

一种利用动态规划和左右一致性的立体匹配算法

张浩峰 赵春霞

(南京理工大学计算机科学与技术学院, 南京 210094)

摘要 立体匹配是计算机视觉领域研究的一个重要课题,为了得到准确、稠密的视差图,提出了一种利用动态规划和左右一致性的立体匹配算法。该算法首先分别以左、右图像为基元图像,计算各自的视差空间图像,在视差空间图像上利用动态规划,计算得到左视差图和右视差图;然后通过使用左右视差图之间的一致性关系,消除误匹配点,得到较为准确的部分视差图;最后利用视差图的顺序约束关系,给出未匹配视差点的搜索空间计算方法,并利用一种简单有效的方法来计算这些点的视差值。在一些标准立体图像对上所做的实验结果表明,该算法效果良好。

关键词 立体匹配 动态规划 左右一致性 视差搜索策略

中图分类号:TP391.4 文献标识码:A 文章编号:1006-8961(2008)04-0756-05

A Stereo Matching Algorithm Using Dynamic Programming and Left-right Consistency

ZHANG Hao-feng, ZHAO Chun-xia

(Department of Computer Science & Technology, Nanjing University of Science & Technology, Nanjing 210094)

Abstract Stereo matching is one of the most important researches in computer vision. In order to obtain dense and correct disparity, a stereo matching algorithm using dynamic programming and left-right consistency is presented. Firstly, the left and right disparity space images are computed using the left and right images as basic image separately. Secondly, in the computed disparity space images, the disparity images are computed using dynamic programming. Then the left-right consistency of the disparity images is used to remove the mismatching pixels, and to generate the part of near real disparity images. At last, according to the ordering constraint of disparity image, a method detecting the searching space of unmatched pixels is presented. And an ordinary but efficient strategy is proposed to finalize these pixels. The experiments on some standard stereo pairs are executed, and the results show the algorithm is effective.

Keywords stereo matching, dynamic programming, left-right consistency, disparity searching strategy

1 引言

计算机立体视觉是场景理解的重要部分,可以广泛应用于图像分割、物体3D重建、立体视频流的编码、车辆导航等领域。

立体匹配是立体视觉算法中最重要任务之一,它通过双目或者多目图像匹配得到视差图,再根据三角测量关系进一步得到景物的深度。根据约束

方式的不同,立体匹配方法可以大致分为两种^[1,2],一种是只对像素周围小区域进行约束的局部匹配方法;另一种是对扫描线甚至整幅图像进行约束的全局匹配方法。局部匹配方法的优点是计算量小,速度快,易于硬件实现和加速,许多立体视觉系统采用了局部匹配方法,但是由于局部寻优,对遮挡区、无纹理区比较敏感,和全局匹配方法相比其视差图的误匹配率比较高,这一类算法主要有基于区域的匹配方法^[3,4]、基于梯度的优化方法^[5]、基于特征的匹

配方法^[6];全局匹配方法的优点是匹配准确性较高,但是计算量较大,耗时较长,这一类算法主要有动态规划^[7-9]、信任度传播^[10]、Graph Cut^[11]等。

Birchfield 等人提出了一种点对点的深度图计算以及遮挡检测方法,其主要原理就是利用动态规划在视差空间图像中寻找最优路径^[8],但是该算法对于一些块状的误匹配区域并不能很好地消除。为此在其基础上,提出了利用左右视差图的一致性关系,消除误匹配点,并利用视差图之间的顺序约束来给出未匹配点的搜索空间计算方法。

2 动态规划立体匹配原理

自从1957年Bellman提出动态规划理论以来,它在各个科学领域得到了广泛的研究和应用。动态规划在立体匹配中的应用主要是在视差空间图像(DSI)上寻找一条最优路径来表示一条扫描线上的所有像素点的视差。

