

从双视图到多视图的协同优化 立体视觉匹配算法

王道累, 吴懋亮, 陈军

(上海电力学院 能源与机械工程学院, 上海 200090)

摘要: 采用基于视图协同优化算法解决视觉匹配问题, 通过对双视图立体匹配的研究到多视图匹配的推广, 利用视差平面拟合进行多个视差集的合并, 利用奇异值分解法解决线性方程, 建立排除异常点的判别规则, 解决遮挡及稀疏纹理的问题. 实验结果表明: 该协同优化算法是有效的, 提高了匹配的精度.

关键词: 立体匹配; 协同优化; 图像分割; 视差平面拟合

中图分类号: TP 391.41

文献标志码: A

20 世纪 80 年代, Marr 等^[1]提出了一种使计算机通过二维图像恢复三维场景信息的能力视觉计算理论, 并应用在双目或多目匹配上, 从而奠定了双目或多目立体视觉发展的理论基础. 立体匹配是计算机视觉关键技术之一, 现有解决立体匹配问题的算法主要分为两类: 基于局部约束的算法和基于全局约束的算法^[2-4]. Wu 等^[5]采用张量投票算法对 PMVS 中的特征扩展和噪声过滤进行优化. Nguyen 等^[6]提出局部稠密解码方法, 利用 45° 和 90° 两方向对该方法进行求导, 获得视差图. Xu 等^[7]提出优化局部自适应算法, 求解图像对视差. 上述算法都是基于局部算法求解视差, 但这类算法对图像噪声较敏感, 对重复纹理区域、遮挡区域和视差不连续区域等往往匹配效果不理想, 错误匹配率较高. 在此基础上, 很多计算机视觉研究专家提出基于全局约束算法, 对整个图像数据信息进行计算, 这样降低对局部区域敏感度, 解决重复纹理的影响, 同时, 也可以求解视差不连续区域, 最终达到全局最优解. 基于全局优化算法的本质是把匹配问题转化为求解能量函数, 然后, 通过求解该能量方程的全局最优解, 并可寻找到匹配点. 全局算法一般有动态规划算法、图割算法^[8]和人工智能算法等. 本文研究双视图到多视图立体匹配算法, 提出基于全局协同算法解决多视图的匹配问题, 从而准确获得并恢复三维场景深度.

1 双目立体视觉匹配

1.1 图像分割

在图像分割之前, 通过两个假设: 每个分割区域内的视差值变化细微, 平滑过渡; 只有在分割区域边界存在深度距离的不连续性. 采用 mean-shift 算法^[9]对参考图像(两幅图像任选一副作为参考图像)进行分割, 能很好满足图像的匹配假设条件.

1.2 视差的初始化

由于局部匹配算法速度快、运算简单, 局部匹配算法被用来计算视差初始值, 采用自适应相似算法来计算双视图匹配视差值. 自适应相似算法具体表示为

$$\text{SelfAd}(x, y, d) = (1 - t) | I(x, y) - I'(x + d, y) | +$$

收稿日期: 2015-05-07

通信作者: 王道累(1981-), 男, 讲师, 博士, 主要从事计算机视觉、图像处理和 CAD/CAM 的研究. E-mail: alfredwdl@shiep.edu.cn.

基金项目: 上海发电环保工程技术研究中心项目(11DZ2281700)

$$t(|\nabla_x I(x, y) - \nabla_x I'(x + d, y)| + |\nabla_y I(x, y) - \nabla_y I'(x + d, y)|). \quad (1)$$

式(1)中: I 和 I' 分别代表左右图像中像素坐标 (x, y) 处的灰度; t 表示灰度值差和梯度差之间的权值.

视差值初始化代价方程为

$$C_{\text{data}}(i, j, d) = \sum_{(i, j) \in (i, j)} \text{SelfAd}(x, y, d). \quad (2)$$

$N(x, y)$ 是以像素的 (x, y) 为中心 5×5 的矩形窗, 使用 WTA (winner-takes-all) 方法最小化代价方程, 求得像素点 (x, y) 的视差值.

