

文章编号: 1001-0920(2007)01-0069-04

金字塔双层动态规划立体匹配算法

赵 杰, 于舒春, 蔡鹤皋

(哈尔滨工业大学 机器人研究所, 哈尔滨 150080)

摘 要: 针对控制点修正的动态规划立体匹配算法存在控制点求取时间长、实时性差的问题, 提出一种金字塔双层动态规划立体匹配算法. 采用金字塔算法求取低、高分辨率图像, 然后分别在低、高分辨率图像上求取候选控制点集和最终控制点集, 并用最终控制点集修正高分辨率图像上的动态规划立体匹配. 由于候选控制点集的求取在低分辨率图像上进行, 算法用时大为减少. 实验证明, 此算法匹配率高、速度快.

关键词: 控制点; 动态规划; 立体匹配; 视觉导航

中图分类号: TP242

文献标识码: A

Pyramid two-level dynamic programming stereo matching algorithm

ZHAO Jie, YU Shu-chun, CAI He-gao

(Robot Research Institute, Harbin Institute of Technology, Harbin 150080, China. Correspondent: YU Shu-chun, E-mail: yscjs@sohu.com)

Abstract: To the problems of dynamic programming stereo matching algorithm modified by ground control points (GCPs), a pyramid two-level dynamic programming stereo matching algorithm is proposed. The low-resolution images and high-resolution images are gained by pyramid algorithm. Then, candidate GCPs and ultima GCPs are obtained from low-resolution and high-resolution images. Dynamic programming stereo matching algorithm in high-resolution images is directed by ultima GCPs. Total time of this algorithm is reduced because candidate GCPs are gained in low-resolution images. Experiments show that the algorithm has a high matching-rate and a fast speed.

Key words: Ground control points (GCPs); Dynamic programming; Stereo matching; Vision navigation

1 引 言

立体视觉技术是移动机器人视觉导航领域的关键技术之一^[1,2]. 机器人在未知环境中运动, 必须能准确、快速地辨别障碍物的位置、大小, 找出可行走的道路区域. 这就要求应用于视觉导航系统中的立体视觉匹配算法具有较高的准确性和实时性^[3]. 基于特征的立体匹配只能得到稀疏的视差图^[4,5], 难以保证三维场景信息重构时的准确性和精度; 基于区域的立体匹配得到的视差场虽然足够致密^[6,7], 但视差图中杂点较多, 质量较差. 因此, 这二者都不理想. Ohta 等提出了动态规划立体匹配算法^[8], 在用区域匹配求得初始匹配代价之后, 用动态规划方法进行全局优化, 使视差图的质量得到明显提高, 但视差图中存在条纹瑕疵. Stephen 等引入了控制点修正技术^[9], 进一步改善了视差图质量, 大大减少了条纹瑕疵, 但控制点的求取时间较长, 使匹配算法

整体时间增加, 影响了算法的实时性.

本文以移动机器人的视觉导航技术为背景, 紧紧把握立体匹配算法的准确性和实时性要求, 针对目前算法存在的不足, 对控制点修正的动态规划立体匹配算法进行改进, 提出了一种金字塔双层动态规划立体匹配算法. 此算法在保证匹配精度的前提下, 能大幅度提高实时性.

2 控制点修正的动态规划立体匹配算法

传统的动态规划立体匹配算法分为两个阶段:

1) 计算左右两幅图像的原始匹配代价, 建立视差空间图像;

2) 在视差空间图像上运用动态规划算法进行全局优化, 求取视差.

2.1 视差空间图像的建立

假定立体视觉系统为平行立体视觉系统, 且整个系统已经标定. 采集后的图像已经过外极线校准

收稿日期: 2005-10-11; 修回日期: 2005-12-05.

基金项目: 国家高技术研究发展计划项目(2001AA422250).

作者简介: 赵杰(1968—), 男, 河北卢龙人, 教授, 博士生导师, 博士后, 从事机器人多传感器集成控制技术等研究;

于舒春(1980—), 男, 吉林榆树人, 博士生, 从事机器人立体视觉技术、三维重构技术的研究.

和去躁等预处理.

