

doi:10.11918/j.issn.0367-6234.2015.03.010

非参数变换和改进动态规划的立体匹配算法

门宇博¹, 马 宁^{1, 2}, 张国印¹, 李 香¹, 门朝光¹, 孙鹏飞²

(1.哈尔滨工程大学 计算机科学与技术学院, 150001 哈尔滨; 2.哈尔滨师范大学 计算机科学与信息工程学院, 150001 哈尔滨)

摘 要: 针对传统稠密立体匹配方法在非纹理区、深度不连续处和遮挡处存在错误匹配率过高的问题,提出一种基于非参数变换和改进动态规划相结合的立体匹配算法.采用稀疏非参数变换相关方法计算初始局部匹配代价,并利用行列双向约束动态规划算法对匹配代价进行全局优化,在获取初始视差后分别对原始图像每一像素点进行可信性与纹理性检测,最后利用视差平面拟合结果代替非纹理与非可信区域像素点的原始视差,得到稠密视差图.实验表明,该算法具有较高的鲁棒性与匹配精度,尤其在处理图像的非纹理区、深度不连续处和遮挡处,可获得精确的匹配结果.

关键词: 立体匹配; Census 变换; 动态规划; 双向约束; 视觉导航

中图分类号: TP391.41

文献标志码: A

文章编号: 0367-6234(2015)03-0060-06

A stereo matching algorithm based on Census transform and improved dynamic programming

MEN Yubo¹, MA Ning^{1, 2}, ZHANG Guoyin¹, LI Xiang¹, MEN Chaoguang¹, SUN Pengfei²

(1.College of Computer Science and Technology, Harbin Engineering University, 150001 Harbin, China;

2. College of Computer Science and Information Engineering, Harbin Normal University, 150001 Harbin, China)

Abstract: A stereo matching algorithm based on Census transform and improved dynamic programming is proposed to the problems of traditional dense stereo matching methods, which have high false matching rate in the textureless areas, depth discontinuities and occlusion. The initial local matching cost is calculated by sparse Census transform correlation, and the raw cost is also optimized by a dynamic programming method by involving bidirectional constraints of row and column simultaneously. Meanwhile, the confidence and texture of each pixel are measured for reference image. Finally, the disparities of non-confident or textureless pixels are estimated by fitting parameters of a plane model for the corresponding segment, and the dense disparity map was obtained as well. Experiment results demonstrate that the proposed algorithm achieves high matching accuracy and robustness, especially in the textureless areas, depth discontinuities, and occlusion as well.

Keywords: stereo matching; census transform; dynamic programming; bidirectional constraint; vision navigation

立体匹配作为计算机视觉领域中的热点研究内容,在机器人自主导航、物体识别与跟踪、工业控制与检测、地理信息系统、生物医学成像及虚拟现实等领域有着广泛的应用^[1-3].Scharstein 等^[4]

把立体匹配过程概括为匹配代价计算、匹配代价集成、初始视差计算和视差优化.根据匹配基元的不同,立体匹配算法可以分为基于区域的匹配算法^[5],基于特征的匹配算法^[6]和基于相位信息的匹配算法^[7].在基于区域的算法中,Zabih 等^[8]提出的传统 Census 算法,利用像素间的大小关系作为相似性测度,解决了因相机在拍摄过程中受到光照等外部条件影响而产生的参考图像与目标图像之间的灰度差异所造成的匹配效果降低的问题.但传统 Census 的相似性测度由变换窗口的中

收稿日期: 2014-05-02.

基金项目: 国家自然科学基金(61100004); 黑龙江省自然科学基金(F201320).

作者简介: 门宇博(1989—),男,博士研究生;

张国印(1962—),男,教授,博士生导师.

门朝光(1963—),男,教授,博士生导师.

通信作者: 门朝光, menchaoguang@hrbeu.edu.cn.