2.1 视差空间图像的生成

生成视差空间图像有两种方法,一是以右图像像素为基元在左图像相应的极线上搜索;另一种是以左图像像素作为基元,在右图像相应的极线上搜索,如图1所示。设待匹配的左右图像分别为 f 和 g 。图像的大小为 $W \times H$, 图像对的最大视差为 D 。以右图像像素为基元可以得到视差空间图像为

$$S(d, i) = |f(i + d, j) - g(i, j)| \quad (1)$$

$$i \in [0, W), j \in [0, H), d \in [0, D]$$

以左图像像素为基元可以得到的视差空间为

$$S(d, i) = |g(i - d, j) - f(i, j)| \quad (2)$$

$$i \in [0, W), j \in [0, H), d \in [0, D]$$

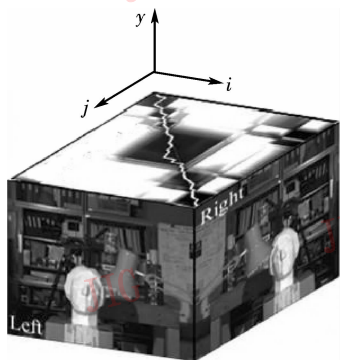


图1 视差空间图像

Fig. 1 Disparity space image

2.2 动态规划寻径

动态规划主要是用来处理多阶段决策问题的,即将决策问题的全过程恰当地划分为若干个相互联系的子过程(每个子过程为一个阶段),以便按照一定的次序去求解。

在视差空间图像中,将图像的每一列看成是一个阶段,从第0阶段开始,到第 k 阶段,计算所得的最优路径所花费的能量为

$$E_k = \sum_{i=0}^k Q(d_i) + \lambda \sum_{i=0}^k O(d_i, d_{i-1}) - \alpha \sum_{i=0}^k M(d_i) \quad (3)$$

式中, $Q(d_i)$ 为所选择的第 i 个像素视差为 d 的需要的能量; $O(d_i, d_{i-1})$ 为第 i 个像素是否遮挡需要的能量; $M(d_i)$ 为第 i 个像素是否有匹配点所获得的能量。当能量最小的时候,即为最优路径,也就是所求得的最优视差图。所以动态规划的方法即为求路径 $P_k(d) = \operatorname{argmin}(E_k)$ 。

3 算法描述

本文算法包含3个步骤,首先设定立体图像对之间的视差约束关系,并给出最终的能量函数,利用动态规划的方法分别计算左右视差图;然后根据左右视差图之间的关系,去除错误匹配点;最后在上一阶段所得到的视差图上,填充错误匹配所留下的空隙,得到完整的视差图。

3.1 视差约束关系以及能量函数

算法遵循以下几个视差约束关系:

(1) 设所拍摄的景物离摄像机都有一定的距离,那么所有的视差值都应该处于一定的范围之内,即 $0 \leq d = j - i \leq D$, 其中, i, j 分别为图像中的匹配点, D 为允许的最大视差值。

(2) 图像匹配点不会交叉,即 $x_i < x_j \&\& y_i < y_j$, $0 \leq i \leq j < W$, 其中 x_i, x_j 为左图像上的像素点, y_i, y_j 为右图像上的像素点, x_i, y_i 和 x_j, y_j 为两对匹配点。

(3) 在同一段匹配上,当存在遮挡时,或者左图像连续有视差,没有遮挡,而右图像有遮挡;或者右图像有连续视差,没有遮挡,而左图像有遮挡。 $x_{i+1} = x_i + 1$ or $y_{i+1} = y_i + 1$, 其中 $i = 0, 1, \dots, k, \dots, W - 2$ 。

