

基于互信息的亚像素级立体视觉点匹配方法研究

孙少杰, 杨晓东, 任继昌
(海军潜艇学院, 山东 青岛 266042)

摘要: 在立体视觉测量中,为获得更高精度,往往需要将视差计算精确到亚像素级。将互信息理论引入双目图像配准并结合多分辨率技术实现亚像素级点匹配。采用 Bouguet 立体校正算法对左右图像进行极线校正,利用 Harris 角点探测器检测目标并将获取角点作为待匹配点,采用最大互相关法进行搜索确定像素级匹配点。然后对以左右匹配点为中心的 20×20 邻域图像进行插值并分别放大 10 倍和 100 倍,采用互信息方法先对低分辨率图像进行配准,再在高分辨率图像上进一步细化求精,结合像素级匹配的整数视差可得最终亚像素级视差。实验结果表明,该方法能将视差精度提高到 0.01 像素。

关键词: 立体视觉; 图像配准; 互信息; 多分辨率; 亚像素

中图分类号: V271.4 **文献标志码:** A **文章编号:** 1671-637X(2015)04-0023-04

Sub-Pixel Point Matching Method of Stereo Vision Based on Mutual Information Theory

SUN Shao-jie, YANG Xiao-dong, REN Ji-chang
(Navy Submarine Academy, Qingdao 266042, China)

Abstract: In order to improve the range-measuring accuracy in stereo vision, a sub-pixel parallax is needed. The Mutual Information (MI) theory is combined with multi-resolution method to realize the goal of sub-pixel point matching. Firstly, Bouguet algorithm is used to rectify the left and right images making their epipolar line forward-parallel. Then Harris corner detector is brought to find a most characteristic corner as candidate matching point. After that, Most Cross Correlation matching rule is introduced to search the matching point. The 20×20 areas whose center is the left and right matching point are magnified by 10 and 100 times respectively. The low-resolution image is registered with MI theory, followed by the seeking of higher precision in the high-resolution image. Finally, combined with the integer-grade parallax, we can get the sub-pixel parallax. The experimental result shows that the method used in this article can improve the precision to 0.01 pixel level.

Key words: stereo vision; image matching; mutual information; multi-resolution; sub-pixel

0 引言

图像匹配^[1]是立体视觉的关键环节,其主要目的在于寻找左右图像对应特征点从而计算出目标视差,为视觉测量和三维场景重建等立体视觉应用提供基础。在基于视觉的远距离目标测量中,匹配点之间的视差一般很小,常规像素级匹配已经无法满足高精度测量的需求,因而需要寻找方法实现亚像素级匹配,即

将视差的计算精度提高到亚像素级。

1995 年文献[2]首次提出将信息论应用于图像配准中;近年来基于互信息的图像配准技术在诸多领域取得较快发展,文献[3-5]将互信息理论用于医学影像的配准中;文献[6-7]运用互信息理论实现遥感图像的配准,均获得良好的匹配效果,可达到亚像素级精度,实验表明这是一种鲁棒性很强的图像配准技术。但在配准过程中出现的大量互信息计算制约了其处理速度,不能保证实时性需求。

本文将基于互信息的图像配准技术引入立体视觉点匹配中,并与多分辨方法结合,以实现亚像素级配准并提高算法实时性。首先根据最大互相关测度对双目

收稿日期:2014-05-12 修回日期:2014-09-11

基金项目:海军装备部军内科研项目

作者简介:孙少杰(1985—),男,湖北咸宁人,博士生,研究方向为机器视觉导航。

图像进行像素级点匹配,然后以左右匹配点为中心选取相同大小邻域并做插值放大,再利用互信息算法进行二次配准以实现亚像素级精度。此外,为提高配准速度,本文采用多分辨率方案,先对低分辨率图像配准,再将该值作为高分辨率图像初始值进一步细化配准结果,同时采用优化算法减少搜索数据量。