心像素的灰度值决定,不能完整的描述矩形变换窗口内的信息,匹配精度不高,并且在中心点像素发生畸变的情况下会出现误匹配。

本文提出一种基于非参数变换和改进动态规划的立体匹配算法,考虑到相邻像素间的视差约束,利用改进的 Census 算法计算匹配代价,增加区域内像素点与邻域像素均值的比较,将区域内全部像素均值代替了中心点像素,有效地解决了中心点像素的畸变问题,在代价集成中利用全局的改进动态规划方法,大幅提高算法精度.此算法具有较高的匹配精度与鲁棒性,在视差优化中处理图像的非纹理区、深度不连续处和遮挡处,最终获得稠密视差图,得到较好的匹配效果。

1 初始视差图计算

1.1 稀疏 Census 非参数变换的初始匹配代价计算

Census 变换的思想是以待匹配像素点为中心取一个矩形区域,把待匹配像素点灰度值与其邻域的像素灰度值进行对比生成匹配模板,对该模板进行非参数变换获得比特串作为匹配基元,最后利用 Hamming 距离完成立体匹配^[9].但是由于 Census 变换匹配法是一种局部算法,没有考虑相邻像素间的视差约束,视差结果的平滑约束受到影响,所以本文提出的基于稀疏 Census 非参数变换法作为初始代价的计算方法.并且传统的 Census 变换的相似性测度只依赖于变换窗口的中心像素的灰度值,不能完整的描述矩形变换窗口内的信息,本文算法将邻域像素灰度的平均值作为另一个约束条件,用矩形窗口内所有像素点灰度均值代替中心点的灰度值,并且每隔一行以及一列选择一个像素点与两个均值做比对,获得两位二进制码构成的比特串作为新的匹配基元.改进后的 Census 变换可以更完整地描述所选取支持区域的像素信息,清晰地区分出两个矩阵窗口图像的不同,同时提高了算法的执行效率,解决了由于中心点因噪声问题发生畸变所带来的误匹配。

稀疏 Census 非参数变换的原理如图 1 所示,图 1(a)为稀疏 Census 非参数变换的模板,其中阴影部分为参加变换的像素,图 1(b)为传统 Census 非参数变换的模板.对图像 I 中的像素 (u,v) 进行稀疏 Census 非参数变换,首先以像素 (u,v) 为中心选取一个矩形变换窗口 W ,其窗口尺寸为 $n \times m$,计算窗口内全部像素点的均值 $\overline{I(u,v)}$,计算公式为

$$\overline{I(u,v)} = \frac{\sum_{i=-n'/2}^{n'/2} \sum_{j=-m'/2}^{m'/2} I(i+u, j+v)}{n \times m}.$$

式中: $I(i,j)$ 为像素 (i,j) 的灰度值; $n' = \lfloor n/2 \rfloor$, $m' = \lfloor m/2 \rfloor$. 以此均值代替中心像素. 计算窗口内邻域像素的均值 $\overline{I(u,v)'}$ 为

$$\overline{I(u,v)'} = \frac{\left(\sum_{i=-n'/2}^{n'/2} \sum_{j=-m'/2}^{m'/2} I(i+u, j+v) \right) - I(u,v)}{n \times m - 1}.$$

将矩形变换窗口内每隔一行以及一列的邻域像素灰度值 $I(u+i, v+j)$ 与窗口内邻域像素点的均值 $\overline{I(u,v)'}$ 和窗口内全部像素点灰度均值 $\overline{I(u,v)}$ 进行比较,可获得一组 bit 字符串. 稀疏 Census 变换表示为

$$T(u,v) = \bigotimes_{i \in N} \bigotimes_{j \in M} \xi(\overline{I(u,v)'}, I(u+i, v+j), \overline{I(u,v)}).$$

式中运算符 $\bigotimes$ 为位连接运算,辅助函数 ξ 定义为

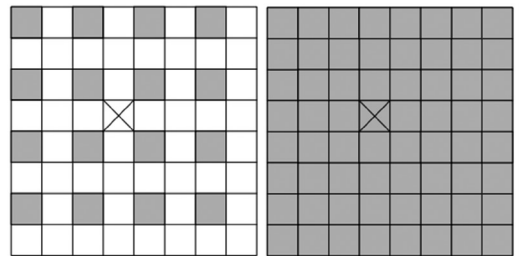
$$\xi(x,y,z) = \begin{cases} 01, & \text{if } x < y < z; \\ 10, & \text{if } x > y > z; \\ 00, & \text{if } y \leq \min(x,z); \\ 11, & \text{if } y \geq \max(x,z). \end{cases}$$

