

基于视差点的大遮挡检测和立体匹配方法*

文贡坚[†], 周秀芝

(国防科学技术大学 电子科学与工程学院 ATR 重点实验室, 湖南 长沙 410073)

Stereo Matching and Large Occlusion Detection Based on Disparity Points

WEN Gong-Jian[†], ZHOU Xiu-Zhi

(ATR Key Laboratory, School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: Phn: +86-731-4576343, E-mail: wengongjian@sina.com, <http://www.nudt.edu.cn>

Received 2004-04-07; Accepted 2004-07-06

Wen GJ, Zhou XZ. Stereo matching and large occlusion detection based on disparity points. *Journal of Software*, 2005,16(5):708-717. DOI: 10.1360/jos160708

Abstract: An algorithm based on disparity points to solve the occlusion problem in the process of building high-quality stereo disparity map is presented in this paper. It is firstly proved that the disparity curve corresponding to a pair of epipolar-line images may be approximated by a group of piece-wise straight lines, and then the definition of disparity point is introduced. In the parameterization of a disparity point, two parameters are used to describe left and right occlusions so that the occlusion problem can be successfully solved in the approach. By analyzing intensity property of a disparity point and its neighbor points, an approach which combines stepwise hypothesis-verification strategy and Marquardt-Levenberg (M-L) algorithm is devised to extract the candidate disparity points from the epipolar images, and then aperiodic dynamic programming is employed to search the epipolar-optimal disparity function. The proposed method is tested by using the international standard image data and compared with other methods, and the experimental results show that its performance is the best among epipolar-optimal methods and worse than some excellent global-optimal approaches, but its complexity is much lower than the global-optimal approaches.

Key words: stereo matching; occlusion detection; epipolar-optimal; hypothesis-verification strategy; dynamic programming; Marquardt-Levenberg algorithm

摘要: 提出一种基于视差点的方法来解决在高质量立体视差图生成过程中所出现的遮挡问题。首先证明同名核线对应的视差函数曲线可近似为一条分段直线,然后在此基础上引出视差点的概念。在视差点的结构中利

* Supported by the National Natural Science Foundation of China under Grant No.60275008 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2002AA135200 (国家高技术研究发展计划(863)); the ATR Key Laboratory Foundation of the National University of Defense Technology of China under Grant No.wkl(02)0101 (国防科学技术大学 ATR 重点实验室基金)

作者简介: 文贡坚(1972 -),男,湖南宁乡人,博士,副教授,主要研究领域为摄影测量与遥感,图像分析与理解,模式识别和信息融合;周秀芝(1977 -),女,博士生,主要研究领域为图像理解,信息融合。

用两个参数分别描述左右遮挡量,使得所提出的方法能够很好地解决遮挡问题.通过分析视差点及其邻域的灰度特性,提出一种分层假设证实和 Marquardt-Levenberg(M-L)算法相结合的方法从同名核线图像中提取出候选视差点,然后采用不定期的动态规划(dynamic programming,简称 DP)算法获得核线最优的视差函数.利用国际标准数据对提出的方法进行了测试,并与其他方法作比较,实验结果表明,它的匹配效果是目前核线最优方法中最好的,仅差于几种优秀的全局最优方法,但其计算复杂度要远低于全局的方法.

关键词: 立体匹配;遮挡检测;核线最优;假设证实;动态规划;M-L 算法

中图法分类号: TP391 文献标识码: A

立体匹配是由二维图像获得三维物体结构信息的主要技术手段.早期的影像匹配算法主要是针对于小比例尺的影像,在这种情况下,物体之间的遮挡问题可以忽略,典型的算法有基于区域匹配和基于特征匹配的方法^[1].然而,随着图像比例尺的增大,物体之间的相互遮挡问题变得越来越严重,遮挡成为产生高质量立体视差图中必须解决的重要问题,近年来已有大量的研究者提出不少的算法来解决它^[2-11].

目前用于解决遮挡问题的方法大致可以分为两类^[2]:一类是核线最优的方法^[5,6],这类方法是在一对核线上计算通过所有匹配像对的最短路径,部分遮挡的问题通过由一幅图像的多个像点与另一副图像的一个像点匹配来解决;另一类是全局最优的方法^[2,9-11],这类方法是首先在整个图像上定义一个包含数据误差、模型约束以及遮挡在内的能量函数,然后采用最优化的方法获得全局最优的视差图.从产生视差图的效果上来说,全局最优的方法一般要优于核线最优的方法,但是核线最优的方法的计算复杂度要低于全局最优的方法.本文提出的算法属于第 1 类.

在现有核线最优的立体匹配算法中,有 3 个突出的问题:一是如何设置合理的遮挡匹配代价.目前,核线最优方法一般是将匹配代价设为常值,若该值太大,检测出来的遮挡部分要小,反之则大;二是用于最优求解的动态规划算法要求匹配遵循 Order 约束;三是求得的每个像素的视差值只能是整数.为了解决上述 3 个问题,本文提出了一种基于视差点的立体匹配方法.本文假设同名核线图像是一条空间分段直线在左右图像上的投影,由此推导出这条核线对应的视差函数可近似为一条分段直线,这样求解视差函数的关键是如何获得分段直线的转折点.基于此点,本文引出了视差点的概念,而 3 类特殊的视差点正好对应着分段直线的转折点.在描述视差点的 6 个参数中,融入了对左右遮挡的处理,使得提出的方法能够很好地解决遮挡问题.在分析视差点及其附近的灰度特性后,可以获得一个描述这种特性的能量函数,视差点就是这个能量函数的局部极小值点,本文首先用一种分层假设证实和 M-L 算法相结合的方法从左右核线影像中提取所有候选的视差点,然后采用不定期的动态规划算法获得核线最优的视差函数.

