

# 一种快速双目视觉立体匹配算法

狄红卫<sup>1,2</sup> 柴颖<sup>2</sup> 李逵<sup>2</sup>

(<sup>1</sup> 重大工程灾害与控制教育部重点实验室, 广东 广州 510632; <sup>2</sup> 暨南大学光电工程系, 广东 广州 510632)

**摘要** 针对目前双目视觉立体匹配算法计算量过大、实时性不强的问题, 提出了一种平行配置系统的快速立体匹配算法。利用两幅视图的差异将视图分为特征点和非特征点, 然后对特征点采用 WTA (winner-take-all) 方法进行匹配, 而对非特征点只进行简单的验证, 最后得出致密的视差图。该算法利用视差的分段连续性, 大大减少了运算量。实验结果表明, 该算法提取的特征点集中于视差不连续区, 与目前其它基于区域的匹配算法相比, 该算法得到的误匹配像素百分比与其它算法相当, 而计算速度却提高了一个数量级, 并且边缘特征较好, 是一种有效可行的高实时性立体匹配算法。

**关键词** 图像处理; 双目视觉; 立体匹配; WTA; 视差

**中图分类号** TP391.41 **文献标识码** A **doi**: 10.3788/AOS20092908.2180

## A Fast Binocular Vision Stereo Matching Algorithm

Di Hongwei<sup>1,2</sup> Chai Ying<sup>2</sup> Li Kui<sup>2</sup>

<sup>1</sup>Key Laboratory of Disaster Forecast and Control in Engineering, Ministry of Education, Guangzhou, Guangdong 510632, China

<sup>2</sup>Department of Optoelectronic Engineering, Jinan University, Guangzhou, Guangdong 510632, China

**Abstract** To resolve the problem of large computational cost and bad real-time performance of current binocular vision stereo matching algorithm, a fast stereo matching algorithm of parallel system is presented. Firstly, based on the difference of two images, the images are divided into feature points and non-feature points. Then the feature points are matched using winner-take-all (WTA) method and the non-feature points are validated simply. Finally a dense disparity map is yielded. With the fact that the disparity map consists of piecewise-smooth surfaces, the computational cost is decreased greatly. Experimental results show that the extracted feature points are concentrated in the depth discontinuity regions. Compared with other area-base matching algorithms, the computational speed of the presented algorithm has been improved for one order of magnitude while the percentage of bad matching pixels is equivalent to other algorithms. Besides, the edge feature is better, and thus this method is an effective and feasible stereo matching algorithm with good real-time performance.

**Key words** image processing; binocular vision; stereo matching; winner-take-all (WTA); disparity

## 1 引言

双目立体视觉是计算机视觉的一个重要分支, 即由不同位置的左右两个摄像机获得同一场景的两幅视图, 通过计算空间点在两幅图像中的视差, 获得该点的三维坐标值。立体匹配是双目立体视觉中最关键、最困难的问题, 一直是立体视觉领域研究的焦点。立体匹配算法的目的就是找出两幅图像对应的匹配点, 进而获得距离信息。目前, 还没有一种方法

能够完美地解决图像对应点匹配问题。因此, 如何实现图像间快速、准确的匹配对立体视觉的研究具有重要意义。

常见的立体匹配算法有基于特征的匹配算法和基于区域的匹配算法。基于特征的匹配算法只匹配图像中的一些特征点, 如角点、边缘等, 因此基于特征的匹配算法产生稀疏的视差图<sup>[1~3]</sup>。如基于 SIFT 特征描述子的立体匹配算法<sup>[4]</sup>, SIFT

收稿日期: 2008-07-18; 收到修改稿日期: 2008-11-20

作者简介: 狄红卫(1969-), 男, 教授, 博士, 主要从事信息处理与模式识别、信息与图像融合等方面的研究。

E-mail: tdhw@jnu.edu.cn

特征描述符采用 128 维向量表征每个特征点,具有良好的尺度、旋转、光照不变性,将其应用到立体匹配中可以得出较好的效果,匹配结果有较高的精确度和鲁棒性,但时间复杂度高,实时性不强。

