

双目协同多分辨率主动跟踪方法

崔智高, 李艾华, 姜柯, 苏延召, 金广智

(第二炮兵工程大学 502 教研室, 陕西 西安 710025)

摘要: PTZ(pan-tilt-zoom)相机由于其具有可变视角和可变分辨率能力,在视频监控领域得到了广泛的应用。该文针对智能监控的需求,提出了一种基于双目 PTZ 相机的多分辨率主动跟踪方法。该方法分为离线标定和在线协同跟踪两部分。离线标定部分,提出了一种基于图像特征匹配的单目自标定和基于目标运动信息的双目自标定方法,该方法操作简单,无需标定物,在最大程度上减小了对人工干预的依赖,在此基础上推导了系统所具有的两个重要性质;在线协同跟踪部分,设计了一种分段静止的协同跟踪策略。通过实际监控场景下的视频实验,验证方法的有效性和可行性。实验结果表明,该方法可以在复杂环境下有效的主动跟踪目标,在智能监控领域具有较广泛的应用前景。

关键词: 目标跟踪; 主动跟踪; PTZ 相机; 双目协同; 自标定

中图分类号: TP391 **文献标志码:** A **文章编号:** 1007-2276(2013)12-3509-08

Multi-resolution active tracking algorithm using two-camera collaboration

Cui Zhigao, Li Aihua, Jiang Ke, Su Yanzhao, Jin Guangzhi

(Faculty 502, The Second Artillery Engineering University, Xi'an 710025, China)

Abstract: PTZ (pan-tilt-zoom) cameras have been widely used in visual surveillance domain due to their ability to cover wide field of view and large scale range. For the large demand of intelligent visual surveillance, a novel multi-resolution active tracking method was presented based on two PTZ cameras. The proposed method consists of two stages: off-line calibration and on-line cooperative tracking. In off-line calibration stage, a novel method was presented that combines feature based single-PTZ camera self-calibration and object motion based dual-PTZ cameras self-calibration. The proposed approach doesn't require calibration tools and manual operation. Two important properties of proposed system were also inferred based on above calibration method. In on-line tracking stage, a piecewise cooperative tracking strategy was designed. The verification of proposed algorithm framework was implemented on real outdoor surveillance environment. Experimental results show that the proposed method can track the moving object effectively. The proposed framework can be used in intelligent visual surveillance with

收稿日期: 2013-04-12; 修订日期: 2013-05-13

基金项目: 国家自然科学基金(61021063, 61225008)

作者简介: 崔智高(1984-), 男, 博士生, 主要从事双目视觉监控等方面的研究。Email: cuizg10@126.com

导师简介: 李艾华(1966-), 男, 教授, 博士生导师, 博士, 主要从事信号处理、图像处理、模式识别、智能视觉监控等方面的研究。

Email: L863@163.com

wide application prospect.

Key words: object tracking; active tracking; PTZ camera; two-camera collaboration; auto-calibration

0 引言

运动目标跟踪^[1-2]是智能视频监控^[3]中的核心部分,具有重要的研究意义和应用价值。根据摄像机是否运动,目标跟踪可以分为两类^[4]:静止相机的被动目标跟踪和 PTZ 相机的主动目标跟踪。前者摄像机与监控场景保持相对固定,一般采用背景建模方法检测目标运动,并通过颜色、纹理以及轮廓等属性建立目标之间的轨迹关联。由于静止相机视场(Field of view)的限制和其被动性,当目标运动出其视场范围时,无法实现对目标的长时间持续跟踪。目前,已经有很多学者针对此类情况进行了研究,算法相对成熟。

为了扩大系统的监控范围,基于 PTZ 相机的主动跟踪系统得到了广泛研究与应用,该类系统通过比较目标理想位置与实际位置的差异,自动调整 PTZ 参数,从而使跟踪目标始终处于其可见视场内。此类情况研究较少,算法也相对复杂。Xue 等^[5]利用全景高斯混合模型辅助确定目标区域,并利用 Mean-shift 方法主动跟踪室外场景中的行人、车辆等目标;Kumar 等^[6]通过贝叶斯假设确定每个像素是否为目标像素,并用 Camshift 方法进行非刚体目标的主动跟踪;Varcheie 和 Bilodeau 等^[7]将光流法和模糊粒子滤波方法相结合,提出了一种低帧率条件下基于网络 PTZ 摄像机的人体主动跟踪方法;Bernardin 等^[8]将 KLT(Kanada-Lucas-Tomasi)算法与模糊控制相结合,提出了一种室内场景下人体上半身的主动跟踪方法。参考文献^[5-8]虽在一定程度上扩大了监控与搜索范围,但分辨率比较单一。相比于单目 PTZ 视觉系统,双目 PTZ 视觉系统可获得更丰富的信息,如同时获取目标运动的低分辨率全景图像和高分辨率近景图像等。虽然目前关于双目 PTZ 视觉系统的研究较少,但其具有很大的应用前景。

