

doi: 10.3969/j.issn.1674-8425(z).2015.02.014

双目立体视觉匹配技术综述

曹之乐, 严中红, 王 洪

(重庆理工大学 药学与生物工程学院, 重庆 400054)

摘 要: 立体视觉匹配技术是立体视觉研究的核心问题。立体匹配方法在不同的应用领域有着不同的应对策略。介绍了双目视觉立体匹配技术的原理和一些研究关键点。通过总结现有各种算法的优缺点, 对基于区域的立体匹配算法和基于全局的立体匹配算法进行了深入论述和研究, 最后对立体匹配算法的未来发展做了展望。对研究者选择立体匹配算法有一定参考意义。

关 键 词: 双目立体视觉; 立体匹配; 动态规划; 图像分割; 置信传播

中图分类号: TP39

文献标识码: A

文章编号: 1674-8425(2015)02-0070-06

Summary of Binocular Stereo Vision Matching Technology

CAO Zhi-Le, YAN Zhong-Hong, WANG Hong

(College of Pharmacy and Biomedical Engineering,
Chongqing University of Technology, Chongqing 400054, China)

Abstract: The core part of the stereo vision is stereo matching. For different application fields, there are different stereo matching methods to deal with. First, the paper briefly introduced the principle of binocular vision stereo matching technology and some key points. Secondly, we summed up the advantages and disadvantages of various algorithms and had deeply discussed and studied the stereo matching algorithm based on district and stereo matching algorithm based on overall situation. Finally, the stereo matching algorithm of the future was prospected, which provides some valuable ideas for those who study in this field.

Key words: binocular stereo vision; stereo matching; dynamic programming; image segmentation; belief propagation

立体视觉技术广泛应用于探测车导航、机器人导航、3D 场景图绘制、3D 场景重建、汽车预警系统、大型机械位姿感知等领域。自 20 世纪 70

年代 Marr 首次提出完整的机器视觉系统计算理论框架^[1]以来, 三维视觉信号备受人们的关注。早期的立体匹配是基于小窗口的区域匹配, 提取的

收稿日期: 2014-10-10

基金项目: 重庆市教委科学技术研究项目(KJ120807)

作者简介: 曹之乐(1989—), 男, 河南濮阳人, 硕士研究生, 主要从事生物医学信号及图像处理方面的研究; 通讯作者 严中红(1964—), 男, 重庆垫江人, 博士, 教授, 主要从事生物医学信号及图像处理方面的研究。

引用格式: 曹之乐, 严中红, 王洪. 双目立体视觉匹配技术综述[J]. 重庆理工大学学报: 自然科学版, 2015(2): 70-75.

Citation format: CAO Zhi-Le, YAN Zhong-Hong, WANG Hong. Summary of Binocular Stereo Vision Matching Technology [J]. Journal of Chongqing University of Technology: Natural Science, 2015(2): 70-75.

特征值大多为灰度、颜色等信息。后来开发出自适应窗口技术,针对特征值也开发出对光照鲁棒性的变换值,如 Census 变换等。近年来,基于全局的立体匹配技术成为人们研究的热点,其大致思路为:提取特征值光流场特征、边缘特征、Harris 角点特征、SIFT 特征向量、SURF 特征等可靠特征,计算出初步视差图,再由全局优化函数(动态规划、图像分割、置信传播等)对整幅图像的视差值进行迭代优化分配计算。目前,主流的立体视觉匹配技术分为 2 种:基于区域的立体匹配和基于全局的立体匹配。本文主要就以上 2 种技术的一些经典立体匹配方法和近年来发展起来的立体匹配新方法进行较深入的讨论和评述。