$$N = M = \{-5, -3, -1, 1, 3, 5\}.$$

基于稀疏 Census 非参数变换的初始代价计算采用 Hamming 距离的方法, Hamming 距离表示两个比特串之间在相同位置上不同比特数的数目总和,相同的位数越多, Hamming 距离越短,匹配代价越小,两个像素点越相似. 初始代价计算为

$$C(u,v,d) = \text{Hamming}(T_r(u,v), T_t(u+d,v)).$$

式中 $T_r(u,v)$ 、 $T_t(u,v)$ 分别为参考图像和目标图像的像素 (u,v) 对应的稀疏 Census 变换 bit 字符串. 至此基于稀疏 Census 变换的初始代价计算完毕,将 $C(u,v,d)$ 保存在视差空间图 DSI 中,作为下一步动态规划算法的输入数据。



(a) 稀疏 Census 模板

(b) 传统 Census 模板

图 1 稀疏 Census 非参数变换示意

1.2 行列双向约束动态规划的匹配代价集成

由于空间场景的复杂性以及局部算法的局限性,立体匹配需要借助全局优化技术获得满足平

滑约束的视差图. 动态规划算法是一种对扫描线进行约束的全局优化算法. 该类算法匹配准确率较高, 可以获得较高精度的稠密视差图^[10]. 但是传统的动态规划立体匹配算法缺乏对扫描线方向上的视差连续性的融合, 每一条扫描线被单独处理, 缺少了对扫描线之间的约束限制, 使得生成的视差图上带有明显的“条纹效应”. 本文采用基于行列双向约束的动态规划方法, 在已求得的初始匹配代价的基础上, 分别在扫描线行方向和扫描线列方向上进行全局优化以获得平滑的视差图. 首先构造全局能量代价函数为

$$E(u, v, d) = E_{\text{data}}(u, v, d) + E_{\text{smooth}_h}(u, v, d) + E_{\text{smooth}_v}(u, v, d) + E_{\text{occlude}}(u, v, d).$$

式中: $E_{\text{data}}(u, v, d)$ 为对应像素点 (u, v) 与 $(u + d, v)$ 之间的初始匹配代价, 即已求得的 $C(u, v, d)$; $E_{\text{smooth}_h}(u, v, d)$ 为扫描线行方向上的平滑约束项, 用来约束扫描线行方向上相邻像素点之间的不连续性; $E_{\text{smooth}_v}(u, v, d)$ 为扫描线列方向上的平滑约束项, 用来约束扫描线列方向上相邻像素点之间的不连续性; $E_{\text{occlude}}(u, v, d)$ 为遮挡惩罚项, 用来对遮挡像素点进行惩罚.

依据该能量函数分别在扫描线行方向和扫描线列方向上进行全局优化. 由于在扫描线行方向上暂不考虑列方向平滑性约束, 行方向能量函数 $E_h(u, v, d)$ 可简化为由数据项 $E_{\text{data}}(u, v, d)$, 扫描线行方向上的平滑约束项 $E_{\text{smooth}_h}(u, v, d)$ 以及遮挡惩罚项 $E_{\text{occlude}}(u, v, d)$ 组成.

$$E_h(u, v, d) = E_{\text{data}}(u, v, d) + E_{\text{smooth}_h}(u, v, d) + E_{\text{occlude}}(u, v, d).$$

利用动态规划算法对行方向上每一条扫描线内对应像素点之间的匹配代价值进行求解:

$$E_h(u, v, d) = \sum_{v \in V} C(u, v, d) + \sum_{v \in V} \lambda_{\text{disc}}(d_{(u, v)}, d_{(u+1, v)}) + \sum_{v \in V} \lambda_{\text{occ}}(N_{\text{occl}} + N_{\text{occr}}).$$

式中: $\sum_{v \in V} C(u, v, d)$ 为扫描线 V 内对应像素点之间的初始匹配代价之和; λ_{disc} 为所设置的平滑惩罚系数; λ_{occ} 为所设置的遮挡惩罚系数; N_{occl} 为左遮挡像素点, 即只在参考图像中可见的像素点的个数; N_{occr} 为右遮挡像素点, 即只在匹配图像中可见的像素点的个数.