1 立体视觉测量精度分析

在平行双目视觉测量模型中,目标的深度信息可表示^[8]为

$$z = \frac{Tf}{D} \quad (1)$$

式中: T 为摄像机基距; f 为有效焦距; D 为视差。由式(1)可得给定深度值的精度为

$$\Delta z = \frac{z^2}{Tf} \Delta D \quad (2)$$

由摄像机图像采样原理可知,图像采样过程中最大化误差可达 0.5 像素,因而,双目图像视差的理论最大化误差 ΔD 为 1 个像素。对于本文实验所用双目摄像机(如图 1 所示),焦距 $f = 35 \text{ mm}$,两摄像机基距 $T = 30 \text{ cm}$,CCD 像元尺寸为 $6.45 \mu\text{m} \times 6.45 \mu\text{m}$,若要测量深度为 500 m 的目标,由式(2)可知,当前深度最大误差为 $\Delta z = 153.57 \text{ m}$,最大测距误差达到 30.7%,这显然无法满足远距离目标测量要求。因此,在不改变双目测量系统硬件结构(焦距、基距、摄像机分辨率等)的情况下,迫切需要设法减小 ΔD 以达到提高测量精度目的,这显示了将图像配准精确到亚像素级水准的必要性。

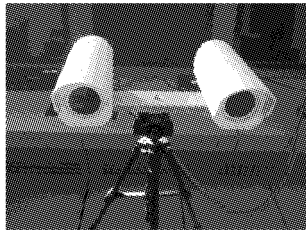


图 1 双目视觉测量系统

Fig. 1 Binocular vision measuring system

2 互信息配准与多分辨率插值

2.1 互信息配准原理

互信息是信息论中的一个基本概念,用来描述两个随机变量间的统计相关性^[9],可作为一个变量包含另一个变量信息量多少的度量。它可用熵来描述

$$I(A, B) = H(A) + H(B) - H(A, B) \quad (3)$$

式中: $H(A)$ 和 $H(B)$ 分别为系统 A 和 B 的熵; $H(A, B)$ 为二者的联合熵,分别表示为

$$H(A) = - \sum_a p_A(a) \text{lb} p_A(a) \quad (4)$$

$$H(B) = - \sum_b p_B(b) \text{lb} p_B(b) \quad (5)$$

$$H(A, B) = - \sum_{a,b} p_{AB}(a, b) \text{lb} p_{AB}(a, b) \quad (6)$$

式中: $a \in A$; $b \in B$; $p_A(a)$ 和 $p_B(b)$ 分别是 A 和 B 完全独立时的概率分布,由分别计算图像 A 和 B 的归一化直方图得到; $p_{AB}(a, b)$ 是 A 和 B 的联合概率分布,可从图像 A 和 B 的归一化联合直方图得到。

当两幅图像在空间的位置完全一致时,其中一幅图像表达的关于另一幅图像的信息(也就是互信息) $I(A, B)$ 为最大值,本文即根据这个原理对双目图像进行亚像素配准。

2.2 图像插值与多分辨率配准

在计算两图像间的互信息时,为提高配准精度,需对图像进行插值放大以提高分辨率。常用的插值方法有最近邻插值、双线性插值和双三次插值^[10-11]。一般来说,在这 3 种插值技术中,最近邻插值速度最快,效果最差,而双三次插值效果最佳,但速度最慢,双线性插值效果和速度介于两者之间。

图 2b~2d 是对图 2a 中方框区域分别采用最近邻插值、双线性插值和双三次插值法放大 50 倍后的结果,通过对比可以看出图 2d 的插值效果要优于其他两种方法。