1 双目立体视觉系统

最常见的立体视觉系统是双目立体视觉系统,也可以是多目视觉系统,但双目视觉系统是其基础。因此,本文首先简要阐述双目视觉系统,其几何原理如图 1 所示。

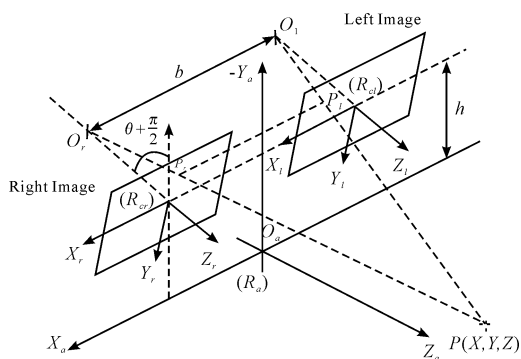


图 1 双目视觉系统几何原理

在三维空间坐标系 $O_a X_a Y_a Z_a$ 中,左摄像头 O_l 和右摄像头 O_r 处于理想位置关系。其中,左摄像头光轴 Z_l 和右摄像头光轴 Z_r 平行,且直线 $Z_l Z_r$ 与三维空间坐标系中的 X_a 轴垂直。2 个成像中心 O_l, O_r 的连线为基线 b 。三维空间中的一点 P 与两成像中心点 O_l, O_r 以及映射点 p_l 和 p_r 处于同一平面,平面与两成像平面的交线被称为极线。根据立体视觉的几何成像原理,双目立体视觉系统对第三维(即深度)的测量由式(1)得出。其

中: f 为焦距, d 为视差,即点 p_l 和 p_r 两图像在 X 轴方向上的差值,深度为 Z 。基线 b 和焦距 f 可在双目标定时计算得出,因此双目立体视觉定位的核心内容是对视差 d 的计算。视差 d 由立体匹配得出,故对立体匹配的研究和讨论成为双目视觉系统研究的核心内容。

$$Z = \frac{b \cdot f}{d} \quad (1)$$

1.1 立体匹配约束

1) 唯一性约束:三维空间中的点映射到左右摄像头时,只会在图像上映射出唯一的对应点。在匹配时左右图像只有唯一一点相匹配。

2) 连续性约束:三维空间中的物体一般是连续光滑的,在映射到左右摄像头时这种特性也被保留下来。在连续的图像上,其视差也一般是连续的。

3) 极线约束:如图 1 所示,对于一个图像上的映射点来说,其匹配点必定落在另一图像的极线上。而对双目视觉系统来说,通常将其校正为上文原理中的假设情况,在这种情况下匹配点对处于同一水平线,即点在图像中的坐标 y 值相等。

4) 顺序一致性约束:三维空间中点的位置关系会在映射到成像平面时保留下来,即原物体的位置顺序在两幅图像中不会改变。

1.2 立体匹配的匹配单元

匹配单元是立体匹配中的单位匹配对象,也就是立体匹配中的匹配特征对象。一般有以下几种:① 点:在双目立体匹配中最常用的特征就是点特征,例如像素点的灰度、颜色值、Harris 角点、SIFT 特征点、SURF 特征点等;② 块:图像块也可以认为是图像区域,块的特征提取可以是直观的块内所有像素点的灰度或颜色值,也可以是对块内像素点的统计特征,还可以是对块内像素点的变换特征;③ 线:线一般只是边缘线,是最能体现图像的纹理区域特征。在有些方法中将角点或特征点之间的连线作为特征线来提取。

2 立体匹配算法

在立体匹配视觉系统的原理基础上,根据匹

配的约束条件就可以对选定的匹配基元进行匹配。匹配算法是立体匹配技术的核心。

2.1 基于区域的立体匹配算法

基于区域的立体匹配通常是在一个设定的窗口基础上来完成,以窗口中提取的特征向量为基础进行匹配。这种匹配可以是逐像素的匹配,也可以是成块的匹配,更进一步的,有根据图像自适应调节窗口大小的区域匹配算法。经典的有 Lazaros Nalpantidi 等^[2]以灰度、颜色为特征,通过对窗口内的点的空间距离加权、颜色加权以及边缘加权来提高匹配精度。对于窗口来说,窗口越大信息量越多,匹配精度越高,但会带来大量的复杂计算,影响整体算法的运行速度^[3]。