经外极线校准后,左右两幅图像满足外极线约束条件,即左右两幅图像中对应的匹配点必在同名极线上.假定图像宽度为 m ,最大搜索视差为 d_{\max} .现在取出左右两幅图像上的一组同名极线 x_L 和 x_R ,以右极线 x_R 上一点 $x_{Ri} (0 \leq i \leq m)$ 为参考点,依次与左极线 x_L 上 $n (n = d_{\max})$ 个点 $x_{Lj} (i - j \leq i + d_{\max})$ 进行相似性测度计算,即可得到 n 个初始匹配代价.遍历整条极线,即可得到 nm 个初始匹配代价.将这些代价值存放于一个二维图像中,即可得到二维的视差空间图像,如图1所示.同理,当遍历整个图像时,便可得到三维的视差空间图像.

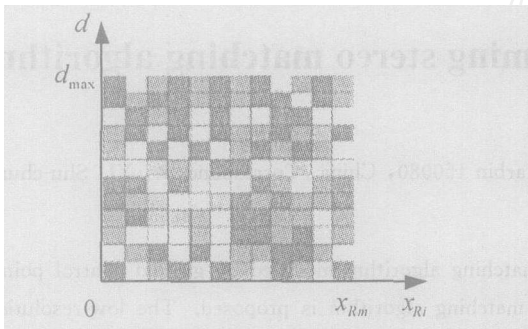


图1 二维视差空间图像

图1中:横坐标 x 表示像素的位置,纵坐标 d 表示视差,图中的子图像块 $I(x, d)$ 表示右图像中像素点 x_{Ri} 与左图像中像素点 $x_{L(i+d)}$ 的匹配代价.颜色越浅表示匹配代价越低;反之,匹配代价越高.

计算视差空间图像中的匹配代价,可用多种相似性测度.下面给出一种常用的相似性测度^[9]:

$$W^L(x, d, w_x, w_y, c_x, c_y) = \frac{w_y - c_y}{w_x - c_x} \left[(I_R(x + t, i + s) - I_L(x + d + t, i + s))^2 + (I_R(x + d, i) - I_L(x + d, i))^2 \right] \quad (1)$$

式中: (c_x, c_y) 为匹配窗口的中心坐标,即参考像素点或待匹配像素点的坐标; $w_x \times w_y$ 为匹配窗口的尺寸; M^L 为

$$M^L(x, y) = \frac{1}{w_y w_x} \sum_{i=-c_y}^{w_y-c_y} \sum_{j=-c_x}^{w_x-c_x} I_L(i, j); \quad (2)$$

M^R 的计算与 M^L 相似.

区域匹配法直接从视差空间图像的每一列中选择匹配代价最小的点,将其纵坐标 d 作为最终的视差.这种方法没有考虑遮挡问题,即由于遮挡的存在,左图像中的某些点在右图像中并没有匹配点,反之亦然.也就是说,将匹配代价最低作为视差选取的唯一约束条件必然会导致很多误匹配.因此,用区域匹配法得到的视差图在遮挡区域往往得不到正确的

视差,从而导致输出的视差图有较多的杂点.

基于动态规划的立体匹配算法,将遮挡代价加入匹配代价中,同时应用唯一性约束和顺序性约束求得在全局范围内都具有最优特性的视差,从而得到高质量的视差图.

2.2 用动态规划算法求取视差

动态规划算法是用于降低全局优化问题复杂度的数学方法,它将整体的优化问题分解为分阶段决策进行.每一个阶段用变量 k 表示,每一个阶段的状态用 x_k 表示.从一个状态转换到下一个状态的方案称为决策,用 u_k 表示.一种决策的好与坏用准则函数 $V_{k,n}$ 衡量.

动态规划算法的一个重要特点是具有无后效性,即当某一阶段的状态一旦确定,此后过程的演变不再受此前各状态和决策的影响.动态规划算法的另一个重要特点是其准则函数具有可分性,即

$$V_{k,n} = \min_{j=1}^n v_j(x_j, u_j) + V_{k-1,n}, \quad (3)$$

式中 $v_k(x_k, u_k)$ 表示准则函数在第 k 阶段的取值.动态规划算法全局优化的实质就是找出全局最优的准则函数 $V_{k,n}$.