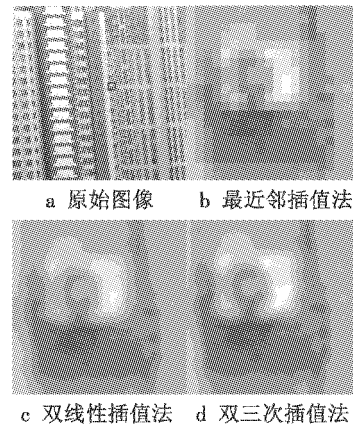


图 2 三种插值方法效果比较

Fig. 2 Comparison of three interpolate methods

为保证插值效果以提高亚像素级配准精度,本文采用双三次插值算法。同时,本文经过实验比较了不同插值放大倍数的计算速度,对图 2a 中区域分别进行插值放大 100 倍、500 倍和 1 000 倍,其对应的时间开销分别为 94 ms, 2.1 s 和 7.5 s,考虑到图像配准的实时性,本文对图像插值放大的最大倍数为 100 倍。

另一方面,由于利用互信息配准方法需计算联合直方图和进行图像插值,如果直接对整幅图像进行操作,处理时间较长。本文将多分辨率方法应用到图像配准中,先对低分辨率图像进行配准,再将该值作为高

分辨率图像初始值进一步细化配准结果,即运用由粗到精的配准策略以减小搜索的计算量,从而提高处理速度。

3 亚像素级匹配方法及实现

3.1 极线校正

在双目图像匹配中,当两图像平面极线水平对准时,匹配效率最高,视差计算也最简单。然而现实中的两台摄像机很难达到完全水平状态,也就几乎不可能产生准确行对准的成像平面,一个好的解决方法是对双目图像进行立体校正,采用一定的数学方法使立体图像行对准到前向平行结构上。

本文采用 Bouguet 立体校正算法^[8]对双目图像进行校正。该算法通过对立体标定获取的两摄像机间旋转矩阵和平移向量信息进行数学处理,使两幅图像中每一幅投影次数最小化,同时使得观测面积最大化,以达到让极线水平对准的效果。图3为经校正后左右摄像机对应的两幅图像,由此可以看出,校正后的图像处于极线平行对准状态。在后续匹配搜索过程中,只需在左右图像同一行进行搜索即可找到对应匹配点,从而提高匹配效率。



图3 极线校正结果

Fig. 3 Results of epipolar line rectification

3.2 像素级点匹配

3.2.1 特征角点检测与提取

本文采用文献[12]提出的角点检测方法对感兴趣区域进行角点检测。如图4所示,选取图3中右图像某窗户角为特征点,先确定一个固定大小的矩形邻域为角点检测区域(大致包含所选特征点即可),然后进行角点检测。按照上述方法对右图中目标角点进行坐标提取,得到该点像素坐标为 $P_r(209,347)$ 。



图4 角点检测及坐标提取

Fig. 4 Corner detecting and coordinate computing

3.2.2 匹配相似性测度

图像经极线校正后即可进行立体匹配,本文首先采用最大互相关测度作为像素级匹配准则。定义某邻域上的灰度相关值 $C(x,y)$ 来表示灰度相似性,具体操作方法是将在右图中已检测出角点 $P_r(x,y)$ 为中心的 $m \times n$ 灰度邻域作为范本,在左图中以互相关函数作为相似性测度搜索具有同样或相似灰度值分布的对应点邻域,实现同名点的匹配。互相关函数为

$$C(x,y) = \frac{\sum_{i=1}^m \sum_{j=1}^n [S_k(i,j) - \mu_k] \cdot [T(i,j) - \mu_T]}{\sigma_k \cdot \sigma_T} \quad (7)$$

式中: $S_k(i,j)$ 表示右图像样本窗口第 i 行 j 列图像点的灰度值; $T(i,j)$ 表示左图像待匹配窗口第 i 行 j 列图像点的灰度值; $\mu_k = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n S_k[i,j]$; $\sigma_k = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [S_k(i,j) - \mu_k]^2$; $\mu_T = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n T[i,j]$; $\sigma_T = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [T(i,j) - \mu_T]^2$; 当 $C(x,y)$ 取最大值时对应的左图像点 $P_l(x,y)$ 即为 $P_r(x,y)$ 最佳匹配点。

