

# 宽基线主动视觉中感兴趣目标的对应技术

刘博文 余松煜 徐奕 杨小康

(上海交通大学电子工程系图像通信与信息处理研究所, 上海 200240)

(上海交通大学上海市数字媒体处理与传输重点实验室, 上海 200240)

**摘要** 在主动视觉系统中, 通常需要多个代理对同一场景中的感兴趣目标进行协同处理, 以提高系统智能分析感兴趣目标的能力。其中, 基于多视几何关系解决感兴趣目标的对应问题是协同处理的基础。一方面, 主动视觉系统一般工作在宽基线条件下, 这增加了对应问题描述的复杂性; 另一方面, 主动视觉系统以最佳视角观察目标, 因此摄像头需做实时的姿态调整, 由此导致的视间几何关系变化进一步加深了对应问题的解决难度。本文基于仿射不变的几何特征, 建立宽基线条件下的多视几何关系, 并针对频繁使用几何特征不能满足主动视觉系统实时要求的问题, 提出一种快速更新多视几何关系的方法, 并在多视几何约束下实现对感兴趣目标的鲁棒标识。实验结果表明, 该方法能解决宽基线主动视觉系统中感兴趣目标的复杂对应问题, 并能达到实时要求。

**关键词** 主动视觉 宽基线 感兴趣目标 对应问题 多视几何关系

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2007)10-1917-05

## Correspondence of Object-of-Interest in Wide Baseline Active Vision

LIU Bo-wen, YU Song-yu, XU Yi, YANG Xiao-kang

(Institute of Image Communication and Information Processing, Department of Electronic Engineering,  
Shanghai Jiaotong University, Shanghai 200240)

(Shanghai Key Laboratory of Digital Media Processing and Transmissions, Shanghai Jiaotong University, Shanghai 200240)

**Abstract** In active vision system, intelligent analysis of objects-of-interest (OOI) in 3D scene generally needs multi-agent cooperation, which depends heavily on the accuracy of multi-view geometry computation and the robust correspondence of OOI. In most cases, active vision system captures multiple views under wide-baseline stereo model, which results in distinct affine distortions between views and complex description of correspondence problem. Furthermore, the difficulty of depicting correspondence problem in active vision is aggravated when the cameras perform real-time pose adjustment for active OOI tracking with the best viewpoint. For abovementioned issues, this paper proposes a method to achieve real-time multi-view geometry updating on the basis of constructing geometric feature invariants. Then a robust OOI correspondence algorithm is advanced with the acquired multi-view geometry constraint. The experimental results demonstrate that our method can obtain robust OOI correspondence under wide baseline stereo model and be suitable for real-time application in active vision system.

**Keywords** active vision, wide baseline, correspondence of object-of-interest, multi-view geometry

## 1 引言

近年来, 主动视觉系统已经受到广泛关注, 它是监控与跟踪场景中感兴趣运动目标 (OOI) 的强有力

工具, 可用于各种监控系统<sup>[1]</sup>, 如机器人伺服系统<sup>[2]</sup>等。系统拥有多个代理 (一个代理包括 PTZ (pan tilt zoom) 摄像头与 PC 机, 完成摄像头控制与图像处理的功能), 它们扫描工作场景中的感兴趣目标, 并互相通信以协同处理这些感兴趣目标的有

基金项目: 上海博后基金项目 (06R214138)

收稿日期: 2007-07-05; 改回日期: 2007-07-19

第一作者简介: 刘博文 (1983 ~ ), 男。现为上海交通大学硕士研究生。主要研究方向为多视图图像匹配与应用。E-mail: bowenforever@hotmail.com