基于区域的匹配方法产生致密的视差图。基于区域的匹配算法主要有优胜者全选(winner-take-all, WTA)、动态规划(dynamic programming)、扫描线优化(scan-line optimization)、模拟退火(simulated annealing)、图像分割(graph cut)等<sup>[5]</sup>。其中 WTA 是在一个视图中取一个窗口,然后在另一视图中计算视差范围内所有匹配代价,选取匹配代价最小点的差值作为视差,这种方法的误匹配率较高,原因是得到的匹配点是局部解而非全局最优解。动态规划引入了平滑约束,计算出所有点的匹配代价,然后求出视差的全局最优解,这种方法的结果较为精确,而模拟退火、图像分割算法运算量很大。为了解决无纹理区域、遮挡区域视差不准确问题,文献[6]采用三目视觉方法,文献[7]采用自适应窗口的方法,计算小尺度窗口内的局部差异,如果差异小,则加大窗口。文献[8]采用在不同尺度上将图像分割成矩形子区域进行匹配。文献[9]利用马尔可夫网络和可信度传播进行立体匹配,获得了较好的结果。文献[10]则采用了遗传算法进行立体匹配,增加了立体匹配算法的途径。

基于区域的立体匹配算法一般有下列步骤:第一步为计算匹配代价;第二步,对匹配代价进行均值滤波;第三步,根据不同的算法确定视差。然而目前基于区域的立体匹配算法复杂度非常高,假设图像大小为  $M \times N$ ,最大视差为  $d_{\max}$ ,则第一、二步的复杂度为  $MNd_{\max}$ ,第三步算法若选 WTA 或动态规划,则复杂度为  $MNd_{\max}$ ,其它算法复杂度更高,因此这类算法实时性不强。针对基于区域的立体匹配算法实时差的问题,文献[11]采用小波变换的方法,文献[12]采用减小矩阵大小、由粗到精、汇编语言优化、编译优化等方法,提高了实时性。

本文提出了一种有效减少计算量、提高实时性的新算法。该算法将以上三步合为一步,并改进了 WTA 算法,利用视差的分段连续性,大大减少了计算量,得出致密的视差图。实验结果表明提出的算法匹配精度与大部分基于区域的匹配算法相当,而计算速度却提高了一个数量级。

## 2 快速双目立体匹配算法

为了获得参考图像上所有像点在另一幅图像中的对应位置,需要将该像点的邻域与另一幅图像中对应位置附近所有像点的邻域进行相似值计算,即匹配代价计算,然后在一些约束条件(如内极线约束、唯一性约束、连续性约束、顺序性约束以及视差梯度极约束等等)的指导下选取出一个最可能的像素作为对应的像素来完成匹配搜索。由于在匹配计算前需要完成镜头的畸变恢复、配准等操作,最后得到等效于平行配置的成像系统输出的两幅左右视图。平行配置的成像系统在对对应点的搜索过程中,只需要在相同行坐标约束下的一个左右邻域内进行即可。

本文考虑平行配置的成像系统,因此,只在水平扫描线上进行匹配<sup>[11]</sup>。由于图像中存在无纹理区域,以此区域中的点进行匹配往往是局部最优解而非全局最优解,因此会出现误匹配,这正是 WTA 方法误匹配率高的原因。由于视差的分段连续性,在场景中大部分区域的视差差别不大,因此本文算法将邻域的视差代入,验证两幅视图间的差异,若小于设定的阈值,则认为视差不改变,若大于阈值,则此点可能是视差跳变的边缘,从两幅视图中选择一点记为特征点,用 WTA 的方法进行匹配。该算法不但简化了运算,而且避免由局部最优解带来的误码差。

图像匹配中常用的相似性测度函数有 SSD(sum of square difference)、SAD(sum of absolute difference)、NCC(normalized cross correlation)等,用在匹配中三者之间差别不大,本文采用 SAD 作为相似性测度函数。