此外,为避免匹配搜索过程中因存在相同特征结构而造成歧义匹配,即一幅图像中某特征点在另一幅图像中存在多个相似候选匹配点的情况,引入额外的约束条件作为辅助判据,以便得到唯一准确的匹配。对于水平平移相机获得的左右图像 I_1 和 I_2 ,根据双目成像光学理论,同一场景点在左图 I_1 上的列坐标为 y_1 ,在右图 I_2 上的列坐标为 y_2 ,若将 y_1 和 y_2 在同一坐标尺度下重叠,则 $y_2 < y_1$ 一定成立,再结合欧氏距离的最近邻准则,作为图像配准相似特征进一步度量。

本文将右图提取的窗户左上角点作为待匹配点,将以该点为中心的 7×7 邻域作参考邻域,以相同大小窗口在左图对应行按上述匹配准则进行搜索,得到左图匹配点整像素坐标为 $P_l(421,347)$ 。

3.3 亚像素级配准

完成像素级初步匹配之后即可采用最大互信息理论结合多分辨率法进行亚像素级配准。分别以左右匹配点为中心选取一定大小的邻域作为待配准图像,例如本文选取匹配角点的 20×20 邻域,采用双三次插值对其进行插值放大10倍,如图5所示。

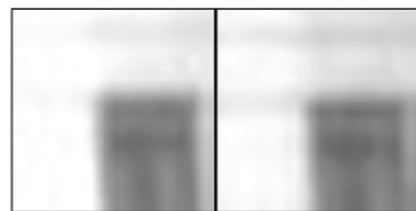


图5 邻域图像放大10倍

Fig. 5 The choosing area magnified by 10 times

在进行亚像素配准时,以左图作为参考图像,右图作为浮动图像进行平移,并计算每次平移后两图像重合区域的互信息值。当互信息值为最大时,说明两图实现完全配准,其平移量即为双目图像视差。将该视差除以插值放大的倍数,再结合整像素级匹配时得到的视差值即可得两图同名匹配点之间的亚像素级视差。

常规寻找最大互信息值的方法是让图像在全图范围进行平移搜索,并计算每次平移的互信息值,然后找出最大值对应的平移量,这种方法计算量较大。在实验过程中发现,右图按一定方向浮动配准时其互信息值符合上抛物线特征,即在达到最大值后开始下降。因此可基于这个特征设置停止搜索条件,即当互信息值不断增大并在历经某值后开始变小,则认为该值是最大互信息值。根据这个发现,本文设计如下算法优化平移配准过程。

步骤一 参考匹配点周围邻近点对应的互相关值确定平移方向,并给出平移量估计值。例如对图 5 计算可得 $P_i(421,347)$ 对应的互相关值为 $C=0.982\ 439$,而其左右邻近点 $P'_i(420,347)$ 和 $P''_i(422,347)$ 对应的互相关值分别为 $C'=0.972\ 945$, $C''=0.914\ 373$,如图 6 所示。这里 $C' < C, C'' < C$,且 $C' > C''$ 。根据 C, C', C'' 三者之间的关系可知实际的匹配点应位于 P'_i 和 P_i 之间,因而需将右图向右移动。反之若 $C' < C''$,则右图向左平移。同理,考虑匹配点上下邻近点的互相关值可确定右图向上还是向下平移。

	420	421	422
346		0.944 272 (421,346)	
347	0.972 945 (420,347)	0.982 439 (421,347)	0.914 373 (422,347)
348		0.909 007 (421,348)	

图 6 匹配点邻近点的互相关值

Fig. 6 Correlation value of neighboring matched point

步骤二 本步骤实际过程又可细分为以下 4 个小步骤进行:

- 1) 利用匹配点及其上下邻近点的互相关值确定 y 方向的平移方向并估计平移量 y_0 ,这里的 y_0 可通过对 3 个互相关值的抛物线拟合获得;
- 2) 将右图向上或向下移动 y_0 ,再在 x 方向上平移并依次计算互信息值,达到停止搜索条件后,记录平移量 x_0 ;
- 3) 固定 x 方向的平移量,在 y 方向上重复步骤 1),得到 y_1 ;
- 4) 再次固定 y_1 ,重复步骤 2),得到 x_1 ,检验 $x_1 = x_0$ 是否成立,若成立则停止搜索,否则重复 3),4) 直到 $x_1 = x_0$ 成立。