将动态规划算法具体到立体匹配问题上, k 表示第 k 个像素点, x_k 表示第 k 个像素点的横坐标, u_k 表示视差, $v_k(x_k, u_k)$ 表示匹配代价函数.因为基于动态规划的立体匹配算法考虑了遮挡问题,所以 $v_k(x_k, u_k)$ 中除包括式(1)计算的初始匹配代价,还应包括遮挡代价.即

$$v_k(x_k, u_k) = v_{\text{data}}(x_k, u_k) + v_{\text{smooth}}(x_k, u_k). \quad (4)$$

式中: $v_{\text{data}}(x_k, u_k)$ 表示由相似性测度计算的初始匹配代价; $v_{\text{smooth}}(x_k, u_k)$ 表示遮挡代价函数,在遮挡区域,其值将明显增大.可见,式(4)在遮挡区域的匹配总代价明显增大.因此,基于动态规划的立体匹配算法能较好地保持视差的平滑性,剔除遮挡区域的误匹配点.

结合式(3)和(4)可以看出,基于动态规划的立体匹配算法的实质就是要在视差空间图像上求出一条最优路径.

为进一步限制路径搜索的范围,降低算法的复杂度,文献[10]引入了唯一性约束和顺序性约束,从而限定了从前一阶段到当前阶段的转换只有7种可能的形式.分别用 m, l, r 表示前一阶段的匹配点、左图可见点、右图可见点,用 M, L, R 表示当前阶段的匹配点、左图可见点、右图可见点,则这7种可能形式如图2所示.这样便可以通过动态规划算法进

行视差求取了。

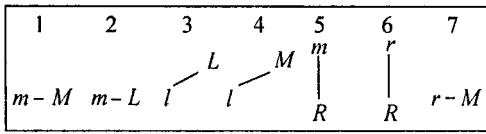


图 2 动态规划立体匹配算法中的 7 种状态转换

由传统的基于动态规划的立体匹配算法得出的视差图有明显的条纹状瑕疵,这是因为一旦某个像素点匹配出现错误,会波及同一扫描线上后续像素的匹配.这种问题的出现正好验证了动态规划算法的特点.因为前一阶段的状态决定当前阶段的状态,而根据无后效性,当前阶段状态确定后,就不再考虑此前的所有阶段的状态,所以当前一阶段的匹配出现错误时,会直接影响当前阶段的匹配,而这种错误在后续匹配中再也没有机会修正.引入控制点修正技术可以较好地解决这个问题^[9].

2.3 控制点修正的动态规划立体匹配算法

动态规划算法在视差空间图像上进行视差求取时,规划出的路径总是趋向于寻找那些具有最小匹配代价的点.如果事先知道某些点为正确的匹配点,并将这些点的总体匹配代价设定为一个较小值,就可以迫使最优路径必须经过这些正确的匹配点,从而减少条纹瑕疵.本文将这些点称为控制点.这种人为计算并设定控制点的动态规划立体匹配算法就是控制点修正的动态规划立体匹配算法.

此种方法的关键是如何正确地求出控制点.控制点的求取应符合以下 3 条基本规则^[9]:

- 1) 控制点必须是能够通过左右一致性校验的视差点;
- 2) 控制点的匹配代价必须小于动态规划时的遮挡代价;
- 3) 控制点不是孤立的.

根据这 3 条规则,为每条扫描线求取 12 个控制点,就能大大改善视差图^[9],较好地去除条纹瑕疵.但是,视差图质量得到改善的同时,出现了新的问题:动态规划算法是众多全局优化算法中较为快速的,在耗时较多的立体匹配中引入此种方法的目的就是为了尽量减少算法的总用时,然而控制点的求取使整个算法的耗时增加了 1 倍以上,极大地影响了整个算法的实时性.因此,如何减少控制点求取的时间便成为提高整个算法实时性的关键.

3 金字塔双层动态规划立体匹配算法

现在分析控制点的求取为什么耗时较长.这是由控制点求取的第 1 条基本规则决定的,即控制点必须是能够通过左右一致性校验的视差点.左右一

致性校验的实质是:先以左图像为参考图像、右图像为待匹配图像进行立体匹配,求得各像素点视差;再以右图像为参考图像、左图像为待匹配图像进行立体匹配,求得各像素点视差,最后比较同一像素点两次求得的视差值,如相等,则通过左右一致性校验.因此,控制点的求取过程相当于进行了两次立体匹配.由于对每条扫描线只求取 12 个控制点就大大减少了条纹瑕疵,说明少量的控制点便可以大大改善视差图的质量.因此,如果在低分辨率图像上求取控制点,其数目足以改善视差图质量,而低分辨率的图像像素数目则大为减少,从而可以大大缩短控制点的求取时间.