对于约束3,当以右图像为匹配基元时,使用后向逐步搜索的动态规划,如图2(b),有条纹的网格

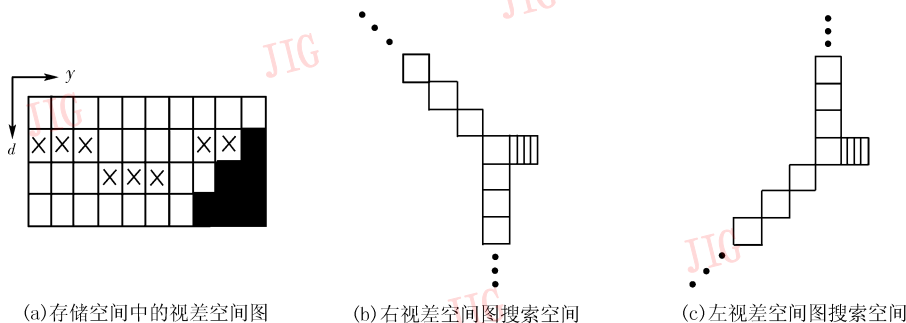


图 2 视差空间图中的搜索空间

Fig. 2 Search space in DSI

为当前搜索像素,空心网格为路径上的上一个视差点。当左图像像素连续有视差时, $x_i = x_{i-1} + 1$, $y_i = y_{i-1} + k$, $x_i = y_i + d_i$, $x_{i-1} = y_{i-1} + d_{i-1}$, 可以求解得到 $d_{i-1} = d_i + (k - 1)$, $y_{i-1} = y_i - k$, 那么规划路径上的上一个可能的点为 $(d_i + (k - 1), y_i - k)$ 。当右图像像素连续有视差时 $x_i = x_{i-1} + k$, $y_i = y_{i-1} + 1$, 可以求解得到 $d_{i-1} = d_i - (k - 1)$, 那么规划路径上的上一个可能的点为 $(d_i - (k - 1), y_i - 1)$ 。

当以左图像为匹配基元时,使用后向逐步搜索的动态规划,如图 2(c),当左图像连续有视差的时候 $x_i = x_{i-1} + k$, $y_i = y_{i-1} + 1$, $x_i = y_i + d_i$, $x_{i-1} = y_{i-1} + d_{i-1}$, 可以求解得到 $d_{i-1} = d_i - (k - 1)$, 那么规划路径上的上一个可能的点为 $(d_i - (k - 1), y_i - 1)$ 。当右图像像素连续有视差时, $x_i = x_{i-1} + 1$, $y_i = y_{i-1} + k$, 可以求解得到 $d_{i-1} = d_i + (k - 1)$, $y_{i-1} = y_i - k$, 那么规划路径上的一个可能点为 $(d_i + (k - 1), y_i - k)$ 。

通过以上几个匹配约束,将式(3)简化成如下形式:

$$E_k = \sum_{i=1}^{N_m} S(d_i) + \lambda_{occ} N_{occ} - \alpha_m N_m \quad (4)$$

式中, $Q(d_i)$ 为其确定为匹配点的消耗, N_m 和 N_{occ} 分别为匹配的像素数和遮挡数, λ_{occ} , α_m 分别为对遮挡点的惩罚和对匹配点的奖励。当使用右图像为匹配基元时,那么对上式使用动态规划求最佳路径,就可以转化为

$$E_i(d_i, y_i) = Q(d_i) - \alpha_m + \min \begin{cases} E(d_i, y_i - 1) & \text{both continue} \\ E(d_i - (k - 1), y_i - k) + \lambda_{occ} & \text{right continue} \\ E(d_i + (k - 1), y_i - 1) + \lambda_{occ} & \text{left continue} \end{cases} \quad (5)$$

同样,当用左图像为匹配基元时,

$$E_i(d_i, y_i) = Q(d_i) - \alpha_m +$$

$$\min \begin{cases} E(d_i, y_i - 1) & \text{both continue} \\ E(d_i - (k - 1), y_i - k) + \lambda_{occ} & \text{left continue} \\ E(d_i + (k - 1), y_i - 1) + \lambda_{occ} & \text{right continue} \end{cases} \quad (6)$$

3.2 消除错误匹配点

由 3.1 节可以得到分别以左右图像为基元的视差图,设 $fd(x, y)$, $gd(x, y)$ 分别为以左、右图像为匹配基元的视差图。对于原始图像上的一对匹配点 $f(x_i, y)$, $g(x_j, y)$, 设其视差为 d , 则有 $d = gd(x_j, y)$, $d' = fd(x_j + d, y)$, 因为 d 与 d' 是相对应的一对视差, 所以应该相等。使用下式计算可以得到最终的去除错误匹配点的视差图。

$$\text{disparity}(x_j, y) = \begin{cases} (d + d')/2 & |d - d'| \leq 1 \\ 0 & \text{其他} \end{cases} \quad (7)$$

3.3 填充剩余视差点

在 3.2 节得到了去除了错误匹配点的视差图,但是在去除了错点之后,这些点全部被赋值为 0, 这样就不是完整的视差图,下面利用视差图的顺序约束给出填充这些剩余点视差值的方法。