本文第 1 节给出视差点的定义.第 2 节描述从左右核线影像中提取视差点的方法.第 3 节给出实验结果.最后给出结论.

1 视差点的定义

本节,首先分析了同名核线对应的左右视差函数的曲线形状,然后讨论了视差函数在遮挡情况下的特性,由此引出了视差点的定义,将求解视差函数的问题转化为从核线影像上提取几类特殊的视差点的问题.

1.1 视差函数的性质

我们是在核线影像上进行匹配的.设点 P 是一个坐标为 (X, Y, Z) 的空间点,在左右图像中投影分别为点 $p_L(x_L, y)$ 和点 $p_R(x_R, y)$.由三角几何可以得到:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda \begin{bmatrix} x_L \\ y \\ -f \end{bmatrix} = \lambda \begin{bmatrix} x_R \\ y \\ -f \end{bmatrix} + \begin{bmatrix} B \\ 0 \\ 0 \end{bmatrix} \quad (1)$$

其中 f 为左右相机的焦距, B 为摄影基线, λ 是尺度因子,由上式可得:

$$\lambda = \frac{B}{x_L - x_R} \quad (2)$$

在式(1)和式(2)中, x_L, x_R 和 y 是图像坐标,而在数字影像中,通常采用扫描坐标来描述投影点位置,因此需要将扫描坐标转化为像平面坐标.设点 p_L 和 p_R 的扫描坐标分别为 (x_l, y') 和 (x_r, y') ,扫描坐标与像平面坐标之间有如下关系:

$$\begin{cases} x_L = x_{l0} + \mu \times x_l \\ x_R = x_{r0} + \mu \times x_r \end{cases} \quad (3)$$

其中 x_{l0}, x_{r0} 是左右图像左下角点在像平面坐标系下的横坐标, μ 是一个像素水平方向的大小,由此可得

$$\lambda = \frac{B}{(x_{l0} - x_{r0}) + \mu(x_l - x_r)} = \frac{B}{d_0 + \mu(x_l - x_r)} \quad (4)$$

其中 $d_0 = (x_{l0} - x_{r0})$.

定义左视差函数 $d_l(x_l)$ 和右视差函数 $d_r(x_r)$ 为

$$\begin{cases} d_l(x_l) = x_l - x_r \\ d_r(x_r) = x_r - x_l \end{cases} \quad (5)$$

由此可得高程函数 $Z(x)$ 与左右视差函数的关系如下:

$$\begin{cases} Z(x) = -\lambda f = -\frac{C_0 f}{1 + C_1 d_l(x)} \\ Z(x) = -\frac{C_0 f}{1 - C_1 d_r(x)} \end{cases} \quad (6)$$

其中 $C_0 = \frac{B}{d_0}, C_1 = \frac{\mu}{d_0}$.

定理 1. 若同名核线对应的空间曲线是一条分段的空间直线,那么它对应的左右视差函数可近似认为是一条分段直线.

证明:将式(6)作泰勒展开得:

$$Z(x) = -C_0 f + C_0 f C_1 d_l(x) - 2C_0 f C_1^2 d_l^2(x) + \dots \quad (7)$$

为了保证左右图像覆盖着同一区域,应有 $x_{l0} > x_{r0}$ 且要求 $|\alpha/d_0| \ll 1$, 即 $|C_1| \ll 1$, 这样可忽略上式的高次项,可得

$$Z(x) \approx -C_0 f + C_0 f C_1 d_l(x) \quad (8)$$

由上式可知,高程函数 $Z(x)$ 与左视差函数 $d_l(x)$ 之间近似为线性关系.因此,若同名核线对应的空间曲线是一条分段的空间直线,则左视差函数也可近似为一条分段直线.

若采用像点的像平面坐标值,根据式(2),一个点的视差值与它的高度值(或深度值)成反比,这与文献中[1]的结论一致.但在实际处理中,一般采用像点的扫描坐标,由上面的推导可知,视差值与高度值之间的关系近似为线性.

假设左视差函数 $d_l(x), x \in X_l$ 由 N 条分段直线组成,包括两个端点在内共有 $N+1$ 个分割点 $p_{l_i} (0 \leq i \leq N)$, 其坐标分别为 $x_{l_i} (0 \leq i \leq N)$, 对应的左视差值分别为 d_{l_i} , 则左视差函数 $d_l(x)$ 可表示如下:

$$d_l(x) = d_{l_k} + m_{l_k} \times (x - x_{l_k}), x \in X_l \quad (9)$$

其中 k 的选择使得 x 落在 x_{l_k} 和 $x_{l_{k+1}}$ 之间, m_{l_k} 是由分割点 p_{l_k} 和 $p_{l_{k+1}}$ 组成的直线 l_{l_k} 的斜率,即

$$m_{l_k} = \frac{d_{l_{k+1}} - d_{l_k}}{x_{l_{k+1}} - x_{l_k}} \quad (10)$$

同样,我们可以定义右视差函数 $d_r(x), x \in X_r$,

$$d_r(x) = d_{r_k} + m_{r_k} \times (x - x_{r_k}), x \in X_r \quad (11)$$

其中 $x_{r_i}, m_{r_i}, d_{r_i}$ 与左视差函数中的 $x_{l_i}, m_{l_i}, d_{l_i}$ 意义相同.