3.1 用金字塔算法求解多分辨率图像

金字塔算法属于小波理论中多分辨率分析范畴,在图像处理中的应用非常广泛.

构造图像金字塔,并根据需要确定金字塔级数 k 和逐级缩放比例 r .高层图像(低分辨率图像)的像素值可直接用低层图像(高分辨率图像)的对应点像素值,也可通过对低层图像的 $r \times r$ 区域的像素求平均获得.通常 r 可取 2,则由下到上每一级金字塔图像的像素数目减少到原来的 $1/4$.

用金字塔算法求得左右两幅图像低分辨率图像后,在低分辨率图像上求取候选控制点集;然后在高分辨率图像上求得最终控制点集,并用此控制点集修正高分率图像上的动态规划立体匹配算法.

3.2 求取控制点集

控制点求取的准确与否,直接决定着后续动态规划的正确与否,其重要性显而易见.故此,文献[9]要求对控制点的求取必须满足 3 条基本规则.本文将控制点的求取过程分为两个阶段,以确保其准确性.

第 1 阶段是在低分辨率图像上求取候选控制点集 GCP_{Sfirst} .该阶段分为以下两个步骤:

- 1) 在低分辨率图像上采用传统的动态规划立体匹配算法进行立体匹配,得出每一点的视差值,记录其匹配代价;
- 2) 将各点的匹配代价与预先设定的阈值(一般选取高分辨图像上进行动态规划立体匹配时的遮挡代价作为此阈值)进行比较,小于 的全部纳入候选控制点集 GCP_{Sfirst} ,否则淘汰之.

第 2 阶段是在高分辨率图像上求取最终控制点集 GCP_{Slast} .该阶段也分为两个步骤:

- 1) 先对左右两图图像的高分辨率图像进行初始匹配代价计算,建立视差空间图像;
- 2) 将低分辨率图像上记录的候选控制点的匹配代价与此点在高分辨率视差空间图像上对应列的

所有匹配代价进行逐一比较,如果其值仍为最小,则纳入最终控制点集 GCP_{Slast} , 否则淘汰之。

本文用最终控制点修正高分辨率图像上已经求出的视差空间图像;然后用动态规划法进行立体匹配。由于此方法对控制点求取的主要耗时工作都在低分辨率图像上进行,算法用时大为减少。

该方法先后在低、高分辨率图像上进行两次动态规划算法,故称为金字塔双层动态规划立体匹配算法。

4 实验与实验结果分析

为了验证本文算法的正确性,用 Tsukuba 人工合成立体图对进行实验。Tsukuba 立体图对满足平行立体视觉测距原理和外极线约束条件,且无噪声,故可直接对其施加各种立体匹配算法。Tsukuba 立体图如图 3(a) 所示,宽为 384,高为 288,其真实视差如图 3(b) 所示。

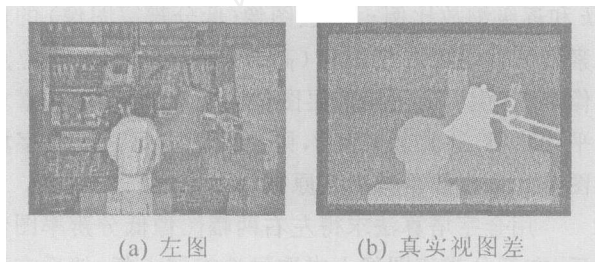


图 3 Tsukuba 人工合成图

本文先后对 Tsukuba 立体图对施加 4 种不同的立体匹配算法,具体实验如下:

1) 基于区域的立体匹配算法实验。实验中采用 SAD 相似性测度算子和 7×7 匹配窗口求取初始匹配代价,视差搜索范围为 16,如图 4(a) 所示。

2) 基于动态规划的立体匹配算法实验。实验中采用 SAD 相似性测度算子和 7×7 匹配窗口求取初始匹配代价,视差搜索范围为 16,动态规划时所用的遮挡代价为 20,如图 4(b) 所示。