关信息。一方面,代理的主动跟踪克服了传统的被动视觉系统中固定摄像头视野范围有限、观测视角不佳等诸多不足;另一方面,多代理的协同处理能提高整个系统的跟踪置信度。各个代理的最佳观测视角不同,所使用的摄像头传感器也可能不同,因此每个代理关于感兴趣目标的主动跟踪置信度会不一致,比如感兴趣目标处在某个代理视角范围的边界区域而导致该代理的主动跟踪过程不够稳定,或者安装红外传感器的代理在伪目标的干扰下往往比安装可见光传感器的代理在主动跟踪性能上更可靠,因此各个代理之间应借通信对感兴趣目标信息进行协同处理,使具备高跟踪置信度的代理能引导低跟踪置信度代理的主动跟踪过程,从而使整个主动视觉系统实现对感兴趣目标的可靠跟踪。图 1 所示为红外代理引导可见光代理的一个典型的两代理主动视觉系统。

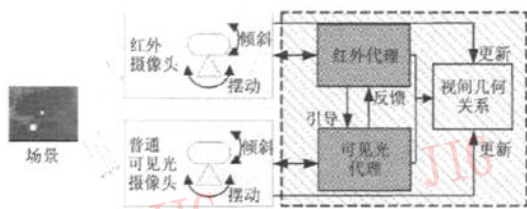


图 1 一个典型的两代理主动视觉系统

Fig. 1 A typical two-agent active vision system

当场景中存在多个感兴趣目标时,多代理的协同处理必定要求在各个视图中能逐一标识出所有对应的感兴趣目标。通常需要借助多视之间的几何约束将感兴趣目标的对应问题变成适定的,显著缩减对应解的搜索空间,降低匹配结果的歧义性。但是,主动视觉系统的监控范围较大,而代理的个数十分有限,代理之间的分布间隔相对于监控距离往往不能忽略不计,由此产生的宽基线成像模型会导致视间存在显著的仿射畸变,使得传统方法中相似尺度的匹配基元失去了选择意义,从而增加了对应问题的复杂性。同时,协同式主动跟踪要求各个代理调整摄像头的姿态,以达到更佳综合观测和分析感兴趣目标的效果,因此视间几何关系是随时间不断变化的。如果按照惯常的思路,通过特征匹配重复计算多视几何关系来约束感兴趣目标的对应过程,则难以达到系统的实时要求。

针对以上问题,本文通过构造几何特征不变量建立宽基线成像条件下的视间几何关系,并提

出一种随摄像头姿态调整实时更新视间几何关系的方法。另外,外极几何约束通常只能将对应点的搜索空间限制在 1 维的外极线上,仍然存在一定的匹配歧义性。为进一步降低算法的复杂度而提高系统的实时性能,文中提出一种点对点的几何约束方法,能快速鲁棒地确定多个视图中的对应感兴趣目标。实验结果表明,本文方法能解决宽基线主动视觉系统中感兴趣目标的复杂对应问题,并能达到实时要求。

## 2 感兴趣目标的对应技术

匹配基元的选择和描述是解决对应问题的关键,大多数对应技术选择相似尺度的匹配基元<sup>[3]</sup>,或者采用自适应窗技术缓解仿射畸变问题<sup>[4]</sup>。但是,全局变换、局部扰动和部分遮挡等复杂现象是宽基线成像模型下的固有问题,前述方法中的特征度量将随这些现象呈现难以预测的变化。目前,一些研究者将几何特征不变量的概念引入到对应技术中,较好地解决了宽基线视觉系统的匹配问题,其中 SIFT(scale invariant feature transform)几何特征不变量尤为广泛应用<sup>[5]</sup>。本文通过 SIFT 特征匹配的鲁棒结果初始化宽基线主动视觉系统的多视几何关系,采用四元代数理论解决时变多视几何关系的实时更新问题,并基于多视几何的计算结果提出一种双外极线约束下的感兴趣目标的对应技术。

### 2.1 基于几何特征不变量建立多视几何关系

本文选择 SIFT 几何特征不变量初始化主动视觉系统中的多视几何关系。SIFT 算法对各视图提取连续尺度空间上的特征点,然后每一特征点的邻域由一个 128 维的矢量来描述,其中每 1 维对应着一个特征方向统计量。SIFT 算法检测出的特征点及其描述子能够在缩放和旋转变换下保持不变,并对光照变化不敏感,适于解决宽基线主动视觉中的对应问题。之后特征点应用于多视几何关系,可得到最典型的约束关系——基础矩阵。