在计算出的视差层面上通常采用赢者通吃 (winner-takes-all, WTA) 的方法来聚合匹配像素或块。以亮度特征为例,在立体匹配过程中可以选取绝对差值和 SAD、差值平方和 SSD、归一化相关系数 NCC、零均值绝对差值和 ZSAD 等匹配函数,详见公式(2)。

$$\begin{aligned} SAD &= \sum_{i=n} \sum_{j=m} |I_1(u+i, v+j) - I_2(u+d+i, v+j)| \\ SSD &= \sum_{i=n} \sum_{j=m} (I_1(u+i, v+j) - I_2(u+d+i, v+j))^2 \\ NCC &= \frac{\sum_{i=n} \sum_{j=m} I_1(u+i, v+j) I_2(u+d+i, v+j)}{\sqrt{\sum_{i=n} \sum_{j=m} I_1(u+i, v+j)^2 \sum_{i=n} \sum_{j=m} I_2(u+d+i, v+j)^2}} \\ ZSAD &= \sum_{i=n} \sum_{j=m} |I_1(u+i, v+j) - \bar{I}_1| - |I_2(u+d+i, v+j) - \bar{I}_2| \end{aligned} \quad (2)$$

上述函数中: I_1, I_2 为左右图像对; u, v 为像素的坐标点; n, m 为窗口的长和宽; \bar{I}_1, \bar{I}_2 为窗口中像素亮度的平均值。这些匹配函数可以对窗口中的亮度等特征进行匹配。

区域匹配由于匹配的特征为独立的像素或像素块,因此在噪声和误匹配的影响下,匹配出的图像很不平滑,更重要的是图像中的无纹理区域因为像素的亮度相似会产生大量的误匹配。然而区域匹配因其计算复杂度低,运算速度快,故在大量的实时系统中获得应用。为克服区域匹配算法的缺点,通常采用的特征也由简单的亮度转变为更

为丰富的特征,如多尺度对频段上的特征^[4-5],基于梯度的加权特征^[3]等。

另外,对于受光照变化影响严重的图像,一般采用变换的方法进行处理。例如,Census 变换可以将窗口中的亮度值变换为一个序列,匹配时只需计算两幅图像中窗口特征序列的汉明距离 (Hamming distance);Rank 变换则是统计 0 或 1 的个数作为特征值,因而可以使用上文中的匹配函数。

2.2 基于全局的立体匹配算法

基于全局的立体匹配通常是先找出图像中显著的特征点、特征线或者扫描线,初步匹配得出一个初始的视差值,然后使用全局能量函数进行约束,通过匹配算法不断地对分配的视差值迭代优化,直到最优。此迭代优化的过程以全局能量函数最小化为目的。

全局能量函数一般包括 2 项,数据项和平滑项,见式(3)。数据项用来表示该匹配基元的视差值,平滑项用来约束邻域基元之间的差值^[6]。近年来,硬件技术的飞速发展使得这种复杂匹配方法的实时性成为可能,未来的立体匹配方法也必将以基于全局的匹配方法为主流。

$$E(f) = E_{\text{smooth}}(f) + E_{\text{data}}(f) \quad (3)$$

2.2.1 动态规划全局优化算法

动态规划是一种常用的算法,是解决问题的一种思路。其基本思路是将给出的问题分为一个个的小问题,通过解决这些小问题从而得出问题的解。一般这些子问题很相似,为了减小计算量,动态规划试图对每个子问题只计算一次。当得到一个子问题的解时,可将该结果直接保存,当再次需要相同子问题解时可直接查找此结果。

在动态规划的使用方面,2010 年,J. Cal 在文献^[7]中结合光流法和动态规划,以光流法计算视差图,再基于此视差图使用动态规划进行优化。2011 年,Mohamed El Ansari 等^[8]依据提取出的边缘线上的点做动态规划,以边缘点的梯度幅值和方向作为引导动态规划的全局能量函数。2013 年,Minh Nguyen, Yuk Hin Chan 在文献^[9]中使用图像块进行动态规划匹配,并进行双向匹配以降

低误差。

应用立体匹配时,一般先匹配出一个初始的视差图,然后使用动态规划算法,在一条扫描线上(通常为极线)寻找使立体匹配能量函数值最小的路径(视差值)。扫描线上的每一点的视差为得到的最小路径值(视差值)^[7]。