1.3 视差图平面拟合

在参考图像分割之后, 以分割区域为单位进行视差图平面拟合. 由于假设在同一个分割区域内没有太大视差变化, 所以分割区域内每一个像素的视差可以用一个模型来表示. Tao 等^[10] 给出了视差图平面的拟合方程, 每一个分割区域视差模型表示为

$$d(x, y) = ax + by + c. \quad (3)$$

式(3)中: a, b, c 是平面的参数; d 是像素点 (x, y) 相对应的视差值.

对于分割区域内的各个像素点可以表示为

$$\mathbf{A}[a, b, c]^T = \mathbf{B}. \quad (4)$$

式(4)中: 矩阵 \mathbf{A} 第 i 行为 $[x_i, y_i, 1]$; 向量 \mathbf{B} 第 i 行为 $d(x_i, y_i)$.

对线性方程(4)进行变换, 可得

$$\mathbf{A}^T \mathbf{A}[a, b, c]^T = \mathbf{A}^T \mathbf{B},$$

即有

$$\begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i y_i & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i y_i & \sum_{i=1}^m y_i^2 & \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m y_i & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i d_i \\ \sum_{i=1}^m y_i d_i \\ \sum_{i=1}^m d_i \end{bmatrix}. \quad (5)$$

通过奇异值(singular value decomposition, SVD)求解, 有

$$[a, b, c]^T = \mathbf{A}^T \mathbf{A} + \mathbf{A}^T \mathbf{B}. \quad (6)$$

一般情况下, 线性方程(5)的对奇异点非常敏感, 所以要对奇异点进行排除. 这里制定 3 种规则来排除奇异点.

1) 视差相互比较. 对左右视图分别通过自适应相似算法就得视差图, 然后进行相互比较. 给定一个误差阈值 ξ , 如果小于该阈值 ξ , 该像素点为可靠的像素点.

2) 可靠区域判断. 在排除奇异点之后, 分割区域内的可靠像素点总数是否可以到达构建拟合平面的要求, 即构成可靠区域. 这里规定可靠像素点总数是整个区域像素点数一半以上.

3) 视差距离判断当前获得视差值与前一次获得值进行比较, 在规定的误差范围内, 该像素点为可靠像素点.

通过判断之后, 当前获得平面参数与前一次获得值进行比较, 即

$$|a - a'| + |b - b'| + |c - c'| < \epsilon. \quad (7)$$

如果在规定的误差内, 循环结束, 或者继续进行参数优化.

1.4 视图协同优化视差值

视图协同优化是将一个复杂目标分解为一些相对简单子目标, 相连区域对该子目标影响进行优化^[11]. 通过 mean-shift 分割后视图, 如图 1 所示. 图 1 中: s_1, s_2, \dots, s_n 为分割区域.

设 $E(x)$ 为整幅图像的能量方程, 根据协同优化原理, 有

$$E(x) = \sum_{j=1}^n E_j(x). \quad (8)$$

式(8)中: $E_j(x)(j=1,2,\cdots,n)$ 是第 j 个分割区域的能量方程.

将一个复杂的能量分解为相对简单的子能量方程,通过迭代方程,有

$$(1-\kappa)E_j(x)+\kappa_j\sum_{i\neq j}\mu_{ji}E_i(x),\quad j=1,2,\cdots,n.\quad (9)$$

式(9)中: $E_i(x)$ 表示第 i 个分割区域的能量方程,分割域 $s_j(j=1,2,\cdots,n;i\neq j)$ 是分割域 s_j 的相连域; $0\leq\kappa_j\leq 1, 0\leq\mu_{j,1}\leq 1$ 分别是相对应的权值.