3) 基于控制点修正的动态规划立体匹配算法实验。实验中采用 SAD 相似性测度算子和 7×7 匹配窗口求取初始匹配代价,视差搜索范围为 16,动态规划时所用的遮挡代价为 20,每条扫描线的控制点为 12 个,如图 4(c) 所示。

4) 金字塔双层动态规划立体匹配算法实验。在用金字塔算法求解低分辨率图像时,设定金字塔级数为 2,缩放比例为 2。实验中在低分辨率图像上采用 SAD 相似性测度算子和 3×3 匹配窗口求取初始匹配代价,视差搜索范围为 8,动态规划的遮挡代价为 20,候选控制点选取阈值为 20;在高分辨率图像上采用 SAD 相似性测度算子和 7×7 匹配窗口求

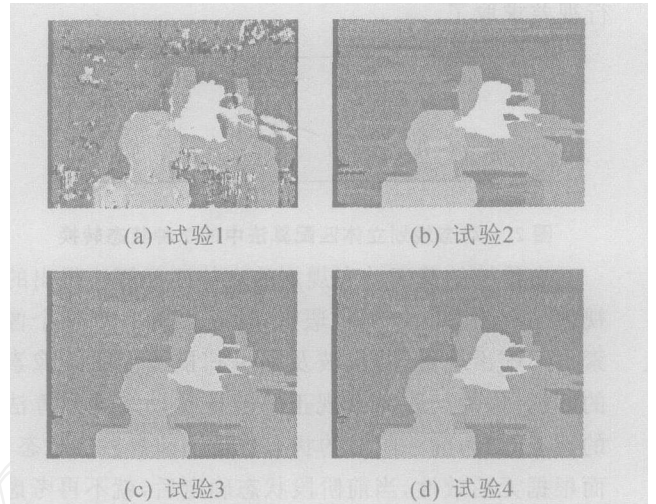


图 4 4 组实验得到的视差图

取初始匹配代价,视差搜索范围为 16,动态规划遮挡代价为 20。实验结果如图 4(d) 所示。

由图 4 可以看出,实验 1 输出的视差图在遮挡区域存在较多的误匹配,视差图中杂点较多;实验 2 输出的视差图质量要好得多,但存在条纹瑕疵;实验 3 输出的视差图则显然消除了条纹瑕疵;实验 4 与实验 3 的效果相差无几。

除定性比较外,本文还对这 4 种算法进行了定量比较,如表 1 所示。

表 1 4 组实验的实验性能比较

	实验 1	实验 2	实验 3	实验 4
匹配率/ %	68.5	80.4	86.1	86.0
时间/ ms	1 290	1 673	2 935	1 804

由表 1 可以看出,实验 3 和实验 4 的匹配率远远高于实验 1 和实验 2,但实验 3 的用时过长,而实验 4 则匹配率高、耗时低,最为理想,从而验证了本文提出的金字塔双层动态规划立体匹配算法的有效性。

5 结 论

本文针对基于控制点的动态规划立体匹配算法耗时过大的问题展开研究,在精确分析控制点求取用时过多的原因之后,提出了一种金字塔双层动态规划立体匹配算法。此算法首先通过金字塔算法求得高、低分辨率图像,然后在低分辨率图像上求取候选控制点集;之后在高分辨率图像上求取最终控制点集,并用此控制点集修正高分辨率图像上的动态规划立体匹配,得到了质量较好的视差图。而候选控制点集的求取在低分辨率图像上进行,大大缩短了算法用时,从而在保证匹配精度的前提下,大大提高了匹配算法的实时性,为完成后续的移动机器人视觉导航任务奠定了基础。

(下转第 77 页)

5 结 论

本文结合减法聚类和标准 FCM 对聚类算法进行改进,改进后的算法收敛速度快且更加稳定,减少了由用户确定聚类中心个数的盲目性.将改进后的算法用于模糊系统的输入、输出空间划分和隶属度函数类型及参数的确定,适用范围较广,且便于实现,提高了模糊系统的应用效率.