由左、右视差函数的定义,可以容易得到左、右视差函数之间关系如下:

$$\begin{cases} x_{r_i} = x_{l_i} - d_{l_i} \\ d_{r_i} = -d_{l_i} \\ m_{r_i} = m_{l_i} / m_{l_i} - 1 \end{cases} \quad (0 \leq i \leq N) \quad (12)$$

1.2 视差点的定义

1.2.1 视差函数曲线上点的特性

由上面公式可知,左、右视差函数曲线上的任何一点总可以恢复一个空间点的位置,反过来,空间曲线上的任何一点能否在左、右视差函数曲线上找到其对应点,其对应点具有什么性质呢?下面,我们分 4 种情况来讨论这一问题:

第 1 种情况是空间点落在一段空间直线上,但不是两条相邻直线的交点;高程函数在它的左右导数是相等的,由式(8)可得,视差函数在其对应点上的左右导数也是相等的,即落在视差函数曲线一段直线上;

第 2 种情况是空间点是两相邻空间直线的交点;此时高程函数在它的左右导数是不相等的,视差函数在其对应点上的左右导数也不相等,即为视差函数曲线上两相邻分段直线的交点;

第 3 种情况是空间点在右影像上存在投影,而在左影像上的投影被其他空间点遮挡,称这种情况为左遮挡;下面我们来分析左遮挡情况下视差函数曲线上点的特性。

我们假设存在两个三维空间点 $P_1(X_1, Y_1, Z_1)$ 和 $P_2(X_2, Y_2, Z_2)$, 它们对应在右图像上的投影点分别为 $p_1(x_1, y)$ 和 $p_2(x_2, y)$ 。如果 P_1, P_2 之间存在左遮挡,那么 3 个空间点 S_1, P_1, P_2 必须共线,即

$$\frac{X_2}{X_1} = \frac{Y_2}{Y_1} = \frac{Z_2}{Z_1}.$$

利用上面公式,化简得到:

$$x_2 - x_1 = d_r(x_2) - d_r(x_1).$$

上式意味着右视差函数在 x_1, x_2 之间的斜率 $m_r = 1$, 根据式(12)可知, $m_l = \pm\infty$, 这样,在右视差函数一阶导数等于 1 和左视差函数发生跳跃的位置出现左遮挡。

第 4 种情况是空间点在左影像上存在投影,而在右影像上的投影被其他空间点遮挡,称这种情况为右遮挡;同样可得,在左视差函数一阶导数等于 1 和右视差函数发生跳跃的位置出现右遮挡。

根据上述分析,视差函数曲线上的点大致可分为 4 类:第 1 类点是视差函数在这些点的左导数等于右导数,且该导数值不等于 1 或 $\pm\infty$, 我们称这类点为平滑视差点;第 2 类点是它的左导数不等于右导数,且左、右导数值都不等于 1 或 $\pm\infty$, 我们称这类点为转折视差点;第 3 类点是左视差函数在该点的导数等于 $\pm\infty$ 或右视差函数在该点的导数等于 1, 在这种情况下,可以将其视为一个整体,当作一个点来对待,我们称这类点为左遮挡视差点;第 4 类点是左视差函数在该点的导数等于 1 或右视差函数在这些点的导数等于 $\pm\infty$, 同样也可以将其视为一个整体,当作一个点来对待,我们称这类点为右遮挡视差点。从物理意义上讲,由平滑视差点换算出来的空间点是落在空间直线上;由转折视差点换算出来的空间点是相邻两条空间直线的交点;左(或右)遮挡视差点表示空间一段区域在左右图像上的投影,它在左(或右)影像上只有一个投影点,但在右(或左)影像投影是一段区域,也就是说,空间区域内部的点在左(或右)影像上的投影不存在,即发生左(或右)遮挡现象,因此只能恢复出这段空间区域的两个端点位置。

1.2.2 视差点的定义

为了用统一的数据结构来描述 4 类视差点,我们首先分别用两个参数描述它们在左、右核线上的投影:一个是投影的坐标,另外一个为投影的长度,然后用两个参数描述左视差函数在左坐标处的左导数和右坐标的右导数。

定义 1. 设 Ω 表示空间中一点或一个区域,它在左或右核线上的投影只有一个点,那么 Ω 对应在视差函数的位置就称为一个视差点,如图 1 所示,该视差点用 6 个参数来描述: Ω 在左核线上投影 $a_{LL}a_{LR}$ 的左坐标 x_l 及长度 Δx_l , Ω 在右核线上投影 $a_{RL}a_{RR}$ 的右坐标 x_r 及长度 Δx_r , 以及左视差函数在 x_l 的左导数 m_{ll} 和在 $x_l + \Delta x_l$ 处的右导

数 m_r , 记视差点为 DP , 它的参数可用向量 $f_{DP} = [x_l, \Delta x_l, x_r, \Delta x_r, m_l, m_r]^T$ 表示, $[]^T$ 表示对向量进行转置.