经多次实验表明,第一次检验 $x_1 = x_0$ 就已经成立,这样运算次数一般不超过 3×20 ,远小于 20×20 ,从而提高了配准速率。

按照上述优化算法对左右插值 10 倍图像进行初配准,得到平移量为 $(x, y) = (-6, -3)$ (左负右正,下同),归算到原图像即为 $x = -0.6$ (视差只需考虑 x 方向)。再按照同样的插值方法将匹配点邻域图像放大 100 倍,由第一次匹配得到的平移量可给定 x 方向搜索范围 $D_x = [(0.6 - 0.1) \times 100, (0.6 + 0.1) \times 100] = [50, 70]$,同理可确定 y 方向的搜索范围 D_y 。按照上述优化算法的步骤二再次进行配准得到平移量为 $(x, y) = (-63, -31)$,归算到原图像即为 $x = -0.63$,最终可得左右图像的视差为 $D = 421 - 209 + (-0.63) = 211.37$ 。

图 7 显示的是浮动图像(10 倍插值)在 $x-y$ 方向同时平移的互信息值变化三维图,从中可得出平移向量为 $(x, y) = (6, 3)$ 。同时还可看到,互信息值在 x, y 方向的变化均符合上抛物线特征,这也验证了上述优化过程中搜索策略的正确性。

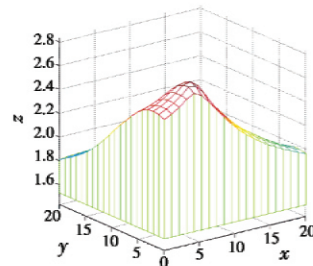


图 7 配准结果

Fig. 7 Results of stereo matching

4 结 语

本文将互信息理论应用于立体视觉图像点匹配过程中,结合多分辨率方法实现双目图像的亚像素级点匹配,进而获取更为精确的视差值。该方法无需对图像进行预处理,具有较强的鲁棒性,且经过算法优化之后,运算量得以减小,提高了图像配准速度。实验结果表明,该方法能有效实现双目图像亚像素级点匹配,并将视差精度提高到 0.01 像素,适用于高精度的三维测量等双目视觉应用场合。

参 考 文 献

[1] 张广军. 机器视觉[M]. 北京:科学出版社,2005:112-122. (ZHANG G J. Robot vision [M]. Beijing: Science Press, 2005:112-122.)

[2] COLLIGNON A, MAES F, DELAERE D, et al. Automated multi-modality image registration based on information theory[J]. Computational Imaging and Vision, 1995 (3): 263-274.