经典的动态规划算法步骤如下:

1) 初始化。初始化是对图像的视差值进行初始的设定。这个初始设定可以全部为0,也可以预先简单匹配出来的视差值。后者在动态规划优化运算时迭代速度更快。

2) 循环求解。在一条扫描线上逐点计算,在视差范围内循环取值,得出前一点的全局能量函数值和前一点的路径参数值。循环在扫描线的末尾前一点结束。

3) 返回视差值。每一点的视差值都为其路径参数值。具体步骤详见文献[7]。

双目视觉立体匹配中,在上文假设前提下,极线处于两张图像的同一水平线上,即上图中左图极线和右图极线的 y 轴坐标值相等。因此,经常使用的扫描线就是极线,可以进行单向匹配,即由一图作为基准去和另一张图匹配;也可以进行双向匹配,即从左向右匹配,而后再从右向左匹配,取匹配较高的视差值作为最终视差值,以提高匹配精度。但是用极线作扫描线时,动态规划仅是对单独的横向极线匹配,忽略了平行极线之间的关系,使得图像的匹配结果呈现水平的条纹状,因此动态规划法在使用中常以消除此条纹作为提高视差图质量的核心。同时,由于动态规划仅是一种优化算法,直接使用其进行匹配会耗费大量的运行时间,其迭代要达到一定次数才会有较好的效果,因此这种算法一般和其他匹配算法联合使用^[10]。

2.2.2 图像分割全局优化算法

图像分割是根据图像的自身结构特性,将图像分为不同的区域,而后再将这些区域的立体匹配问题转化为能量函数最小化问题^[11]。这种做法既可以达到全局算法的精确匹配,又可以保持原图像中的结构信息。基于全局的立体匹配算法一

般都会根据算法及约束的原理构造一个全局能量函数。同样的,图像分割法也有这样的全局能量函数来实现迭代优化。

2006年,Andreas Klaus等^[12]基于颜色特征进行图像分割,对划分出来的图像区域使用自适应相似性度量函数计算最优视差。2007年,张娟^[13]对图像的特征空间做均值漂移计算,使用分水岭分割模型得出分割的图像块。2010年,Bleyer等^[10]提出了一种图像分割的基本算法框架,首先将图像分割为不同的区域,根据不同区域的最小匹配函数计算出分割区域的视差,然后拼合视差相同的区域,最终获取区域视差的最优化匹配。2011年,颜轲等^[11]对基于MRF(Markov随机场)的图像分割算法进行了改进,先对图像进行分割,在此基础上建立立体匹配的MRF模型,以充分保留图像的构架信息,进而得到平滑的视差图像。

图像分割法的目的是根据图像特征划分得到一个分割的块集合,而后对这些块分别进行优化计算视差,最大限度地保留图像的结构信息^[12-14]。为实现这一目标,目前的算法大致分为以下几类:①以像素为基准的分割法,不考虑图像的空间构架,只以像素点的颜色特征等来划分图像;②以区域为基准的分割法,以图像中区域块的连通特性为约束,使用区域生长、分裂和聚合技术实现;③以边界为基准,依靠测定出的边缘划分图像;④以模型为基准,大多为马尔科夫(Markov)随机场或吉布斯(Gibbs)随机场等空间交互模型来对图像建模^[15]。近年来多采用图像分割算法方法。

图像分割在大部分领域中都有很好的应用,其原理是基于颜色、边缘等信息进行分割,而当这些信息与深度信息不能统一时(即深度信息与颜色、边缘等无关时),匹配就会产生许多误差。

2.2.3 置信传播

置信传播算法建立在马尔科夫随机场模型的基础上,更复杂的模型可以是一个或者几个马尔科夫随机场构成的马尔科夫网络^[16]。像素点与马尔科夫随机场模型中的节点一一对应。置信传播算法采用消息传输机制实现节点之间置信度的