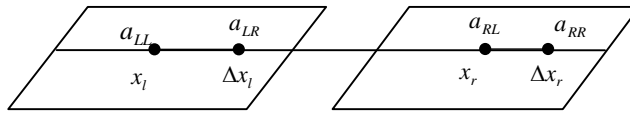


Fig.1 Structure of a disparity point

图 1 视差点的结构

视差点的参数必须满足下面的约束条件,且可以利用视差点的 6 个参数对它进行分类.

- (I) 视差点的约束条件 1: 视差点的参数 Δx_l 和 Δx_r 中至少有一个为 0;
- (II) 若视差点的参数满足 $\Delta x_l = \Delta x_r = 0$ 且 $m_l = m_r$, 那么该点为平滑视差点;
- (III) 若视差点的参数满足 $\Delta x_l = \Delta x_r = 0$ 且 $m_l \neq m_r$, 那么该点为转折视差点;
- (IV) 若视差点的参数满足 $\Delta x_r > 0$ 且 $\Delta x_l = 0$, 则该点为左遮挡视差点;
- (V) 若视差点的参数 $\Delta x_l > 0$ 且 $\Delta x_r = 0$, 则该点为右遮挡视差点.

视差函数曲线上各分段直线的连接点是 3 类特殊的视差点, 即转折视差点、左遮挡视差点和右遮挡视差点, 因此求解视差函数的关键是如何从左、右核线影像上提取出 3 类特殊视差点.

2 提取特殊视差点及视差函数

我们首先分析了视差点及其邻域范围内的灰度特性, 然后设计一种组合 M-L 算法和分层假设证实的方法从左、右核线图像中提取出所有 3 类候选的特殊视差点, 最后用不定期的动态规划选出正确的特殊视差点, 以得到核线最优的视差函数.

2.1 视差点的灰度特性

令 $g_L(x)$ 和 $g_R(x)$ 分别为左右核线图像的灰度函数, 假设在匹配点处灰度差是高斯白噪声, 即

$$g_L(x) = g_R(x - d_l(x)) + n(x) \quad (13)$$

其中 $n(x) \sim N(0, \sigma_n^2)$.

设点 P 是一个视差点, 它的参数为 $f_p = [x_l, \Delta x_l, x_r, \Delta x_r, m_l, m_r]^T$. 假设点 P 的左边存在 M 个平滑视差点, 那么 M 个平滑视差点的灰度差为

$$v_i = g_L(x_l + i) - g_R[x_r + (1 - m_l) \times i] \quad (i = -1, \dots, -M) \quad (14)$$

同样, 假设点 P 的右边存在 N 个平滑视差点, 那么 N 个平滑视差点的灰度差为

$$v_i = g_L(x_l + \Delta x_l + i) - g_R[x_r + \Delta x_r + (1 - m_r) \times i] \quad (i = 1, \dots, N) \quad (15)$$

(I) 视差点的约束条件 2: 视差点左边灰度差 v_i ($i = -1, \dots, -M$) 和右边灰度差 v_i ($i = 1, \dots, N$) 均应服从均值为 0, 方差为 σ_n^2 的高斯分布. 从统计角度而言需要进行两个参数假设检验: 一是在方差未知的情况下, 均值是否等于 0, 采用 t 检验; 二是在均值未知的情况下, 方差是否小于等于 σ_n^2 , 采用 χ^2 检验^[12].

(II) 视差点的约束条件 3: 左、右灰度差的平方和为一个局部极小值, 即

$$E = \sum_{i=-M}^{-1} v_i^2 + \sum_{i=1}^N v_i^2 \rightarrow \min \quad (16)$$

上式中能量值 E 显然是视差点参数 f_p 的函数, 因此可以等价写成

$$DP = \text{Arg min}(E(f_p)) \quad (17)$$

2.2 在初值附近搜索视差点

由视差点的约束条件 3 可知, 视差点对应着能量函数 $E(f_p)$ 的局部极小点. 式(17)是一个典型的非线性最小二乘优化问题, 非常适合采用 M-L 算法^[13]求解.

2.2.1 Marquardt-Levenberg 迭代方程

令 $f_{DP}^{(k)} = [x_l^{(k)} \quad \Delta x_l^{(k)} \quad x_r^{(k)} \quad \Delta x_r^{(k)} \quad m_{ll}^{(k)} \quad m_{lr}^{(k)}]^T$ 为最优解 f_{DP} 的第 k 次近似, $X^{(k)}$ 是 f_{DP} 的第 k 次修改量. Marquardt-Levenberg 迭代方程为下式:

$$X^{(k)} = (A^{(k)T} A^{(k)} + \alpha^{(k)} I)^{-1} A^{(k)T} L \quad (18)$$

其中

$$\left\{ \begin{array}{l} X^{(k)} = [dx_l \quad d\Delta x_l \quad dx_r \quad d\Delta x_r \quad dm_{ll} \quad dm_{lr}]^T \\ V^{(k)} = [v_{l-M}, \dots, v_{l-1}, v_{r_1}, \dots, v_{r_N}]^T \\ A^{(k)} = \begin{bmatrix} \dot{g}_L(x_{ll-M}) & 0 & -\dot{g}_R(x_{lr-M}) & 0 & -M\dot{g}_R(x_{lr-M}) & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dot{g}_L(x_{ll-1}) & 0 & -\dot{g}_R(x_{lr-1}) & 0 & -\dot{g}_R(x_{lr-1}) & 0 \\ \dot{g}_L(x_{rl_1}) & \dot{g}_L(x_{rl_1}) & -\dot{g}_R(x_{rr_1}) & -\dot{g}_R(x_{rr_1}) & 0 & \dot{g}_R(x_{rr_1}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dot{g}_L(x_{rl_N}) & \dot{g}_L(x_{rl_N}) & -\dot{g}_R(x_{rr_N}) & -\dot{g}_R(x_{rr_N}) & 0 & N\dot{g}_R(x_{rr_1}) \end{bmatrix}, \\ L^{(k)} = [\Delta g_{l-M}, \dots, \Delta g_{l-1}, \Delta g_{r_1}, \dots, \Delta g_{r_N}]^T \end{array} \right.$$

I 是六阶单位矩阵, 自由参数 $\alpha^{(k)} > 0$ 决定修改量 $X^{(k)}$ 受高斯-牛顿法还是最速下降法的影响程度.

2.2.2 分步假设证实求解视差点的方法

直接利用上述方法求得的视差点只满足它的约束条件 3, 而不一定满足视差点的约束条件 1 和条件 2. 我们设计了一种分步假设证实方法, 由它求出的视差点能够满足 3 个约束条件. 在整个求解过程中, 参数 x_l 或 x_r 固定不动, 在下面叙述中, 我们假设参数 x_l 固定不动. 方法分为 4 步:

第 1 步. 假设在初值点附近存在一个平滑视差点. 由于这类视差点要求参数 Δx_l 和 Δx_r 均等于 0, 且 $m_{ll} = m_{lr}$, 这时候需要求解的参数只有两个, 即 x_r, m_{ll} . 获得近似结果后, 测试是否满足约束条件 2: 若满足, 就在初值点附近找到一个平滑视差点, 否则进入第 2 步.

第 2 步. 假设在初值点附近存在一个转折视差点. 与第 1 步类似, 不过此时需要求解的参数有 3 个, 即 x_r, m_{ll}, m_{lr} . 获得近似结果后, 测试是否满足约束条件 2: 若满足, 就在初值点附近找到一个转折视差点; 否则进入第 3 步.

第 3 步. 假设在初值点附近存在一个左遮挡视差点. 此时需要求解的参数有 4 个, 即 $x_r, m_{ll}, m_{lr}, \Delta x_r$. 获得近似结果后, 测试结果是否满足约束条件 2: 若满足, 就在初值点附近找到一个左遮挡视差点; 否则进入第 4 步.

第 4 步. 假设在初值点附近存在一个右遮挡视差点. 此时需要求解的参数有 4 个, 即 $x_r, m_{ll}, m_{lr}, \Delta x_l$. 获得近似结果后, 测试结果是否满足约束条件 2. 若满足, 就在初值点附近找到一个右遮挡视差点; 若不满足, 则在初始点附近不存在任何视差点. 导致这种结果出现的原因是由于我们在求解过程中将参数 x_l 或 x_r 固定不动, 若初始点在左核线上落入被遮挡的区域内, 这时其左、右两边的点在右核线上找不到同名点, 从而导致匹配失败.

2.3 从左、右核线影像上提取所有特殊视差点

上面的方法是在一个初值点的附近搜索视差点, 为了从左、右核线影像上提取出所有的 3 类特殊视差点, 直观的方法是将左核线上的每个点与右核线上与其视差小于最大视差值的点组成一个初值点, 然后找出所有的视差点. 假设左核线有 N_L 个像素, 左、右核线的最大视差值为 D_{\max} , 那么按照上面的方法可以组成 $N_L \times D_{\max}$ 个初值点, 也就需要进行 $N_L \times D_{\max}$ 次迭代求解 $N_L \times D_{\max}$ 个局部极小值点, 计算量非常大, 而且这些视差点大部分是重复的, 这样还需要剔除重复的视差点.

我们只要求提取 3 类特殊的视差点, 视差函数在这些点上发生转折或跳变. 一般视差的不连续点对应灰度边缘点. 因此我们可以假设视差点的 4 个坐标点至少有一个边缘点, 从而获得从左、右核线影像上提取所有 3 类特殊视差点的方法.

第 1 步. 从左、右核线影像上提取灰度跳变点 x_l ;

第 2 步.在视差允许的范围内,与右核线上的每一点组成一个初值点,采用固定 x_l 的方法搜索视差点;

第 3 步.在视差允许的范围内,右核线影像上的灰度跳变点 x_r 与左核线上的每一点组成一个初值点,采用固定 x_r 的方法搜索视差点;

第 4 步.剔除从上面两步中得到的视差点中的重复点,剩下的视差点就是从左、右核线影像上提取的所有 3 类特殊视差点.