利用协同原理来优化视差值,首先要构建视差拟合区域的能量方程,该能量方程有三部分组成分别是代价方程、惩罚方程和平滑度方程,具体表示为

$$E_j=E_{j_{\text{data}}}+E_{j_{\text{occ}}}+E_{j_{\text{sm}}}.\quad (10)$$

上式中的各项可分别表示为:

$$E_{j_{\text{data}}}=\sum_{(x,y)\in N_r}c_{\text{data}}(x,y,d);\quad (11)$$

$$E_{j_{\text{occ}}}=\omega_{\text{occ}}\cdot N_{\text{occ}};\quad (12)$$

$$E_{j_{\text{sm}}}=\sum_{(x,y)\in N_r}s_{\text{sm}}(x,y,d);\quad (13)$$

$$s_{\text{sm}}(x,y,d)=\begin{cases}\gamma\sum_{((x,y),(x',y'))\in N}\text{dis}((x,y),(x',y'))^{-1}\exp[-(\beta I_{(x,y)}-I_{(x',y')})^2],\\|d^j(x,y)-d^j(x',y')|\geq 1,\\0, \quad \text{或其他.}\end{cases}$$

各个参数定义如下: $c_{\text{data}}(x,y,d)$ 的定义同节 1.2; ω_{occ} 为遮挡惩罚系数; N_{occ} 为区域内遮挡像素数(通过上面 3 个规则判断得到的不可靠像素数); N_r 是可靠区域像素集; $I_{(x,y)}$ 和 $I_{(x',y')}$ 分别是像素点 (x,y) 和 (x',y') 的灰度; N 以像素的 (x,y) 为中心 5×5 的矩形窗; $\text{dis}(\cdot)$ 是两像素点欧几里得距离; γ 是一个常数; β 定义为

$$\beta=(2\langle(I_{(x,y)}-I_{(x',y')})^2\rangle)-1.\quad (14)$$

式(14)中: $\langle\rangle$ 表示在当前区域内的期望值^[12].

最后,通过 Powell 方法求解能量方程(9),获得第 j 个分割区域的最优视差平面参数(a_j,b_j,c_j),就可以获得优化后视差图.

2 多目立体视觉匹配

利用多视图进行立体匹配求视差,给匹配算法提供更多的信息,有利于提高算法的精度,降低匹配的错误率.选择其中一幅为参考图像,该参考图像与其他图像进行两两匹配求得视差图,形成了多个视差集($d^{(1)},d^{(2)},\cdots,d^{(n)}$),如何优化合并这一系列视差集将是多视图匹配的关键问题之一.

双视图立体算法有 3 个匹配步骤.

步骤 1 对参考图像进行分割.

步骤 2 采用自适应相似算法求初始化视差,并求得($d^{(1)},d^{(2)},\cdots,d^{(n)}$).

步骤 3 利用视差平面拟合对多视图视差集进行合并.

对于多视图视差平面拟合,平面模型表示为

$$d^{(i,j)}(x,y)=ax+by+c.\quad (15)$$

式(15)中:第 j 个分割域在第 i 个视差集中的模型.

对多视差集合并,参考图像第 j 个分割域和其他图像所形成视差,其区域内各个像素点的线性方程表示为

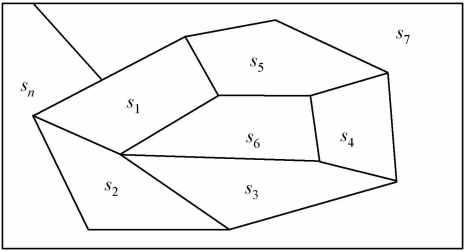


图 1 通过 mean-shift 分割后视图
Fig. 1 Segmented reference image

$$\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_p & y_p & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_p & y_p & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_p & y_p & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} d_1^{(1,j)} \\ d_2^{(1,j)} \\ \vdots \\ d_p^{(1,j)} \\ d_1^{(2,j)} \\ d_2^{(2,j)} \\ \vdots \\ d_p^{(2,j)} \\ d_1^{(n,j)} \\ d_2^{(n,j)} \\ \vdots \\ d_p^{(n,j)} \end{bmatrix}. \quad (16)$$