当以右图像为匹配基元时,设在视差图上查找到长度为 l , 值全部为 0, 即没有匹配的一段像素点。设起始点为 $g(x_i, y_j)$, 其前一点视差为 $\text{disp}(x_i - 1, y_j)$; 终止点为 $g(x_i + l, y_j)$, 其后一点的视差为 $\text{disp}(x_i + l + 1, y_j)$, 那么在左图像上对应的搜索区域则为 $[\text{disp}(x_i - 1, y_j) + x, \text{disp}(x_i + l + 1, y_j) + x_i + l]$, 这样就可以使用加窗的快速 ZNCC 方法进行匹配并填充剩余点。

$$\text{cost} = \min \left(\frac{\text{cov}_{i,j,d}(f, g)}{\text{var}_{i,j}(f) \times \text{var}_{i,j}(g)} \right) \quad (8)$$

式中

$$\text{cov}_{i,j,d}(f, g) = \sum_{m=i-M}^{i+M} \sum_{n=j-N}^{j+N} (f_{m,n} - \bar{f})(g_{m+d,n} - \bar{g})$$

$$= \sum_{m=i-M}^{i+M} \sum_{n=j-N}^{j+N} f_{m,n} \times g_{m+d,n} (2M + 1) \times (2N + 1) \bar{f} \times \bar{g} \quad (9)$$

$$var_{i,j}^2(f) = \sum_{m=i-M}^{i+M} \sum_{n=j-N}^{j+N} (f_{m,n} - \bar{f})^2 \quad (10)$$

$$var_{i,j}^2(g) = \sum_{m=i-M}^{i+M} \sum_{n=j-N}^{j+N} (g_{m+d,n} - \bar{g})^2 \quad (11)$$

通过 ZNCC 匹配可得到完整的视差图,最后使用 Birchfield 等人提出的方法^[8],进行后期处理,得到最终的视差图。

4 实验结果及分析

使用日本筑波大学的 tsukuba 图像对,以及德国波恩大学的 corridor 图像对作为本文算法的测试图像,因为 tsukuba 图像对中含有较多的物体,图像

的视差变化比较大,corridor 图像对中物体较少,图像的视差变化平缓。给出了本文结果与文献[8]方法结果图(图 3(d))和真实的视差图(图 3(e))的对比,并在表 1 中给出了对比数据。由图 3(b)和图 3(c)可以看出,当图像消除了错误匹配点之后,视差图中还含有较多的未匹配点,通过最后的填充,可以得到很好的视差图,并且与真实的视差图非常接近。利用文献[2]的评估方法,测试值 B 表示生成的视差图与真实视差图之间误匹配大于 δ_d 个像素的点在全图所占的比例,其中, d_T 为真实视差, d_C 为计算所得视差,