2.4 利用 DP 算法求解最佳视差函数

一个视差点对应能量函数的一个局部极小值点,但反过来,能量函数的一个局部极小值点不一定是视差点,因此,从上面方法得到的结果中有些点并不一定是真正的视差点.为了从提取出来的视差点中挑选合适的视差点来恢复视差函数,使得视差函数在某种准则下达到最佳,我们需要解决两个问题:一个是如何定义这个准则;另一个是采用什么样的优化算法来求解.

2.4.1 匹配代价

从上面的分析可知,视差函数曲线上两个相邻特殊视差点之间的区域是相互匹配的,也就是说,假设两个视差点为 DP_1, DP_2 , 它们的参数分别为 $f_{DP_1} = [x_{l_1}, \Delta x_{l_1}, x_{r_1}, \Delta x_{r_1}, m_{ll_1}, m_{lr_1}]^T$ 和 $f_{DP_2} = [x_{l_2}, \Delta x_{l_2}, x_{r_2}, \Delta x_{r_2}, m_{ll_2}, m_{lr_2}]^T$, 那么左核线 $x_{l_1} + \Delta x_{l_1}$ 与 x_{l_2} 之间所夹的影像与右核线 $x_{r_1} + \Delta x_{r_1}$ 与 x_{r_2} 之间所夹的影像是相互匹配的.现定义两相邻视差点之间的灰度匹配代价如下:

$$C_M(f_{DP_1}, f_{DP_2}) = \sum_{x=x_{l_1}+\Delta x_{l_1}}^{x_{l_2}} \left(g_L(x) - g_R \left(x_{r_1} + \Delta x_{r_1} + \frac{(x_{r_2} - x_{r_1} - \Delta x_{r_1})}{(x_{l_2} - x_{l_1} - \Delta x_{l_1})} (x - x_{l_1} - \Delta x_{l_1}) \right) \right)^2 \quad (19)$$

一个视差函数的匹配代价是各段分段直线的匹配代价和.假设视差函数曲线由 n 个特殊视差点 $\Phi = \{DP_i (i=1,2,\dots,n)\}$ 连接而成,它们的参数分别为 $f_{DP_i} = [x_{l_i}, \Delta x_{l_i}, x_{r_i}, \Delta x_{r_i}, m_{ll_i}, m_{lr_i}]^T$, 并已按照参数 x_{l_i} 由小到大的顺序排列.我们定义该视差函数的代价函数如下:

$$C(\Phi) = \sum_{i=1}^{n-1} C_M(f_{DP_i}, f_{DP_{i+1}}) \quad (20)$$

2.4.2 利用 DP 算法求解

假设采用上节的方法从左、右核线影像中提取了 N_D 个特殊视差点,记它们为 $\Phi_{DP} = \{DP_1, DP_2, \dots, DP_{N_D}\}$, 每个视差点的参数 $DP_i = \{x_{l_i}, \Delta x_{l_i}, x_{r_i}, \Delta x_{r_i}, m_{ll_i}, m_{lr_i}\} (i=1,2,\dots,N_D)$. 现在要解决的问题是从 N_D 个视差点中挑选出若干个视差点,由这些视差点恢复的视差函数使得匹配代价函数(20)最小,即

$$\Phi_{opt} = \text{Arg min}(C(\Phi)) \quad (21)$$

其中 Φ_{opt} 就是从 N_D 个视差点中挑选出的最佳视差点集合.

假设左、右核线的两个端点是匹配的,故在集合 Φ_{DP} 中再加入两个视差点:一个是 DP_0 , 参数为 $f_{DP_0} = [1,0,1,0,0,0]^T$, 它表示左核线的第 1 点与右核线的第 1 点匹配;另一个是 DP_{N_D+1} , 参数为 $f_{DP_{N_D+1}} = [x_L, 0, x_R, 0, 0, 0]^T$, 它表示左核线的最后一点与右核线的最后一点匹配,其中 x_L, x_R 分别表示左、右核线的长度.这样,集合 Φ_{DP} 变为 $\Phi'_{DP} = \{DP_0, DP_1, \dots, DP_{N_D}, DP_{N_D+1}\}$.

Φ'_{DP} 中任何两个视差点都有一条路径联结,为了满足 Order 约束,这种联结应该是单向的.因此,我们要寻找一条从点 DP_0 到点 DP_{N_D+1} 最短的路径,这条路径所经过的视差点就是所求的最佳视差点集合 Φ_{opt} . 定义从视差点 DP_i 到 DP_j 的路径长度如下:

$$C_{Path}(f_{DP_i}, f_{DP_j}) = \begin{cases} +\infty, & \text{如果 } DP_i \text{ 与 } DP_j \text{ 不满足 Order Constraint} \\ C_M(f_{DP_i}, f_{DP_j}). & \end{cases} \quad (22)$$

如果所求路径是经过视差点 DP_i 到视差点 DP_j , 那么为了满足 Order 约束,参数 f_{DP_i} 与 f_{DP_j} 应满足下式:

$$x_{l_i} < x_{l_j} \text{ 且 } x_{l_i} + \Delta x_{l_i} < x_{l_j} + \Delta x_{l_j} \text{ 且 } x_{r_i} < x_{r_j} \text{ 且 } x_{r_i} + \Delta x_{r_i} < x_{r_j} + \Delta x_{r_j} \quad (23)$$

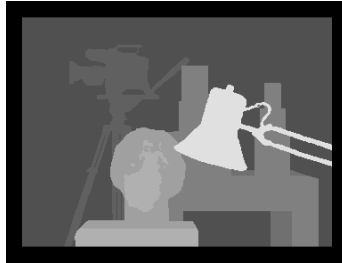
这样,式(21)是一个不定期最优路径问题,可以采用不定期的动态规划算法求解^[14].