根据外极几何约束理论<sup>[6]</sup>,对于图像 1 中的某点  $m_1$ ,它在图像 2 中的对应点  $m_2$  必然位于与  $m_1$  对应的外极线上,即若  $m_1$  与  $m_2$  为同名的对应点,则它们必然要满足下面的公式:

$$\tilde{m}_2^T F \tilde{m}_1 = 0 \quad (1)$$

式中,  $\tilde{m}_1, \tilde{m}_2$  为  $m_1, m_2$  的同构坐标形式,  $F$  即为基础矩阵。为得到足够多(至少 8 对)的对应特征点

对,本文遍历所有特征点的描述子向量并找出其中欧氏距离最近的点作为对应的特征点对。然后在式(1)基础上利用 RANSAC (random sample consensus)方法<sup>[6]</sup>筛选掉置信度低的异常点,就得到了鲁棒的基础矩阵。

### 2.2 基础矩阵快速更新算法

SIFT 算法在提取匹配基元时必须遍历尺度空间中的所有像素点,因此具有较高的计算复杂度,其效率不能满足系统的实时性要求,因此本文只将其作为初始化多视几何关系的基础,基础矩阵的快速更新则需要基于四元数分解来实现。

一般地,主动视觉系统的主动跟踪功能是通过摄像头的姿态调整得以实现。姿态调整包括 3 种情形:摆动(Pan),倾斜(Tilt)和缩放(Zoom),因此它只改变两视图之间的旋转关系,而并不影响视间的平移关系。本文对两视图之间的基础矩阵  $F$  进行分解,得到

$$F = A_2^{-T} E A_1^{-1}, \text{其中 } E = [t]_{\times} R \quad (2)$$

式中,  $3 \times 3$  矩阵  $A_1, A_2$  分别为两摄像机的内部参数矩阵,主要表征摄像机的焦距和中心点参数,可以通过摄像机定标方法获得<sup>[7]</sup>。 $E$  为表征两摄像头位置关系的本征矩阵,它计算摄像头 1 坐标系变换到摄像头 2 坐标系的刚体变换(旋转  $R$  和平移  $t$ ),下标“ $\times$ ”表示向量叉乘运算。本文算法只需要对两视之间的旋转矩阵  $R$  做不断更新就可以更新基础矩阵  $F$ 。为了得到初始的旋转矩阵,本文将  $E$  分解,由式(2),本文将问题归结为  $\min \|E - [t]_{\times} R\|^2$ ,基于四元代数理论,可将矩阵的相乘转化为两个四元数的乘法,从而通过拉格朗日乘子法即可求出满足最小条件的四元数所对应的旋转矩阵  $R$ 。

不失一般性,本文将其中一个代理用于跟踪。以摄像头光心为摄像头坐标系原点,光轴为  $Z$  轴,平行于地面的水平线为  $X$  轴,垂直于  $XOZ$  平面为  $Y$  轴,将摄像头镜头转动统一为先绕  $Y$  轴摆动再绕  $X$  轴转动。则摄像头先绕  $Y$  轴摆动  $\alpha$  角后坐标系变换矩阵  $R_y$ , 以及再绕  $X$  轴转动  $\beta$  角后的坐标系变换矩阵  $R_x$  分别为

$$R_y = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix}, R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \quad (3)$$

若两摄像头间初始的旋转矩阵为  $R_0$ , 则更新后的旋转矩阵为  $R_{\text{NEW}} = R_0 R_x R_y$ , 于是  $E_{\text{NEW}} = [t]_{\times} R_{\text{NEW}} = [t]_{\times} R_0 R_x R_y = E_0 R_x R_y$ , 从而由式(2)可以得到更新

后的基础矩阵为

$$E_{\text{NEW}} = A_2^{-T} E_0 R_x R_y A_1^{-1} = F_0 A_1 R_x R_y A_1^{-1} \quad (4)$$

