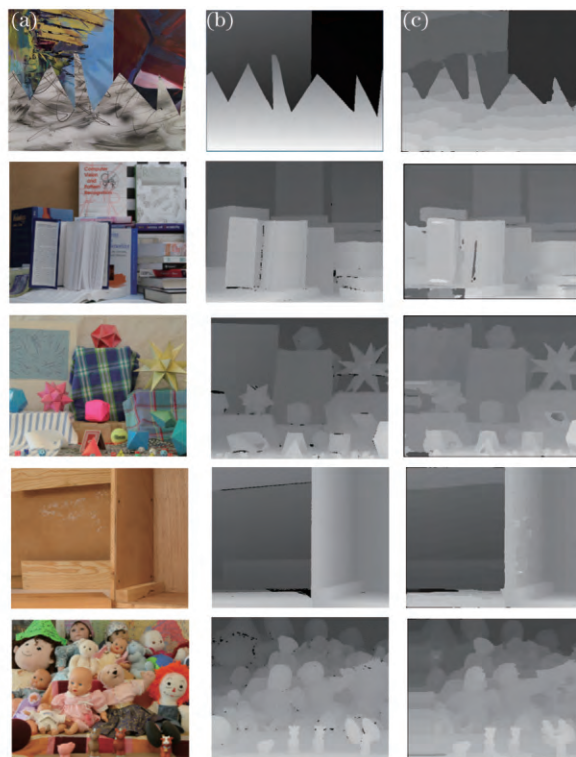


图9 Middlebury标准图像实验结果。(a)左参考图像;(b)真实视差图;(c)改进算法得到的视差图

Fig.9 Results of the Middlebury benchmark images. (a) Left reference images;  
(b) real disparity; (c) disparity obtained by of the proposed method

## 4 结 论

提出了一种基于改进 Census 变换和动态规划的立体匹配算法,在均值基础上增设了噪声容限的三状态 Census 变换进行代价计算,提高了算法的稳健性,减小了误匹配率。在视差选择阶段,利用改进后的动态规划算法,在计算得到的视差图中消除了扫描线效应,减小了搜索半径,降低了算法的运算复杂度。

实验结果表明,在 Middlebury 测试平台上该算法的平均误匹配率为 5.31%,匹配精度优于当前部分优秀的局部算法和全局算法。目前算法主要在 CPU 上实现,在今后的研究中可利用 GPU 并行运算提高算法效率,满足实时性。

## 参 考 文 献

- 1 Zhang Zanzan, Wang Hui, Wu Qiong, *et al.*. 3D shape reconstruction based on digital holography of stereo matching 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 Gu Cheng, Qian Weixian, Chen Qian, *et al.*. Rapid head detection method based on binocular stereo vision[J]. *Chinese J Lasers*, 2014, 41(1): 0108001.  
顾 骋,钱惟贤,陈 钱,等.基于双目立体视觉的快速人头检测方法[J].*中国激光*, 2014, 41(1): 0108001.
- 3 Jae C, Kyoung M, Byoung T. A dense stereo matching using two-pass dynamic programming with generalized ground control points[C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, 9(2): 1075-1082.
- 4 Bleyer M, Gelautz M. Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions[J]. *Signal Processing: Image Communication*, 2007, 22(2): 127-143.
- 5 Besse F, Rother C, Fitzgibbon A, *et al.*. PMBP: Patch match belief propagation for correspondence field estimation[J]. *International Journal of Computer Vision*, 2014, 110(1): 2-13.
- 6 Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. *International Journal of Computer Vision*, 2002, 47(1): 7-42.



- 7 Yoon K J, Kweon I S. Locally adaptive support weight approach for visual correspondence search[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 2: 924-931.
- 8 Birchfield S, Tomasi C. Depth discontinuities by pixel-to-pixel stereo[J]. International Journal of Computer Vision, 1999, 35(3): 269-293.
- 9 Gong Wenbiao, Gu Guohua, Qian Weixian, *et al.*. Stereo matching algorithm based on the inter color correlation and adaptive support weight[J]. Chinese J Lasers, 2014, 41(8): 0812001.  
龚文彪, 顾国华, 钱惟贤, 等. 基于颜色内相关和自适应支撑权重的立体匹配算法[J]. 中国激光, 2014, 41(8): 0812001.
- 10 Xie Zexiao, Lu Wenjuan. A binocular matching algorithm based on similar image geometric features[J]. Chinese J Lasers, 2014, 41(5): 0508002.  
解则晓, 陆文娟. 基于图像相似几何特征的双目匹配算法[J]. 中国激光, 2014, 41(5): 0508002.
- 11 Hirschmuller H. Accurate and efficient stereo processing by semi-global matching and mutual information[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 2: 807-814.
- 12 De-Maeztu L, Villanueva A, Cabeza R. Stereo matching using gradient similarity and locally adaptive support-weight[J]. Pattern Recognition Letters, 2011, 32(13): 1643-1651.
- 13 Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence[C]. In 3rd European Conference on Computer Vision, 1994, 2: 151-158.
- 14 Lee Z, Juang J, Nguyen T. Local disparity estimation with three-mode cross census and advanced support weight[J]. IEEE Transactions on Multimedia, 2013, 15(8): 1855-1864.
- 15 Chang X, Zhou Z, Wang L, *et al.*. Real-time accurate stereo matching using modified two-pass aggregation and winner-take-all guided dynamic programming[C]. International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 2011: 73-79.
- 16 Salehian B, Raie A, Fotouhi A, *et al.*. Efficient interscanline consistency enforcing method for dynamic programming-based dense stereo matching algorithms[J]. Journal of Electronic Imaging, 2013, 22(4): 043028.
- 17 Zhang K, Lu J, 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.
- 18 He K, Sun J, Tang X. Guided image filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6): 1397-1409.
- 19 Rhemann C, Hosni A, Bleyer M, *et al.*. Fast cost-volume filtering for visual correspondence and beyond[C]. 24<sup>th</sup> IEEE Conference on Computer Vision and Pattern Recognition, 2011: 3017-3024.
- 20 Scharstein D, Szeliski R, Hirschmüller H. Middlebury stereo vision page[EB/OL]. [2015-12-16] <http://vision.middlebury.edu/stereo/>.
- 21 Zhu Shiping, Li Zheng. 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.
- 22 Liang W, Miao L. High-quality real-time stereo using adaptive cost aggregation and dynamic programming[C]. 3<sup>rd</sup> International Symposium on 3D Data Processing, Visualization and Transmission, 2006, 30(1): 798-805.
- 23 Wang L, Yang R. Global stereo matching leveraged by sparse ground control points[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2011: 3033-3040.

栏目编辑: 苏 岑