3 实验结果

立体匹配是计算机视觉研究的一个重点,近年来出现了许多方法.为了客观地评价这些算法,文献[2]给出了4幅测试图像,并对现有各种算法进行了比较,关于它们的详细介绍见网址 <http://www.middlebury.edu/stereo>.在这一部分,我们利用这些数据测试了本文的方法,得到的实验结果如图2所示.



(a) Tsukuba image
(a) Tsukuba 图像



(b) True disparities
(b) 真实视差图



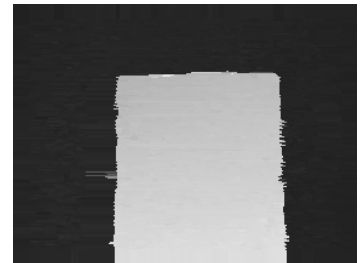
(c) Our results
(c) 本文得到的结果



(d) Map images
(d) Map 图像



(e) True disparities
(e) 真实视差图



(f) Our results
(f) 本文得到的结果



(g) Venus image
(g) Venus 图像



(h) True disparities
(h) 真实视差图



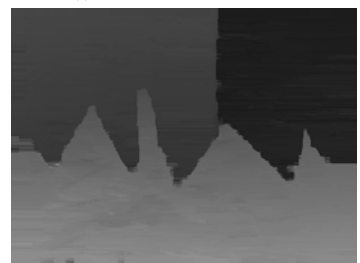
(i) Our results
(i) 本文得到的结果



(j) Sawtooth image
(j) Sawtooth 图像



(k) True disparities
(k) 真实视差图



(l) Our results
(l) 本文得到的结果

Fig.2
图 2

文献[2]中还列出了3个统计指标对各种算法进行了量化比较,这3个量化指标分别为 B_O (非遮挡区域错误匹配百分比)、 B_T (非纹理区域错误匹配百分比)、 B_D (视差不连续区域错误匹配百分比),它们的定义见文献[2].表1给出了多种匹配算法对上述4幅测试图像的3个统计指标比较结果,除了本文方法,其他数据来源于文献[2].

从表1(带有“*”的算法是核线最优的方法,其他属于全局最优算法)可以看出,本文方法的统计误差 B_O 和 B_T 要小于核线最优算法,而误差 B_D 则与它们相当,这表明本文算法是现有核线最优算法中是比较优秀的.而相对于全局算法 Stoch.diff, Graph Cut, Belief prop, 本文算法得到的统计误差是比较大的;这表明本文算法与全局最优的方法相比还有一定差距.

文献[2]中没有进行匹配算法的计算效率评估,因此我们只与 Belief prop^[11]方法作了计算效率比较.对于测试图像 Tsukuba,在 CPU 配置为 PVI1.7GHZ 下, Belief.prop 算法达到表1所示的指标需要 51.6s(文中只给出测试图像 Tsukuba 的结果),而本文的方法只需要 1.86s,这说明本文方法的计算效率是较高的.

Table 1 The results of different stereo matching algorithms testing by four standard images

表1 不同立体匹配算法对4幅标准的测试结果比较

Algorithm	Tsukuba			Sawtooth			Venus			Map	
	B_O	B_T	B_D	B_O	B_T	B_D	B_O	B_T	B_D	B_O	B_D
Our methods*	2.40	1.92	10.26	2.66	1.31	14.14	3.82	5.57	20.36	0.81	10.48
Scanl.opt ^{[2]*}	5.08	6.78	11.94	4.06	2.64	11.90	9.44	14.59	18.20	1.84	10.22
Pix-to-pix ^{[2]*}	5.12	7.06	14.62	2.31	1.79	14.93	6.30	11.37	14.57	0.50	6.83
Dyn.prog ^{[2]*}	4.12	4.63	12.34	4.84	3.71	13.26	10.10	15.01	17.12	3.33	14.07
Stoch.diff ^[10]	3.95	4.08	15.49	2.45	0.90	10.58	2.45	2.41	21.84	1.31	7.79
Graph cut ^[9]	1.86	1.00	9.35	0.42	0.14	3.76	1.69	2.30	5.40	2.39	9.35
Belief.prop. ^[11]	1.15	0.42	6.31	0.98	0.30	4.83	1.00	0.76	9.13	0.84	5.27

4 结 论

本文提出了一种基于视差点的立体匹配方法,试图解决产生高质量立体视差图中出现的遮挡问题.本文在详细描述视差点的概念后,重点研究了从一对核线图像上提取视差点的方法,以得到核线最优的视差函数.该方法基本上能够克服现有核线最优方法中存在的几个问题:(1) 由于遮挡量包含在视差点的结构内,因此求解最优视差函数不必考虑遮挡部分的匹配代价;(2) 由于视差点的结构中引入了视差函数的斜率,因此每个像素的视差值不一定是整数;(3) 利用 DP 算法从候选视差点中选取合适视差点,与以往核线最优方法相比较,对 Order 约束条件有所放宽.实验结果表明,本文的方法与现有核线最优算法相比较,得到的视差图质量较高.虽然它与一些优秀的全局最优方法相比还有一定的差距,但它的计算效率要高得多.我们将在获得图像上所有的候选视差点后,在全局范围内来求解最佳视差函数,以得到一种匹配质量高、速度快的立体匹配算法.