从而可以在初始的  $F_0$  基础上快速更新  $F_{\text{NEW}}$ 。

### 2.3 双外极线约束下的感兴趣目标的对应

由于视图之间存在外极几何约束,即图像 1 中的某点  $m_1$  在图像 2 中的对应点  $m_2$  必然位于与  $m_1$  对应的外极线上,因此我们得到了两幅图像之间点与线的对应关系,使得原本需要在 2 维空间中进行的对应点搜索任务降低至 1 维空间中。虽然这种约束已经大大降低了感兴趣目标对应问题的歧义性和计算复杂度,但是由于这种搜索依赖于图像内容,因此仍然不能满足主动视觉中一些特定应用的要求,例如代理  $A$  引导代理  $B$  实现主动跟踪时,如果感兴趣目标还未进入代理  $B$  的视角范围,那么即使有准确的外极线约束,仍然无法通过匹配确定感兴趣目标相对于代理  $B$  的图像坐标系的有效位置。通常来说,可以通过计算单应矩阵,在单应矩阵约束下实现感兴趣目标坐标的视间一一映射。可是单应矩阵约束仅在两视图的内容所对应的场景处于同一平面时有效,如果各摄像机距离整个场景比较近且视角差异较明显时,该前提往往无法成立。

本文提出一种双外极线约束的对应技术,其巧妙地利用了图像模糊会对 SIFT 算法结果产生影响的事实,使得在不同尺度下同一个感兴趣目标的质心位置能够映射出方向不同的外极线,从而可以直接利用多条外极线的交叉位置估算视间的点对点坐标变换。该方法不仅解决了不可见感兴趣目标的主动跟踪问题,还进一步降低了对应目标的搜索空间。在后续实验中,本文采用了双外极线约束,如图 2 所示。具体算法流程如下:

(1) 对初始的图像 1, 2 使用 SIFT 算法求得它们之间的基础矩阵  $F$ , 从而将图像 1 中的感兴趣目标  $m$  映射成图像 2 中的对应的外极线  $l_m$ 。

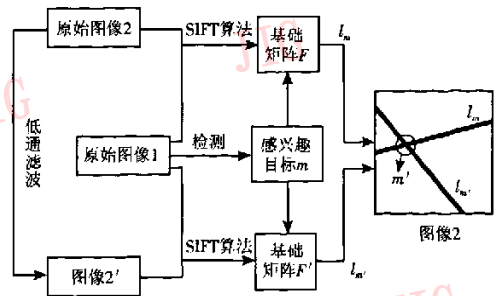


图 2 双外极线交叉方法寻找对应 OOI  $m$  和  $m'$   
Fig. 2 Cross of two epipolar lines for OOI  $m$  &  $m'$

(2) 将原始图像 2 进行高斯低通滤波得到图像 2', 然后再将其与原始图像 1 使用 SIFT 算法求得它们之间的基础矩阵  $F'$ , 从而将图像 1 中的感兴趣目标  $m$  映射成图像 2' 中的对应的外极线  $l_m'$ 。

(3) 由于高斯低通滤波不会改变图像中感兴趣目标的位置, 因此  $m$  在图像 2 与 2' 中的对应点是相同的, 但低通滤波会改变 SIFT 算法中的尺度空间金字塔, 故而会改变匹配特征点的位置与数量, 使得  $F \neq F'$ , 于是  $l_m \neq l_m'$ , 而  $m$  在图像 2 中的对应点既在  $l_m$  上又在  $l_m'$  上, 因此  $l_m$  与  $l_m'$  的交点必然就是  $m$  的对应点。同时由于  $m$  的对应点真实存在, 因此这两条外极线不会出现平行的情况。

(4) 低通滤波高斯窗的选取: 经过实验证明, 高斯窗口的大小以 3 ~ 7 为宜, 若小于 3 则会因为  $l_m$  与  $l_m'$  太相近导致寻找它们的交点时有很大的误差; 若大于 7 则会因为图像 2' 中低频分量太普遍导致 SIFT 算法的鲁棒性明显地降低。