根据扫描线行方向上得到的线内能量函数, 对列方向进行动态规划优化. 列方向能量函数 $E_v(u, v, d)$ 即为全局能量代价函数, 可由数据项 $E_{\text{data}}(u, v, d)$, 扫描线列方向上的视差平滑约束项

$E_{\text{smooth}_v}(u, v, d)$ 和扫描线行方向上的优化结果 $E_h(u, v, d)$ 组成.

$$E_v(u, v, d) = E_{\text{data}}(u, v, d) + E_{\text{smooth}_v}(u, v, d) + E_h(u, v, d).$$

利用动态规划算法对全局能量代价函数进行全局优化为

$$E_v(u, v, d) = \sum_{u \in U} C(u, v, d) + \sum_{u \in U} \lambda_{\text{disc}}(d_{(u, v)}, d_{(u+1, v)}) + E_h(u, v, d).$$

式中 $\sum_{u \in U} C(u, v, d)$ 为扫描线列 U 内对应像素点之间的初始匹配代价之和. 按照上述的全局优化步骤求取初始视差图.

$$d = \arg \min_{d \in D} E_v(u, v, d).$$

将获得的匹配代价和初始视差结果存储于缓存中, 作为后续步骤的输入数据.

2 视差图优化

2.1 图像可信性与纹理性检测

虽然利用行列双向约束的动态规划算法可以增加匹配的可信度, 但是由于大面积的弱纹理区域所导致的误匹配仍然存在. 针对非可信与非纹理的图像区域, 本文提出用于判别图像的纹理性和可信性的方法.

像素 (u, v) 的可信性利用最优与次优匹配代价获得

$$CM(u, v) = \min \left(255, 1024 \cdot \frac{\Delta(u, v)}{\text{maxcost}} \right).$$

式中: $\Delta(u, v)$ 为像素 (u, v) 最优与次优匹配之差; maxcost 为最优匹配代价值.

像素 (u, v) 的纹理性通过尺寸为 $n \times m$ 的矩形方差滤波函数来判断为

$$TM(u, v) = \frac{1}{n \times m} \sum_{i=-n'j=-m'}^{n' m'} I(u+i, v+j)^2 - \left(\frac{1}{n \times m} \sum_{i=-n'j=-m'}^{n' m'} I(u+i, v+j) \right)^2.$$

非可信像素与非纹理像素通过两个阈值 τ_1 和 τ_2 来判断, 只有满足下式的像素才可以直接赋值为初始视差值. 而不满足下式的像素点则利用视差面拟合计算相应的视差, 来代替该像素点的初始视差为

$$DM_{\text{init}}(u, v) = \begin{cases} DM(u, v), & \text{if } CM(u, v) \geq \tau_1, \\ & \wedge TM(u, v) \geq \tau_2; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

2.2 图像分割与视差平面拟合

对非可信像素与非纹理像素利用分割与平面

拟合的方法获取最终视差.拟合的像素包括非可信与非纹理两类,利用 mean-shift 算法^[11]对参考图像进行区域分割.mean-shift 算法利用概率分布的梯度寻找分布峰值的稳定性高的非参数估计方法,将图像聚类得到一系列互不交叉的区域.

对于每个图像分块,可使用平面方程描述其视差值.将分块内的视差 $d(x, y)$ 视作关于像素位置 (x, y) 的函数为

$$d(x, y) = ax + by + c.$$

式中: x, y 分别为像素 (x, y) 的坐标; d 为该像素对应视差值; (a, b, c) 为视差平面的参数,可以用最小二乘法求解平面参数.

$$A[a, b, c]^T = B. \quad (2)$$

式中: 矩阵 A 的第 i 行元素为 $[x_i, y_i, 1]$; B 向量的第 i 个元素为 $d(x_i, y_i)$.式(2) 可以展开为

$$\begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i y_i & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i y_i & \sum_{i=1}^m y_i^2 & \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m y_i & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i d_i \\ \sum_{i=1}^m y_i d_i \\ \sum_{i=1}^m d_i \end{bmatrix}.$$

式中 m 为分割区域中的像素点个数. 本文利用奇异值分解(SVD)求解该表达式.

$$[a, b, c]^T = (A^T A)^{-1} A^T B.$$

式中 $(A^T A)^{-1}$ 为 $A^T A$ 的 Moore-Penrose 逆矩阵,可由 SVD 求得.