References:

- [1] Ma SD, Zhang ZY. Computer Visio - Computation Theory and Algorithm Foundation. Beijing: Science Press, 1998. 72-93 (in Chinese).
- [2] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int'l Journal of Computer Vision, 2002,47(1):7-42.
- [3] Zitnick CL, Kanade T. A cooperative algorithm for stereo matching and occlusion detection. IEEE Trans. on Pattern Anal. Mach. Intel., 2000,22(7):675-684.
- [4] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts. ICCV, 2001,II:508-515.
- [5] Bobick AF, Intille SS. Large occlusion stereo. Int'l Journal of Computer Vision, 1999,33(3):181-200.
- [6] Geiger D, Ladendorf B, Yuille A. Occlusions and binocular stereo. Int'l Journal of Computer Vision, 1995,14(3):211-226.
- [7] Birchfield S, Tomasi C. A pixel dissimilarity measure that is insensitive to image sampling. IEEE Trans. on Pattern Anal. Mach. Intel., 1998,20(4):401-406.
- [8] Cox JJ, Hingorani SL, Rao SB, Maggs BM. A maximum likelihood stereo algorithm. Computer Vision Image Understanding, 1996,63(3):542-567.

- [9] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Trans. on Pattern Anal. Mach. Intel., 2001,23(11):1222-1239.
- [10] Scharstein D, Szeliski R. Stereo matching with nonlinear diffusion. Int'l Journal of Computer Vision, 1998,28(2):155-174.
- [11] Sun J, Zheng NN, Shum HY. Stereo matching using belief propagation. IEEE Trans. on Pattern Anal. Mach. Intel., 2003,25(7): 787-800.
- [12] Wu Y, Li YL, Hu QJ. Application Statistic. Changsha: Press of National University of Defense Technology, 1999. 82-87 (in Chinese).
- [13] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes in C: The Art of Scientific Computing. London: Cambridge University Press, 1992. <http://www.nr.com>
- [14] Bertsekas DP. Dynamic Programming. Xi'an: Xi'an Jiaotong University Press, 1990. 50-70 (in Chinese).

附中文参考文献:

- [1] 马颂德,张正友.计算机视觉——计算理论与算法基础.北京:科学出版社,1998.
- [12] 吴翊,李永乐,胡庆军.应用数理统计.长沙:国防科学技术大学出版社,1999.
- [14] 帕塞克斯.动态规划.西安:西安交通大学出版社,1990.

2005 年全国软件与应用学术会议

征文通知

由中国计算机学会系统软件专业委员会和软件工程专业委员会联合主办,南京大学计算机软件新技术国家重点实验室及计算机科学与技术系承办的“2005 年全国软件与应用学术会议(NASAC 2005)”将于 2005 年 10 月 18 日~20 日在江苏南京召开。

一、征文范围 (但不限于下列内容)

- | | | |
|-----------------|---------------|---------------|
| 1.操作系统的理论与实践 | 6.分布式与嵌入式软件设计 | 11.软件开发方法及自动化 |
| 2.面向对象与软件 Agent | 7.软件工程工具与环境 | 12.软件理论与形式化方法 |
| 3.软件中间件与应用集成 | 8.软件体系结构与设计模式 | 13.软件质量、测试与验证 |
| 4.软件语言与编译技术 | 9.构件技术与软件复用 | 14.软件标准与规范 |
| 5.需求工程的方法与技术 | 10.软件过程管理与改进 | |

二、论文要求

1. 论文必须未在杂志和会议上发表和录用过。
2. 论文篇幅限定 6 页(A4 纸)内。
3. 应以 PDF 或 PS 格式提交论文。有关文章的版心等格式及参考文献样式将在会议网站上公布,敬请关注。

三、重要日期

1. 论文投稿截止日期(邮戳为准): 2005 年 5 月 30 日
2. 论文录用通知日期: 2005 年 7 月 10 日
3. 软件产品发布和演示报名截止日期: 2005 年 9 月 18 日
4. 学术会议及活动日期: 2005 年 10 月 18 日~20 日

四、联系方式

1. 会务组地址: 210093 江苏 南京 汉口路 22 号 计算机软件新技术国家重点实验室
 电话: 025-83593467; 025-83593670 传真: 025-83686596; 025-83593467 E-mail: keysoftlab@nju.edu.cn
 NASAC 2005 网址: <http://keysoftlab.nju.edu.cn/NASAC2005>

2. 联系人

- (1)论文接收: 陶先平 徐煊
- (2)其他会务: 武林红 胡又林 张建莹

3. 注意事项

接受电子投稿。投稿时,随稿附上您的详细地址、邮编、电话(办公室、实验室、家中)、手机以及 E-mail 地址。更详细的内容请访问 NASAC2005 网址。