最大化,同时实现马尔科夫随机场最大概率分布,等同于实现全局能量函数最小化^[17]。

2003 年,Jian Sun 等^[16]使用 3 个马尔科夫随机场构建出一个马尔科夫随机网络,3 个场中平滑场保持视差的连续性,线性场保留视差的断层,二元场保持视差层的闭塞。2013 年,Armin Ahmadzadeh 等^[18]在多个平台(包括 FPGA,GPU,多核 GPU)上实验置信传播算法,并选定多核 GPU 得出一个性能和效率俱佳的置信传播立体匹配算法。

置信传播算法的思路如下:① 构建全局能量函数,以此约束节点的更新,式子构建如上文所说包括数据项和平滑项;② 构建图像对的马尔科夫随机场模型,并确定节点之间的消息传播方式和传播信息;③ 循环迭代直至节点上的信息最优化,根据节点信息算出视差值。具体算法步骤详见文献^[19]。

马尔科夫随机场模型是将图像对的像素点看作节点,而每个节点包含两方面信息:数据信息和消息信息。数据信息是每一个像素点的视差值,消息信息则是每个节点的信息值(初始为零)。

置信算法在每一次迭代过程中都会对节点的信息进行更新,节点间的消息传播一般为 BP-M 和 BP-S。在 BP-M 中,每一次的传播都在邻域 4 个方向(上下左右)点上进行。而在 BP-S 中的传播分为 2 种:一种向前,即消息向下和向右传播直至达到最后一个节;另一种是向后,即消息向上和向左传播直至达到第一个节点^[18]。与动态规划类似,该算法也是一种强大的优化算法,在应用中通常与其他匹配算法联合使用。本算法的复杂性在于模型的建立、全局能量函数的构建,以及信息传递的设定。

3 总结

总体来说,计算视差的过程是一个提取特征并进行特征匹配的过程。基于局部的立体匹配算法在匹配过程中只是简单计算左右图像的特征差异最小化,或者是计算左右图像的特征相似性。

基于全局的立体匹配算法在匹配过程中使用优化算法,同时考虑平滑性和精确性,使全局能量函数最小化。上文中的动态规划和置信传播都是优化视差的算法,图像分割则是重要的保留图像结构的算法,因此对于高精度的视差计算,可以结合两者一起使用。本文算法技术总结见表 1。

表 1 算法技术总结

| 方法 | 核心思想 | 优点 | 缺点 |
|----------------|-------------------------------------|----------------------------------|-------------------|
| SSD, SAD, ZSAD | 窗口内信息求和进行匹配。 | 算法简单,速度快 | 精度低,误匹配等高 |
| NCC | 窗口内信息求相似性进行匹配。 | 算法较为简单,速度快 | 精度低,对无纹理区域误匹配等高 |
| DP, BP | 对视差值迭代优化,以求得最优匹配。一般与其他算法联合使用。 | 精度高,对无纹理区域效果好 | 算法复杂,速度慢,单一使用效果不佳 |
| Graph cut | 依据图像的结构信息分割图像,分别处理后再融合。一般与其他算法联合使用。 | 保留原图像结构信息,剪度高,对无纹理区域效果好,对边界区域剪好。 | 算法复杂,剪度慢,单一使用效果不佳 |

注:SAD,绝对差值和;SSD,差值平方和;NCC,归一化相关系数;ZSAD,零均值绝对差值和;DP,动态规划;BP,置信传播;Graph cut,图像分割;

4 结束语

近年来,新的方法不断被引入立体匹配领域,立体视觉技术也不断应用于新领域。Kajal Sharma 等在 2012 年提出用 SIFT 特征作为输入,使用自组织特征映射模型这种人工神经网络进行立体匹配。Li-Heng Lin 等在 2013 年提出一种使用 harrs 角点群进行立体匹配以获取大型器械的位姿算法。

相比运算速度,近年来更重视对精度和匹配度的要求。在硬件技术迅速发展的现状下,之前许多不能保证实时性的算法都可以很好地实现。

未来立体匹配的研究重点将集中于如何提高精度,对于算法的应用也应是多种算法联合的多层面多尺度上的匹配。

参考文献:

- [1] Marr D C. A Computational Investigation into the Human Representation and Processing of Visual Information [M]. San Francisco: W. H. Freeman and company, 1982.
- [2] Nalpantidis L, Gasteratos A. Stereo vision for robotic applications in the presence of non-ideal lighting conditions [J]. Image and Vision Computing, 2010, 28 (6): 940-951.
- [3] Ambrosch K, Kubinger W. Accurate hardware-based stereo vision [J]. Computer Vision and Image Understanding, 2010, 114: 1303-1316.
- [4] Nuan Shao, Hui Guang Li, Le Liu, et al. Stereo vision robot obstacle detection based on the SIFT [R]. 2010 Second WRI Global Congress on Intelligent Systems, 2010.
- [5] Sizintsev M, Wildes R P. Coarse-to-fine stereo vision with accurate 3D boundaries [J]. Image and Vision Computing, 2010, 28: 352-366.
- [6] Boykov Y, Veksler O, Zabih R. Fast Approximate energy minimization via graph cut [J]. Pattern Analysis and Machine Intelligence, 2002, 23(11): 1222-1239.
- [7] Cai J. Integration of optical flow and dynamic programming for stereo matching [J]. IET Image Processing, 2012, 6(3): 205-212.
- [8] Mohamed El Ansari, Mazoul A, Benshrair A, et al. A Real-time Spatio-Temporal Stereo Matching for Road Applications [C]//14th International IEEE Conference on Intelligent Transportation Systems. 2011.
- [9] Nguyen M, Yuk Hin Chan, Delmas P, et al. Symmetric Dynamic Programming Stereo Using Block Matching Guidance [C]//28th International Conference on Image and Vision Computing New Zealand. 2013.
- [10] Gang Yao, Yong Liu, Bangjun Lei, et al. A Rapid Stereo Matching Algorithm Based on Disparity Interpolation [C]//Proceedings of 2010 Conference on Dependable Computing. 2010.
- [11] 颜轲, 万国伟, 李思昆. 基于图像分割的立体匹配算法 [J]. 计算机应用, 2011, 31(1): 175-178.
- [12] 赵杰, 祁永梅, 潘正勇. 结合边界和区域的水平集超声图像分割算法 [J]. 激光杂志, 2013(6): 46-48.
- [13] 丛超. 医用心脏图像分割算法的量化评估框架 [J]. 重庆理工大学学报: 自然科学版, 2013(7): 71-75, 112.
- [14] 易欣, 贾振红, 覃锡忠, 等. 一种基于多相位水平集的图像分割 [J]. 激光杂志, 2013(6): 37-39.
- [15] 张娟. 基于彩色图像分割的立体匹配算法研究 [D]. 济南: 山东大学, 2007.
- [16] Jian Sun, Nan Ning Zheng. Senior Member. Stereo matching using belief propagation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7): 787-800.
- [17] 李鹤喜, 孙玲云. 基于彩色图像分割的置信传播快速立体匹配算法 [J]. 数字技术与应用, 2013(1): 115-117.
- [18] Ahmadzadeh A, Madani H, Jafari K, et al. Fast and Adaptive BP-based Multi-core Implementation for Stereo Matching [R]. Formal Methods and Models for Code-sign, 2013.
- [19] Felzenszwalb P F, Huttenlocher D P. Efficient Belief Propagation for Early Vision [J]. International Journal of Computer Vision, 2006, 70(1): 41-54.

(责任编辑 杨黎丽)

(上接第 69 页)

- [11] 侯立鹏. Google map 的数据管理和表现机制的讨论研究 [J]. 计算机光盘软件与应用, 2011, 16: 94-94.
- [12] 侯俊杰. 深入浅出 MFC [B]. 第 2 版. 武汉: 华中科技大学出版社, 2001.
- [13] 陶陶, 尹轶华, 陶以欣. GIS 软件共享 CAD 图块的关键技术 [J]. 计算机辅助设计与图形学学报, 2006, 18(9): 1460-1464.
- [14] 赵玲, 谢树春, 汤井田. 动态分段思想在 MapInfo 中的实现 [J]. 测绘学报, 2005, 34(2): 175-178.
- [15] 陈思璘, 魏萍. 如何用 MapInfo 实现油田开发地理查询 [J]. 石油工业计算机应用, 1998(3): 33-36.

(责任编辑 杨黎丽)