### 2.1 特征点提取

假定左、右视图分别为  $L$ 、 $R$ ,则两视图在当前视差  $d_{cu}$  下的绝对差为

$$D(x, y, d_{cu}) = |L(x, y) - R(x - d_{cu}, y)|, \quad x \in (1, N), y \in (1, M) \quad (1)$$

其中  $M, N$  分别为  $L$  和  $R$  的行数和列数,若是彩色图像,则  $D(x, y, d_{cu})$  为 RGB 三色绝对差之和。

符合以下条件的点为特征点  $P$

$$P = \begin{cases} L(x, y), & \delta_L \geq \delta_R, \quad D(x, y, d_{cu}) \geq T_{th} \\ R(x - d_{cu}, y), & \delta_L < \delta_R, \quad D(x, y, d_{cu}) \geq T_{th} \end{cases} \quad (2)$$

其中  $\delta_L$  为视图  $L$  在  $(x, y)$  点的水平梯度,  $\delta_R$  为视图  $R$  在  $(x - d_{cu}, y)$  点的水平梯度,  $T_{th}$  为设定的阈值。

$$\delta_L = \begin{cases} |L(x, y) - L(x + 1, y)|, & \text{匹配方向自右向左} \\ |L(x, y) - L(x - 1, y)|, & \text{匹配方向自左向右} \end{cases} \quad (3)$$

$$\delta_R = \begin{cases} |R(x - d_{cu}, y) - R(x - d_{cu} + 1, y)|, & \text{匹配方向自右向左} \\ |R(x - d_{cu}, y) - R(x - d_{cu} - 1, y)|, & \text{匹配方向自左向右} \end{cases} \quad (4)$$

由于左、右视图中的物体只在水平方向有位移,所以只需求水平梯度。如果匹配方向不同,水平梯度变化点也要区别对待。

## 2.2 匹配算法

假定匹配方向自右向左,以  $L$  为基准视图,当前视差为  $d_{cu}$ 。由于视差的分段连续性,将匹配分为在当前视差  $d_{cu}$  下绝对差大于阈值和小于阈值(即特征点和非特征点)两种情况,匹配结果可表示为

$$d_{result} = \begin{cases} d_{cu}, & D(x, y, d_{cu}) < T_{th} \\ d_{new}, d_{cu} = d_{new}, & D(x, y, d_{cu}) \geq T_{th} \end{cases} \quad (5)$$

其中  $d_{new}$  为对特征点匹配得到的新视差,  $d_{result}$  为视差结果。匹配方法是提取特征点,以该点为中心取一个  $n \times n$  的窗口( $n$  为奇数),分别计算特征点和其在另一视图中视差范围  $0 \sim d_{max}$  内的匹配代价  $C(x_0, y_0, d)$ :

$$C(x_0, y_0, d) = \begin{cases} \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \sum_{j=-\frac{n-1}{2}}^{\frac{n-1}{2}} |L(x_0 + i, y_0 + j) - R(x_0 + i - d, y_0 + j)|, & \text{特征点为 } L(x_0, y_0) \\ \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \sum_{j=-\frac{n-1}{2}}^{\frac{n-1}{2}} |R(x_0 + i, y_0 + j) - L(x_0 + i + d, y_0 + j)|, & \text{特征点为 } R(x_0, y_0) \end{cases} \quad (6)$$

式中  $d \in (0 \sim d_{max})$ 。

取匹配代价最小的  $d$  为  $d_{new}$ , 并记为该点视差结果  $d_{result}$ 。

这里要区分三种情况。

(1)  $d_{new} = d_{cu}$ , 这种情况说明特征点处不存在视差跳变。

(2)  $d_{new} > d_{cu}$ , 视差变大时, 此时的特征点出现在  $L$  视图中, 对应于视差跳变点, 得到的结果正确, 如图 1。

(3)  $d_{new} < d_{cu}$ , 视差变小时, 视差跳变点处可能不出现特征点, 即使出现, 因为大视差的前景总是覆盖在小视差的背景上, 也表现为大视差, 检测到的特征点会出现在视差跳变以后的某个点, 如图 2。因此, 若检测到视差变小, 这时的特征点不能反映视差跳变, 需要再以小视差为当前视差  $d_{cu}$ , 自左向右扫描, 直到找到视差跳变点。