$$B = \frac{1}{N} \sum_{(x,y)} (|d_C(x,y) - d_T(x,y)| > \delta_d) \quad (12)$$

在实验过程中取 $\delta_d = 1.0$ 。

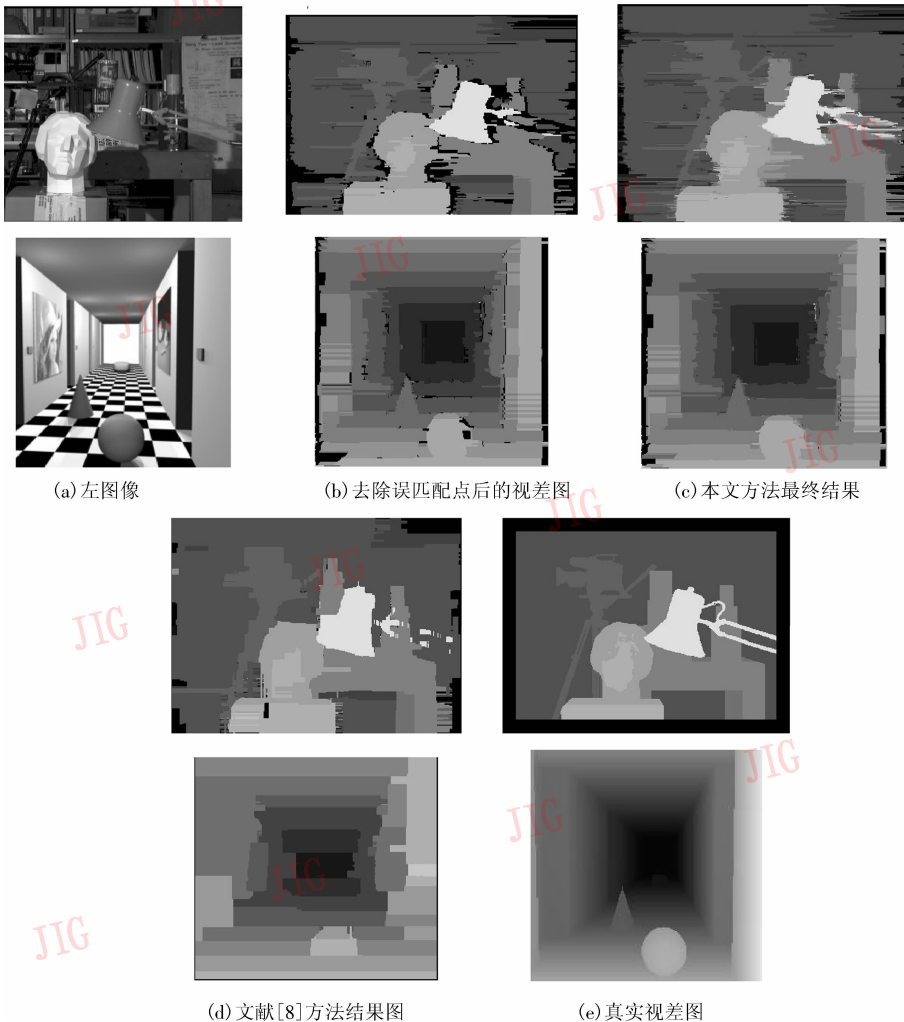


图 3 算法的实验结果与对比
Fig. 3 Experiment and comparison

表 1 本文算法及与文献[8]的误匹配率对比
Tab. 1 Comparison of our algorithm and Ref[8]

算法	tsukuba	corridor	单位: %
本文算法	5.35	2.25	
文献[8]算法	6.82	3.48	

5 结 论

在深入分析了现有算法的基础上,基于动态规划,提出了一种利用左右视差图一致性关系的算法,该算法能够去除错误的视差匹配点,并给出重新计算这些视差点的方法。算法在一定程度上能够很好地解决由动态规划所造成的部分错误匹配点,并达到得很好的效果。

参考文献 (References)

- 1 Brown Z, Burschka D, Hager G. Advance in computational stereo[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, **25**(8):993 ~ 1008.
- 2 Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithm [J]. International Journal of Computer Vision, 2002, **47**(1/2/3):7 ~ 42.
- 3 Bahat D, Nayar S. Ordinal measure for image correspondence[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, **20**(4):415 ~ 423.
- 4 Sun C. A fast stereo matching method [A]. In: Proceedings of Digital Image Computing: Techniques and Applications [C], Auckland, New Zealand, 1997:95 ~ 100.
- 5 Kluth V, Kunkel G, Rauhala U. Global least squares matching [A]. In: Proceedings of International Geosciences and Remote Sensing Symposium [C], San Diego, CA, USA, 1992:1615 ~ 1618.
- 6 Randriamasy S, Gagalowcz A. Region based stereo matching oriented image processing [A]. In: Proceedings of Computer Vision and Pattern Recognition [C], Maui, HI, USA, 1991:736 ~ 737.
- 7 Birchfield S, Tomasi C. A pixel dissimilarity measure that is insensitive to image sampling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, **20**(4):401 ~ 406.
- 8 Birchfield S, Tomasi C. Depth discontinuities by pixel-to-pixel stereo [J]. International Journal of Computer Vision, 1999, **35**(3):269 ~ 293.
- 9 Bobick A, Intille S. Large occlusion stereo [J]. International Journal of Computer Vision, 1999, **33**(3):181 ~ 200.
- 10 Sun J, Zheng N, Shum H. Stereo matching using belief propagation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, **25**(7):1 ~ 14.
- 11 Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, **26**(9):1124 ~ 1137.