式(16)中: p 是在第 j 个分割域中像素点个数; $d_k^{(i,j)}$ 是第 i 个视差集的第 j 个分割域中第 k 个像素点的视差值 ($i=1,2,\dots,n; j=1,2,\dots,m$ 和 $k=1,2,\dots,p$); a, b, c 是第 j 个分割域的拟合平面系数。

从方程(16)获得到第 j 个分割域的拟合视差值是各个视差集中该分割域的平均值, 奇异值分解得到视差拟合平面的系数, 即

$$E_{j_{\text{data}}} = \sum_{i=1}^n \left(\sum_{(x,y) \in N_r} c_{\text{data}}(x,y,d_{(x,y)}^{(i,j)}) \right). \quad (17)$$

以此类推, 可以对参考图像的其他分割域进行同样平面拟合, 最后获得整副参考图像的视差图。

线性方程(16)进行奇异点的排除, 同理, 应用双目视图匹配的规则判断获得可靠分割区域. 采用协调优化方法对多视图视差进行优化, 多视图能量方程的建立, 具体定义为双目视图匹配(方程 10), 各个部分具体定义为

$$E_{j_{\text{occ}}} = \sum_{i=1}^n (\omega_{\text{occ}} \cdot N_{\text{occ}}^{(i,j)}), \quad (18)$$

$$E_{j_{\text{sm}}} = \sum_{i=1}^n \left(\sum_{(x,y) \in N_r} s_{\text{sm}}(x,y,d_{(x,y)}^{(i,j)}) \right). \quad (19)$$

各个参数定义如双视图匹配能量方程, 同理, 根据方程(9)定义, 利用 Powell 算法迭代优化得到最优多视图视差。

3 实验结果及讨论

3.1 匹配算法流程图

算法具体表述为如下 5 个步骤。

步骤 1 视图组输入。

步骤 2 选择参考图像进行分割处理。

步骤 3 运用局部自适应相似算法计算初始化视差。

步骤 4 视差平面拟合, 并根据规则排除奇异点建立可靠区域集。

步骤 5 利用协同优化视差平面参数获得最优视差值。

3.2 实验结果

对述的匹配算法的流程图进行运算, 双视图及多视图的视差图采用 Middlebury^[10] 的视图 Tsukuba, Venus, Cones 和 Teddy, 运行结果如图 2 所示. 图 2(a)为原始图(Tsukuba, Venus, Cones 和 Teddy), 图 2(b)为对应的 ground-truth 图, 图 2(c)为建议的双视图匹配算法结果图。

通过执行协同优化匹配算法获得 Tsukuba, Venus, Cones 和 Teddy 的多视图视差图, 如图 3 所示。

各双视图及多视图的算法结果, 如表 1 所示. 表 1 中: η 为错误匹配率. 由表 1 可知: 利用协同优化匹配算法对多视图获得错误匹配率最小, 双视图次之, 局部算法获得的初始化视差错误率最大. 这样也



图 2 双视图视差图

Fig.2 Disparity map of binocular stereo matching

充分说明了多视图能提高更多信息,有利于算法改进,减少错误的匹配. 局部算法对图像匹配比较敏感,速度虽快,单精确度劣于全局的匹配算法. 同时,也比较了 GF^[13],ST-1^[14]的双视图错误匹配率. 结果表明:匹配算法是可以获得较好的匹配数据,但在参考图 Cones 上,算法表现不够好,需要对处理遮挡问题作进一步优化.



图 3 多视图视差图

Fig.3 Disparity map of multi-ocular stereo matching

表 1 各方法求得视图差比较

Tab.1 Compare with results of stereo matching methods

方法	$\eta/\%$			
	Tsukuba	Venus	Cones	Teddy
初始化视差	1.434	1.04	7.76	7.564
双视图协同优化	1.233	0.87	6.94	6.453
多视图协同优化	1.023	0.52	4.87	4.342
GF ^[13]	2.28	0.91	2.90	8.300
ST-1 ^[14]	1.89	0.76	3.64	7.550

4 结束语

研究立体视觉匹配算法,利用协调优化对双视图及多视图进行匹配优化. 利用 SVD 方法对线性方程进行求解,同时,在多视图在通过视差平面拟合合并视差集. 制定排除奇异点的规则,能很好判断遮挡及非可靠像素点. 本算法不仅可以解决双视图匹配问题,也可以推广到多视图立体匹配问题的解决. 今后,将对算法进一步改进及提出更好的平面拟合算法.