双外极线约束方法大大降低了对应点的搜索范围, 对图像场景的平面性不敏感, 并且两次 SIFT 算法可以同时运行, 基本不增加时间复杂度, 同时很好地解决了对应问题的不稳定性, 使得多代理能够更快速有效的结合在一起。

### 3 实验结果及分析

在验证宽基线主动视觉系统中基础矩阵快速更新方法以及感兴趣目标对应技术的准确性实验中, 使用了多摄像头协同处理与目标跟踪项目的系统平台, 实验中采集了两个摄像头同一帧的两幅图像作为测试输入, 做如下实验: (1) 使用匹配特征点计算两幅图像的视间几何关系, 即基础矩阵, (2) 验证了基础矩阵快速更新方法的准确性, (3) 在两个采集流中分别使用二值图运动信息检测方法检测出这两幅图像中的运动目标, 定为感兴趣目标  $m_1$  和  $m_2$ , 并将  $m_2$  作为  $m_1$  对应点的实际参考值, 与双外极线约束方法计算得到的  $m_1$  的对应点相比较, 得到均方误差作为输出, 并将其与传统的单应矩阵法做比较, 表征双外极线约束法的鲁棒性与优势所在。

图 3(a), 图 3(b) 给出了宽基线条件下求得视间几何关系的结果, 匹配特征点用红色数字标出, 蓝色线条即为每个特征点对应的外极线, 然后使用 RANSAC 方法即可求出基础矩阵; 图 3(c) 为利用图 3(a), 图 3(b) 得到的基础矩阵映射出的外极线, 而

图 3(d) 则是在摄像头向右转过 3° 后由式 (4) 更新得到的基础矩阵映射出的外极线, 它与真实的基础矩阵误差的模值为 2.2849, 误差在可接受范围内, 由图中也可以看出 OOI 橙色小车基本位于外极线的附近, 而其运算速度比 SIFT 算法快了一个数量级, 因此本文提出的基础矩阵更新算法是符合主动视觉应用要求的。