参考文献(References)

- [1] 王士同. 模糊系统、模糊神经网络及应用程序设计[M]. 上海: 上海科技文献出版社, 1998: 25-30.
(Wang S T. Fuzzy systems, fuzzy neural network and application programming [M]. Shanghai: Shanghai Scientific and Technical Document Publishing House, 1998: 25-30.)
- [2] Pedrycz A, Reformat M. Hierarchical FCM in a stepwise discovery of structure in data [J]. Soft Computing, 2005, 10(3): 244-256.
- [3] Bandyopadhyay. Satellite image classification using genetically guided fuzzy clustering with spatial information[J]. Int J of Remote Sensing, 2005, 26(3): 579-593.
- [4] 冯衍秋, 陈武凡, 梁斌, 等. 基于 Gibbs 随机场与模糊 C 均值聚类的图像分割新算法[J]. 电子学报, 2004, 32(4): 645-647
(Feng Y Q, Chen W F, Liang B, et al. An new

algorithm for image segmentation based on gibbs random field and fuzzy C-means clustering[J]. Acta Electronica Sinica, 2004, 32(4): 645-647.)

- [5] Tran D, Wagner M. Generalized fuzzy hidden Markov models for speech recognition[C]. 2002 AFSS Int Conf on Fuzzy Systems. Calcutta: Springer-Verlag GmbH, 2002: 345.
- [6] Lotfi A Zadeh, Berkeley C A. Fuzzy logic toolbox for use with MATLAB version 2. 2. 1 [M]. Natick: The MathWorks, Inc, 2005.
- [7] 侯彩虹, 崔运花, 余润仙, 等. 基于模糊聚类分析的织物质量分级方法[J]. 东华大学学报, 2005, 31(1): 54-58.
(Hou C H, Cui Y H, Yu R X, et al. Fuzzy clustering analysis based method for fabric sample grading[J]. J of Donghua Univeristy, 2005, 31(1): 54-58.)
- [8] Shitong W, Chung K F, Hongbin S, et al. Note on the relationship between probabilistic and fuzzy clustering [J]. Soft Computing, 2004, 8(5): 366-369.
- [9] Carl de Boor. Curve fitting toolbox for use with MATLAB [M]. Natick: The MathWorks, Inc, 2005.
- [10] Thomas F Coleman, Jianguo Liu, Wei Yuan. A new trust-region algorithm for equality constrained optimization [J]. Computational Optimization and Applications, 2002, 21(2): 177-199.

(上接第 72 页)

参考文献(References)

- [1] 李俊, 孙德敏. 月球车的视觉和自主导航系统[J]. 航天控制, 2000, 2(8): 46-51.
(Li J, Sun D M. The vision system and autonomous navigation system for the lunar rover [J]. Aerospace Control, 2000, 2(8): 46-51.)
- [2] Crane C D. Development of an integrated sensor system for obstacle detection and terrain evaluation for application to unmanned ground vehicles[J]. Unmanned Ground Vehicle Technology VII, 2005, 5084(10): 156-165.
- [3] Maarten Vergauwen, Marc Pollefeys, Luc Van Gool. A stereo-vision system for support of planetary surface exploration [J]. Machine Vision and Applications, 2003, 14(1): 5-14.
- [4] 徐奕, 周军, 周源华. 立体视觉匹配技术[J]. 计算机工程与应用, 2003, 15(1): 1-5.
(Xu Y, Zhou J, Zhou Y H. Stereo matching technology [J]. Computer Engineering and Applications, 2003, 15(1): 1-5.)
- [5] Hsieh Y C. Performance evaluation of scene registration and stereo matching for cartographic feature extraction

[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1992, 14(2): 214-238.

- [6] 潘华, 郭戈. 立体视觉研究进展[J]. 计算机测量与控制, 2004, 12(12): 1121-1124.
(Pan H, Guo G. Review of stereo vision[J]. Computer Automated Measurement and Control, 2004, 12(12): 1121-1124.)
- [7] Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence[C]. 3rd European Conf on Computer Vision (ECCV '94). Stockholm, 1994: 151-158.
- [8] Ohta Y, Kanade T. Stereo by intra and inter-scanline search using dynamic programming[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1985, 7(2): 139-154.
- [9] Stephen S Intille, Aaron F Bobick. Disparity-space images and large occlusion stereo [C]. 3rd European Conf on Computer Vision (ECCV '94). Stockholm, 1994: 179-186.
- [10] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. Int J of Computer Vision, 2002, 47(1): 7-42.