这种匹配方法对于非特征点视差的确定一部分来源于邻近的特征点, 一部分来源于对该点的验证, 这与基于可信度传播(Belief propagation)匹配算法思想一致, 因此, 该算法也是对基于可信度传播的匹配算法的简化。本算法的优点在于下面三个方面。

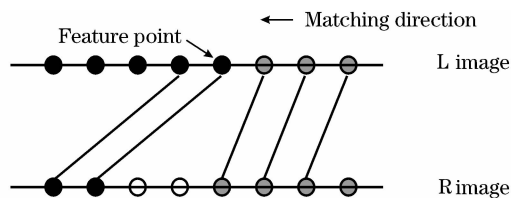


图 1 视差变大

Fig. 1 Disparity increase

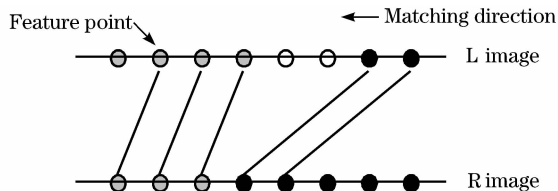


图 2 视差变小

Fig. 2 Disparity decrease

(1) 利用视差的分段连续性, 将匹配集中于视差不连续区域, 而视差连续区域只进行简单的验证, 从而极大地减少运算量;

(2) 可以跳过遮挡区域;

(3) 视差跳变边缘效果较好。

本算法设定初始视差为 0, 从第二行开始, 每行最右端的当前视差  $d_{cu}$  从上一行最右端得到。

### 3 立体匹配算法评价

在已存在真实视差图的基础上,有以下评价指标。

(1) 均方根误差 (Root-mean-squared error)

$$N_{RMSE} = \left[ \frac{1}{MN} \sum_{x=1}^N \sum_{y=1}^M |d_c(x,y) - d_T(x,y)|^2 \right]^{\frac{1}{2}}, \tag{7}$$

式中  $d_c(x,y)$  为估计视差,  $d_T(x,y)$  为真实视差。

(2) 误匹配像素百分比 (Percentage of bad matching)  $P_{BM}$

$$P_{BM} = \left[ \frac{1}{MN} \sum_{x=1}^N \sum_{y=1}^M |d_c(x,y) - d_T(x,y)| > \Delta_d \right], \tag{8}$$

这里  $\Delta_d$  为可接受错误阈值,一般取 1。

为了更有效地进行统计分析,一般将图像分为

3 种区域:无纹理区、遮挡区、视差不连续区。无纹理区是指在一个方窗内,水平梯度的均方值小于阈值。遮挡区为大视差的前景遮挡小视差的背景而产生的。视差不连续区域:若该点与邻域的视差变化大于阈值时,即为视差不连续区域。

### 4 实验结果与分析

根据以上算法,我们进行了仿真实验,实验对象采用 Middlebury 立体图像库中的 Tsukuba、Sawtooth、Venus 图像,阈值  $T_{th}$  分别为 29, 25, 24, 得到的结果如图 3、图 4、图 5。其中(a)为真实视差图,(b)为本文算法得到的视差图,(c)为 WTA 算法,(d)为动态规划算法,(f)为扫描线优化算法。从图中可以看出本文算法视差图边缘特征较好,这是由于本算法提取的特征点与视差跳变边缘接近。