3 算法流程

综上所述,本文算法流程如图2所示.

关键步骤归纳如下:

- 1) 采用稀疏 Census 变换算法对参考图像和目标图像进行局部匹配处理,获取初始匹配代价;
- 2) 利用行列双向约束的动态规划算法对初始匹配代价进行全局优化获取初始视差,其中惩罚系数普遍采用 $\lambda_{\text{disc}} = 20, \lambda_{\text{occ}} = 15$ 来进行试验;
- 3) 对参考图像逐像素点进行可信性和纹理性检测,确定非可信像素与非纹理像素位置.对于非可信像素的检测,由于图像平面的质量很大程度上取决于用于拟合的数据,所以选取 $\tau_1 = 200$

这种高可信度的阈值.对于非纹理像素的检测,为了综合提高算法可行性与准确性,在不失准确性与运算速度的同时能够检测出弱纹理区,选取 $\tau_2 = 20$;

4) 按照纹理性对参考图像进行二值化分割,同时利用 mean-shift 算法对参考图像进行灰度区域分割;

5) 利用奇异值分解计算分割面参数,将非可信像素与非纹理像素的视差值用视差平面拟合的结果替换;

6) 将计算结果合并为最终的稠密视差图.

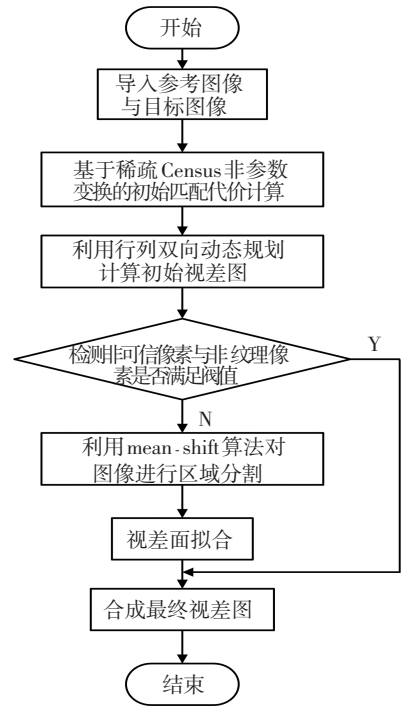


图2 算法流程

4 实验分析

为了验证本文算法的有效性与正确性,本文利用 VC++ 编程实现了所提算法.实验对象采用 Middlebury 大学立体视觉数据库中提供的 4 组标准立体图像^[4]: 包括 Tsukuba、Teddy、Venus 和 Cones 图像,尺寸分别为 384 像素×288 像素、434 像素×383 像素、450 像素×375 像素和 450 像素×375 像素,图像均已进行了极线校正,视差只存在于同扫描线方向,视差搜索范围的最大值分别是 16、20、60 和 60 像素,通过对上述图像分别进行测试验证本文算法的匹配效果.图3给出了本文算法的实验结果.

同时为了更好地评测本文算法的立体匹配结果,本文实现了另外两种文中涉及到的算法,一种算法是经典的动态规划立体匹配算法(DP 算

法)^[4],另一种算法是 Humenberger 等^[12]提出的改进 Census 非参数变换立体匹配算法,该方法是

目前非参数变换类型的算法中效果较好的一种(简称 RTCensus 算法).