参考文献:

[1] MARR D,POGGIO T. A computational theory of human stereo vision[C]// Proceedings of the Royal Society of London. London:IEEE Press,1979:301-328.

[2] FAUGERAS O, LONG Q. The Geometry of multiple images[M]. Cambridge:Cambridge University Press,2001:1-50.

[3] FURUKAWA Y,PONCE J. Accurate, dense, and robust multi-view steropsis[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence,2010,32(8):1362-1376.

- [4] LHUILLIER M, QUAN Long. Aquasi-dense approach to surface reconstruction from uncalibrated images[J]. Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(3): 418-433.
- [5] WU Tai-pang, YEUNG S K, YA Jia-jia, et al. Quasi-dense 3D reconstruction using tensor-based multi-view stereo [C]//Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco: IEEE Press, 2010: 1482-1489.
- [6] VINH D N, DUC D N, SANG J L, et al. Local density encoding for robust stereo matching[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(12): 2049-2062.
- [7] XU Ling-feng, OSCAR C A, SUN Wen-xiu, et al. Stereo matching with optimal local adaptive radiometric compensation[J]. IEEE Signal Processing Letters, 2015, 22(2): 131-135.
- [8] WANG Dao-lei, LIM K B. Obtaining depth map from segment-based stereo matching using graph cuts[J]. Journal of Visual Communication and Image Representation, 2002, 24(5): 603-619.
- [9] COMANICIU D, MEER P. Mean shift: A robust approach toward feature space analysis[J]. Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [10] TAO H, SAWHNEY H S, KUMAR R. A global matching framework for stereo computation[C]//International Conference on Computer Vision, Vancouver: IEEE Press, 2001: 532-539.
- [11] HUANG Xiao-fei. Cooperative optimization for energy minimization: A case study of stereo matching[EB/OL]. [2014-09-23]. <http://front.math.ucdavis.edu/author/X. Huang, cs. CV/071057>.
- [12] BLAKE A, ROTHER C, BROWN M, et al. Torr, interactive image segmentation using an adaptive GMMRF model [C]//Proceedings European Conference of Computer Vision, Prague, Czech Republic, 2004: 428-441.
- [13] RHEMANN C, HOSNI A, BLEYER M C, et al. Gelautz fast cost-volume filtering for visual correspondence and beyond[C]//Conference on Computer Vision and Pattern Recognition, Colorado Springs: IEEE Press, 2011: 3017-3024.
- [14] MEI Xing, SUN Xun, DONG Wei-ming, et al. Segment-tree based cost aggregation for stereo matching[C]//IEEE Conference on Computer Vision and Pattern Recognition, Portland: IEEE Press, 2013: 23-28.

An Algorithm of Stereo Matching Based on Cooperative Optimization from Two Views to Multiple Views

WANG Dao-lei, WU Mao-liang, CHEN Jun

(College of Energy and Mechanical Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: The algorithm of stereo matching using cooperative optimization is developed to extract disparity map from image pairs. The procedure of algorithm is generally extended to apply to multiple views from two views. The disparity sets are merged into one by the disparity plane estimation, and the linear equation is solved by singular value decomposition (SVD). And the rules are built to filter out outliers, which can effectively handle the problem of occlusion and textureless. The experimental results show that the effectiveness of our proposal methods and the accurateness of stereo matching.

Keywords: stereo matching; cooperative optimization; image segmentation; disparity plane fitting

(责任编辑: 陈志贤 英文审校: 杨建红)