- 2010, 27(9):3588-3590.)
- [5] ZHEN C G, HAN P, LIN B H. Improved BP network-based realtime monitor for NO_x releasing from power planturnace[C]//IEEE International Conference on Machine Learning and Cybernetics, 2003. doi: 10. 1109/ICMLC. 2003. 1259638.
- [6] 张欣, 刘英, 高秀艳. 自适应投票快速中值滤波算法研究[J]. 计算机工程与应用, 2010, 46(6): 148-150. (ZHANG X, LIU Y, GAO X Y. Research on auto-adapted voting fast median filtering algorithm[J]. Computer Engineering and Applications, 2010, 46(6): 148-150.)
- [7] 李安安, 何坤, 詹晓倩, 等. 多尺度下的噪声图像边缘检测[J]. 计算机工程与设计, 2011, 32(1): 228-231. (LI A A, HE K, ZHAN X Q, et al. Multiscale edge detection for noise image[J]. Computer Engineering and Design, 2011, 32(1): 228-231.)
- [8] YEN J C, CHANG F J, CHANG S A. A new criterion for automatic multi-level thresholding [J]. IEEE Transactions on Image Processing, 1995, 4(3): 370-378.
- [9] 孟飞, 夏朝辉, 强钰捷, 等. 激光主动成像图像边缘检测算法研究[J]. 激光与红外, 2013, 43(10): 1192-1196. (MENG F, XIA Z H, QIANG Z J, et al. Edge detection algorithm of intensity image for laser active imaging [J]. Laser & Infrared, 2013, 43(10): 1192-1196.)
- [10] 李琦, 傅俊诚, 李自勤, 等. 激光雷达含噪图像边缘检测算法比较[J]. 红外与激光工程, 2003, 32(3): 239-243. (LI Q, FU J C, LI Z Q, et al. Comparison of the edge detection algorithms for lidar image with speckle noise[J]. Infrared and Laser Engineering, 2003, 32(3): 239-243.)
- (上接第26页)
- [3] 曾翎, 刘斐, 乔辉. 基于互信息的功能磁共振图像配准[J]. 电子科技大学学报, 2008, 37(1): 138-140. (ZENG L, LIU F, QIAO H. Functional magnetic resonance image registration based on the mutual information [J]. Journal of University of Electronic Science and Technology of China, 2008, 37(1): 138-140.)
- [4] 李琦, 姬红兵, 同鸣. 一种多模态医学影像鲁棒配准方法[J]. 西安电子科技大学学报, 2011, 38(6): 30-33. (LI Q, JI H B, TONG M. Robust registration of multimodality medical images based on the principal ordinal feature and hybrid entropy[J]. Journal of Xidian University, 2011, 38(6): 30-33.)
- [5] 王丽, 孙丰荣, 王奕琨, 等. 基于互信息的颅脑MR影像序列的三维配准[J]. 计算机工程与应用, 2011, 47(31): 160-163. (WANG L, SUN F R, WANG Y K, et al. 3D registration of MR image sequence of human brain based on normalized mutual information [J]. Computer Engineering and Applications, 2011, 47(31): 160-163.)
- [6] 刘峰, 龚健雅. 一种基于多特征的高光谱遥感图像分类方法[J]. 地理与地理信息科学, 2009, 25(3): 19-22. (LIU F, GONG J Y. A classification method for high spatial resolution remotely sensed image based on multi-feature [J]. Geography and Geo-Information Science, 2009, 25(3): 19-22.)
- [7] 周海芳, 杜云飞, 杨学军, 等. 基于互信息的遥感图像区域配准并行算法的研究与实现[J]. 中国图象图形学报, 2010, 15(1): 174-180. (ZHOU H F, DU Y F, YANG X J, et al. Study and implement of parallel region-based registration algorithm based on mutual information for remote-sensing images[J]. Journal of Image and Graphics, 2010, 15(1): 174-180.)
- [8] BRADSKI G, KAEHLER A. Learning OpenCV: Computer vision with the OpenCV library [M]. Beijing: Tsinghua University Press, 2009: 452-473.
- [9] MAES F, COLLIGNON A, DIRK V, et al. Multimodality image registration by maximization of mutual information [J]. IEEE Transactions on Medical Imaging, 1997, 16(2): 189-198.
- [10] 刘喜平, 龚晓彦, 郭希娟. 基于互信息的医学图像匹配中的改进插值算法[J]. 计算机仿真, 2010, 27(7): 194-197. (LIU X P, GONG X Y, GUO X J. Expanded interpolation in medical image registration based on mutual information[J]. Computer Simulation, 2010, 27(7): 194-197.)
- [11] 张峻豪, 孙焱, 詹维伟. 基于加权互信息的多模图像配准算法[J]. 计算机工程, 2012, 38(16): 207-211. (ZHANG J H, SUN Y, ZHAN W W. Multi-modality image registration algorithm based on weighted mutual information[J]. Computer Engineering, 2012, 38(16): 207-211.)
- [12] HARRIS C, STEPHENS M. A combined corner and edge detector[C]//Proceedings of the 4th Alvey Vision Conference, 1988: 147-151.