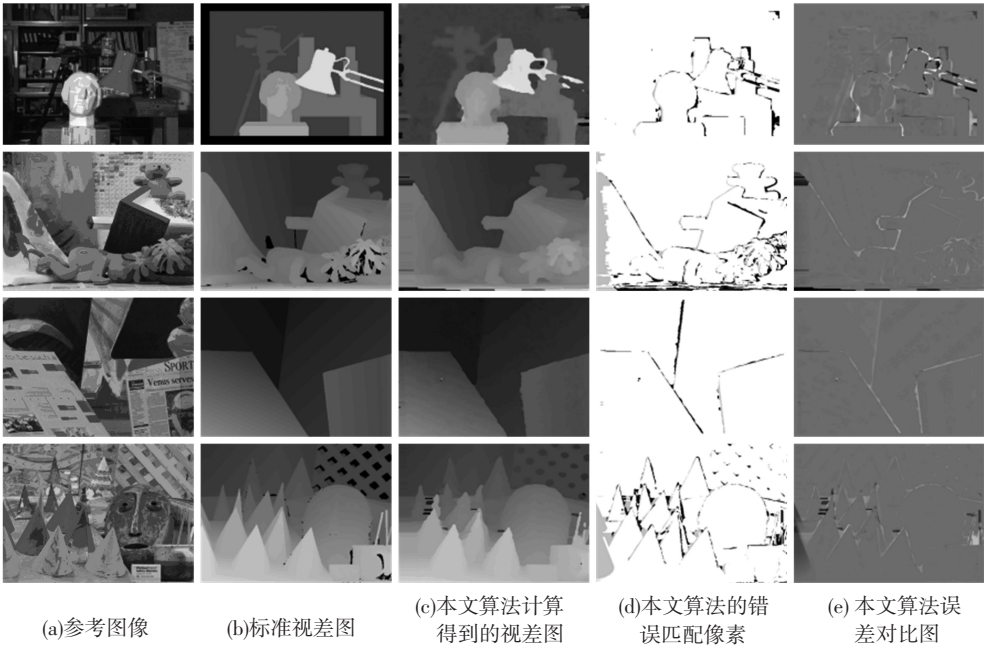


图 3 本文算法的立体匹配实验结果

图 4 给出了 3 种算法的实验结果图. 可以看出,DP 算法生成的视差图平滑性较好,物体边缘较清晰,但视差结果中存在着明显的条纹状效应,导致了错误视差向相邻像素传递;RTCensus 算法生成的视差图在纹理丰富区域效果较好,但在物体边界处(遮挡或不连续区域),产生边界模糊现象,在弱纹理区域同样存在错误匹配;本文算法得

到的视差结果优于 DP 算法和 RTCensus 算法的实验结果,由于同时考虑了行列双方向上的视差不连续性,明显消除了条纹状效应,且利用分割与平面拟合技术,在弱纹理区域、遮挡区域和视差不连续区域得到的匹配效果更加理想.另外,本文算法的初始代价计算利用了 Census 变换,使得算法具有一定的抑制噪声能力.

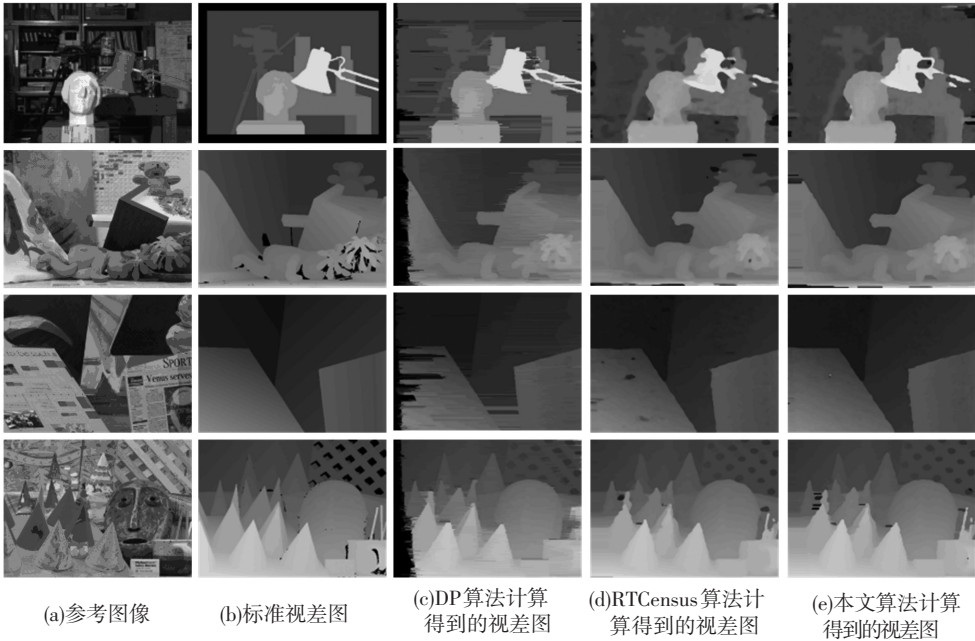


图 4 基于标准立体图像的 3 种算法比较实验结果

表 1 给出了 3 种参加比对实验的算法在错误匹配率方面的量化实验结果. 表 1 为 Nonocc(非

遮挡)区域的错误匹配率;All(包括遮挡区域在内的整体)区域的错误匹配率;Disc(视差不连续)区

域的错误匹配率.由表 1 给出的量化结果显示,本文提出的算法在非遮挡区域、整体区域和视差不

连续区域的错误匹配率都明显低于 DP 算法和 RTConsensus 算法,取得了较高的匹配精度.