图 4 给出了使用双外极线约束法得到对应的感兴趣目标的实验结果。先将图像 1 中检测到的感

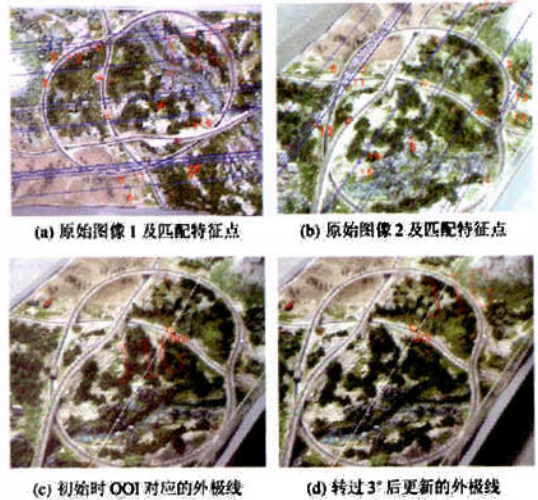


图 3 宽基线条件下的多视几何关系及基础矩阵的快速更新

Fig.3 Multi-view geometry and its fast update

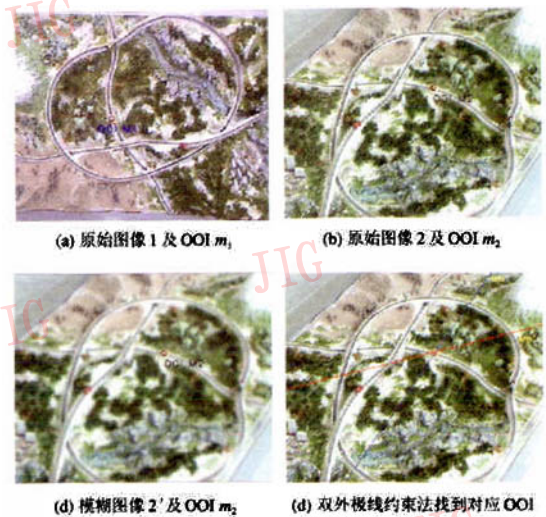


图 4 宽基线条件下感兴趣目标对应结果

Fig.4 Results showing correspondence of OOI

趣目标  $m_1$  利用图像 1,2 间的基础矩阵映射到图像 2 中的外极线  $l_m$  (图 4(d) 中的蓝色(深色)直线),另一方面,实验中统一使用高斯窗大小为 7 的低通滤波器对图像 2 进行模糊,之后利用图像 1,2' 间的基础矩阵将  $m_1$  映射到图像 2 中的外极线  $l_m$  (图 4(d) 中的红色(淡色)直线),两条外极线的交叉点即为求得的  $m_1$  在图像 2 中的对应点。

表 1 给出了单应矩阵法与双外极线约束法的实验比较结果。在实验平台中,橙色小车与场景背景基本在同一平面上,环形轨道上的黑色小车略高于背景,而红色小车位于最高的坡道上。从表中可以看到,双外极线约束法的误差总体小于单应矩阵的误差,这一差距对于红色车更为明显,由于红色车不在背景平面上,单应矩阵法误差很大,但双外极线约束法的结果比较稳定。同时由于黑色车靠近场景边缘,而场景边缘一般是两幅图像重叠区域比较少的位置,因此单应矩阵和基础矩阵在边缘区域的准确性都有所下降,导致两种方法对黑色车的误差都增大了,甚至超过了平面性更差的红色车。在实时性方面,两种方法的运算效率都十分高,都能在一帧时间内完成,因此能满足实时性的要求。

表 1 两种对应技术用于图 3 中各 OOI 的结果比较

Tab.1 Comparison of two correspondence techniques of OOI

坐标	橙色小车	红色小车	黑色小车
OOI $m_1$	(136,147)	(234,183)	(134,49)
OOI $m_2$ (Ref.)	(204,110)	(100,160)	(310,153)
单应矩阵法 $H$	(203.2,108.1)	(95.13,166.9)	(307.3,161.5)
$(H, m_2)$ -MSE	2.06	8.45	8.92
双外极线法 $E$	(205.6,109.7)	(100.2,161.7)	(312.2,151.9)
$(E, m_2)$ -MSE	1.62	1.75	2.47

最后本文得出这样的结论,双外极线约束法普遍好于单应矩阵法,它对场景的平面性不太敏感,但是对于靠近图像边缘的目标准确性会有些许降低;而单应矩阵法对场景的平面性和目标的位置都十分敏感,即使在用于与场景在同一平面上的感兴趣目

标时也差于双外极线约束法,只是在实验中运算时间略微胜出。

## 4 结 论

多代理协同处理特性体现了主动视觉系统的优势所在,它使得基于主动视觉的一系列应用有了更全面更鲁棒的结果。其中感兴趣目标的对应问题是协同处理的基础,本文提出的双外极线交叉法很好地解决了宽基线条件下感兴趣目标的对应问题,同时提出的快速更新视间几何关系的方法满足了主动视觉系统实时性的要求,它们为基于主动视觉系统的应用提供了良好的开发基础。

### 参考文献 (References)

- 1 Bakhtari A, Benhabib B. Agent-based active-vision system reconfiguration for autonomous surveillance of dynamic, multi-object environments [A]. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems [C]. Edmonton, Alberta, Canada, 2005: 1373 ~ 1378.
- 2 Fayman J A, Rivlin E, Christensen H I. A system for active vision driven robotics [A]. In: Proceedings of IEEE International Conference on Robotics and Automation [C]. Minneapolis, MN, USA, 1996: 1986 ~ 1992.
- 3 Sun J, Shum H Y, Zheng N N. Stereo matching using belief propagation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7): 787 ~ 800.
- 4 Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: theory and experiment [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, Sept., 1994, 16(9): 920 ~ 932.
- 5 Lowe D. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91 ~ 110.
- 6 Zhang Z. Determining the epipolar geometry and its uncertainty: a review[J]. International Journal of Computer Vision, 1998, 27(2): 161 ~ 195.
- 7 Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330 ~ 1334.