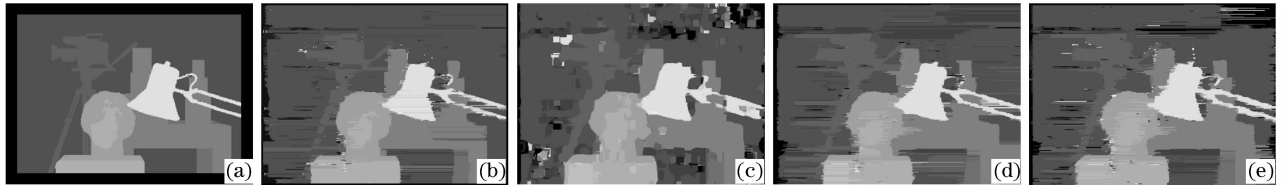


图 3 Tsukuba 图像匹配结果

Fig. 3 Results on Tsukuba data set. (a) Ground truth; (b) our method; (c) WTA; (d) dyn. prog; (e) scan. opt

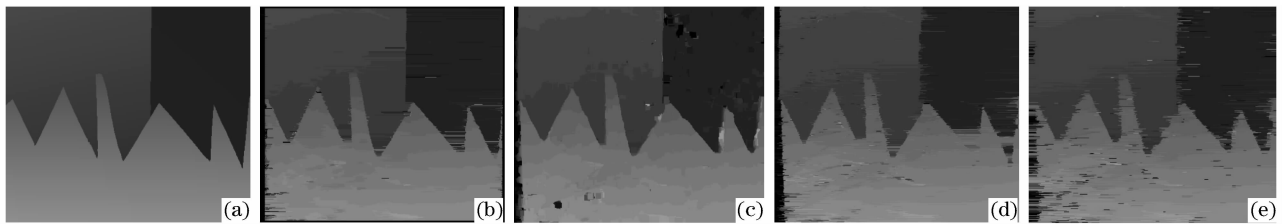


图 4 Sawtooth 图像匹配结果

Fig. 4 Results on Sawtooth data set. (a) Ground truth; (b) our method; (c) WTA; (d) dyn. prog; (e) scan. opt

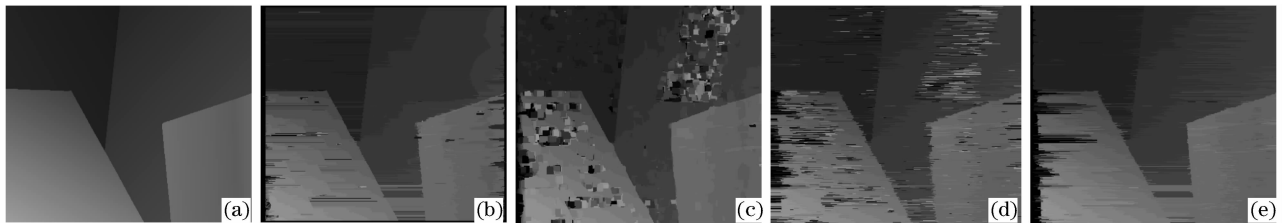


图 5 Venus 图像匹配结果

Fig. 5 Results on Venus data set(a) Ground truth; (b) our method; (c) WTA; (d) dyn. prog; (e) scan. opt

表 1 为误匹配像素百分比(PBM)值比较。其中 all 为所有像素, untex. 为无纹理区域, disc. 为视差不连续区域。由实验数据我们可以看出,本文算法的精度与常用算法相当。

表 2 为上述匹配算法在计算时间上的比较,计算机的配置为 P4 2.4 GHz, 256 MB 内存。由于本

文算法将匹配集中于视差不连续区域,而视差连续区域则只进行简单的验证,从而极大地减少了运算量。图 3 中,用我们算法提取的特征点只占图像的 14.3%。从表中可以看出本文算法在时间上有明显的改进,计算速度提高了一个数量级。

表 1 误匹配像素百分比(PBM)比较

Table 1 Comparison on percentage of bad matching

	Tsukuba			Sawtooth			Venus		
	all	untex.	disc.	all	untex.	disc.	all	untex.	disc.
Proposed method	4.82	4.09	11.28	4.51	3.04	10.87	9.80	15.41	13.12
WTA(SSD+MF)	5.23	3.80	24.66	2.21	0.72	13.97	3.74	6.82	12.94
Scan. opt	5.08	6.78	11.94	4.06	2.64	11.90	9.44	14.59	18.20
Dyn. prog	4.12	4.63	12.34	4.84	3.71	13.26	10.10	15.01	17.12