表 1 本文算法与 DP 算法、RTConsensus 算法的量化实验结果比较

算法	错误匹配率/%												平均百分比/%
	Tsukuba			Venus			Teddy			Cones			
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	
本文算法	3.13	4.20	11.9	1.08	1.87	13.6	5.68	11.6	17.1	3.79	9.26	11.3	6.73
RTConsensus ^[12] 算法	5.08	6.25	19.2	1.58	2.42	14.2	7.96	13.8	20.3	4.10	9.54	12.2	9.73
DP ^[4] 算法	4.12	5.04	12.0	10.1	11.00	21.0	14.00	21.6	20.6	10.50	19.10	21.1	14.20

根据上述实验结果可知,本文算法有效解决了立体影像的匹配问题,算法匹配正确率较高,尤其在弱纹理区域、遮挡区域和视差不连续区域具有良好的稳定性.

5 结 语

本文提出一种基于非参数变换和改进动态规划的立体匹配算法,采用稀疏 Census 变换计算初始局部匹配代价,构造基于行列双向的全局能量函数,并利用改进动态规划算法求解初始视差,对非纹理与非可信区域的像素,通过视差平面拟合代替初始视差结果.算法最大限度地利用了图像纹理信息以及扫描线相邻像素点间的平滑信息,有效克服了弱纹理区域、遮挡区域和视差不连续区域对视差结果的影响.实验结果表明,该算法在匹配效果和匹配精度方面均优于传统的动态规划立体匹配算法和鲁棒性较好的基于 Census 非参数变换立体匹配算法,降低了错误匹配率,可以获得匹配准确率较高的稠密视差图,得到较好的匹配效果.

参考文献

[1] BROWN M Z, BURSCHKA D, HAGER G D. Advances in computational stereo [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(8): 993-1008.

[2] HUMENBERGER M, ENGELKE T, KUBINGER W. A census-based stereo vision algorithm using modified semi-global matching and plane fitting to improve matching quality [C]//Proceeding of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). San Francisco, CA: IEEE, 2010: 77-84.

[3] KUMAR S, KUMAR S, SUKAVANAM N, et al. Human visual system and segment-based disparity estimation [J]. AEU-International Journal of Electronics and Communications, 2013, 67(5): 372-381.

[4] SCHARSTEIN D, SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. International Journal of Computer Vision, 2002, 47(1/3): 7-42.

[5] RHEMANN C, HOSNI A, BLEYER M, et al. Fast cost-volume filtering for visual correspondence and beyond [C]//Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC: IEEE Computer Society, 2011: 3017-3024.

[6] NIERADKA G, BUTKIEWICZ B S. Features stereo matching based on Fuzzy logic [C]//Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference. Lisbon, Portugal, 2009: 1188-1193.

[7] ULUSOY I, HANCOCK E R. A statistical approach to sparse multi-scale phase-based stereo [J]. Pattern Recognition, 2007, 40(9): 2504-2520.

[8] ZABIH R, WOODFILL J. Non-parametric Local Transforms for Computing Visual Correspondence [M]. Springer, Berlin-Heidelberg: Computer Vision—ECCV' 94, 1994: 151-158.

[9] FIFE W S, ARCHIBALD J K. Improved census transforms for resource-optimized stereo vision [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(1): 60-73.

[10] TORR P H S, CRIMINISI A. Dense stereo using pivoted dynamic programming [J]. Image and Vision Computing, 2004, 22(10): 795-806.

[11] COMANICIU D, MEER P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.

[12] HUMENBERGER M, ZINNER C, WEBER M, et al. A fast stereo matching algorithm suitable for embedded real-time systems [J]. Computer Vision and Image Understanding, 2010, 114(11): 1180-1202.