表 2 不同算法匹配的时间比较

Table 2 Running times of several algorithms

Algorithm	Tsukuba	Sawtooth	Venus
SSD+MF	2.344	3.625	3.625
Dyn. prog	2.031	3.547	3.562
Scan. opt	2.359	3.781	3.844
Proposed method	0.093	0.375	0.297

## 5 结 论

本文利用视差的分段连续性,用两幅视图提取特征点,将匹配运算集中于视差不连续区域,其余点则只进行简单的验证。本文算法在匹配精度与常用算法相当的情况下,大大减少了运算量,提高了实时性,并且边缘特征较好,是一种有效可行的高实时性算法。从实验结果看,本文算法得到的视差图存在“线条”特征,在视差缓变处精度略低,对这些问题的改进是今后的研究方向。

## 参 考 文 献

- Xiong Chao, Tian Xiaofang, Lu Qiyong. Real-time binocular matching and depth measurement based on competing schema[J]. *Computer Engineering and Applications*, 2006, (1): 83~85  
熊超,田小芳,陆起涌.基于竞争机制的双目视觉匹配与实时测距[J]. *计算机工程与应用*, 2006, (1): 83~85
- Wang Lichao, Xiong Chao, Wang Chenyi *et al.*. Simplified binocular stereo distance measurement algorithm and a system based on competition [J]. *Chinese Journal of Sensors and Actuators*, 2007, **20**(1): 150~153  
王力超,熊超,王晨毅等.基于竞争机制的简化双目立体视觉测距算法及系统设计[J]. *传感技术学报*, 2007, **20**(1): 150~153

- W. Zhang, X. Cao, E. Sung *et al.*. A feature-based matching scheme; MPCD and robust matching strategy [J]. *Pattern Recognition Letters*, 2007, (28): 1222~1231
- Zai Xiaotao, Zhao Yuming. A stereo matching algorithm based on SIFT feature descriptor [J]. *Microcomputer Information*, 2007, **20**(8): 285~287  
宰小涛,赵宇明.基于 SIFT 特征描述子的立体匹配算法[J]. *微计算机信息*, 2007, **20**(8): 285~287
- Daniel Scharstein, Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *International J. Computer Vision*, 2002, **47**(1/2/3): 7~42
- Gu Zheng, Su Xianyu. An algorithm based on adaptive support-weight and disparity adjustment for trinocular stereo-matching [J]. *Acta Optica Sinica*, 2008, **28**(4): 734~738  
顾征,苏显渝.三目自适应权值立体匹配和视差校正算法[J]. *光学学报*, 2008, **28**(4): 734~738
- C. Georgoulas *et al.*. Real-time disparity map computation module [J]. *Microprocessors and Microsystems*, 2008, **32**(3): 159~170
- Changming Sun. Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques [J]. *International Journal of Computer Vision*, 2002, **47**(1/2/3): 99~117
- Jian Sun, Heung-Yeung Shum, Nan-Ning Zheng. Stereo matching using belief propagation [J]. *Pattern Analysis and Machine Intelligence*, 2003, **25**(7): 787~800.
- Wang Biao, Wang Jinyan, He Yizheng *et al.*. A novel stereo matching algorithm [J]. *Computer Engineering*, 2005, **31**(10): 153~155  
王彪,王金岩,何亦征等.一种立体匹配的新算法[J]. *计算机工程*, 2005, **31**(10): 153~155
- Li Haibin, Zhang Qiang. A new stereo matching algorithm based on sub-line segment [J]. *Acta Optica Sinica*, 2007, **27**(5): 907~911  
李海滨,张强.一种新的基于子线段的立体匹配算法[J]. *光学学报*, 2007, **27**(5): 907~911
- Forstmann S, Kanou Y, Ohya J *et al.*. Real-time stereo by using dynamic programming [C]. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop*. Washington: IEEE Press, 2004: 29~36