针对双目 PTZ 视觉系统,文中提出了一种便于工程实现的多分辨率主动跟踪方法,详细论述了系统的标定过程和协同跟踪策略,并通过实际监控场景下的多组实验,验证了方法的有效性和可行性。

1 离线标定

无论是单目视觉还是双目视觉,一般都需要进行标定。传统的单目视觉或双目视觉标定方法往往依赖于精密加工的 2D 或 1D 标定物^[9],然而,PTZ 相机具有可变焦能力,在相机视场较大(Zoom 较小)时,标定物将覆盖相机较小视场,从而造成巨大的误差,此外,该类方法需要实施相关手工操作。由于监控场景的复杂性和多样性,监控系统往往需要无需标定物和人工干预的自标定方法。为了方法的适用性和可操作性,文中提出了一种改进的基于特征匹配的单目 PTZ 自标定,以及基于目标运动信息的双目 PTZ 共面约束自标定方法。该方法无需标定物,在最大程度上减小了对人工干预的依赖,并在此基础上推导了系统所具有的两个重要性质。文中所提出的标定方法不受 PTZ 参数改变的影响,可以实现两相机在任意视角(对应于 pan-tilt 参数)和任意尺度(对应于 zoom 参数)下的协同跟踪,具有较强的实用性。

1.1 基于特征匹配的单目 PTZ 自标定

PTZ 相机参数固定时,可认为其符合弱透视针孔模型,即空间中一点与该点在图像上的投影满足:

$$x = \kappa K(z)R(p, t)X \quad (1)$$

式中: x 为齐次坐标; κ 为比例因子。 $R(p, t)$ 为相机外参矩阵,仅与 pan-tilt 参数有关,可表示为:

$$R(p, t) = \begin{bmatrix} \cos(p) & 0 & -\sin(p) \\ \sin(p)\sin(t) & \cos(t) & \cos(p)\sin(t) \\ \sin(p)\sin(t) & -\sin(t) & \cos(p)\cos(t) \end{bmatrix} \quad (2)$$

$K(z)$ 为相机内参矩阵,仅与 zoom 参数有关,可表示为:

$$K(z) = \begin{bmatrix} f(z) & s & c_x \\ 0 & \alpha f(z) & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

式中: (c_x, c_y) 为主点坐标; α 和 s 分别为 x 和 y 方向的像素纵横比和倾斜度,一般近似认为 $\alpha=1$ 和 $s=0$,以简化标定过程; $f(z)$ 为等效焦距(单位为像素)随 zoom 参数的函数关系,一般情况下满足非线性模型^[10],可

用指数函数进行表示:

$$f(z) = a_f e^{b_f z} + c_f e^{d_f z} \quad (4)$$

摄像机标定的目的就是要估计主点 c_x, c_y 以及非线性模型系数 a_f, b_f, c_f, d_f 。针对单目 PTZ 相机标定, Sinha 和 Pollefeys^[10]提出了一种基于特征匹配的自标定方法。该方法首先估计 zoom 最小时的焦距, 然后利用其他 zoom 下的图像与最小 zoom 下的图像进行特征匹配, 顺序估计各离散 zoom 下的焦距值。该方法存在如下限制: (1) SIFT^[11]、SURF^[12]等局部不变特征虽具有良好的尺度不变性, 但当两幅图像具有较大的尺度变化时, 特征匹配数目较少, 因此, 当 PTZ 相机具有较大 zoom 变化范围时, 较难或无法估计大 zoom 参数下的焦距; (2) 由于采用了顺序估计焦距的策略 (即首先估计最小 zoom 下的焦距, 再依次确定大 zoom 下的焦距), 有可能形成误差累积。在上述方法的基础上, 文中提出了一种改进的基于特征匹配的 PTZ 相机自标定方法, 通过在不同 zoom 参数下的并行计算估计焦距值, 避免了文中提到的两种问题, 提高了标定的准确性。

文中方法主要包括以下 3 个阶段。

(1) 估计主点 (c_x, c_y) 。保持相机 pan-tilt 参数不变, 改变 zoom 参数, 获得一组不同尺度下的图像序列 (文中所用 PTZ 相机的 zoom 变化范围为 0~18, 因此可通过采集 19 幅图像进行估计)。分别对相邻尺度图像进行 SIFT 特征匹配。设 (x, y) 和 (x', y') 为其中一组匹配特征点, 其所在图像的 zoom 参数为 z 和 z' , 等效焦距分别为 f 和 f' 。则由针孔摄像机模型的几何约束关系, 得到如下等式:

$$\frac{x - c_x}{x' - c_x} = \frac{y - c_y}{y' - c_y} = \frac{f}{f'} \quad (5)$$

约简得:

$$c_x(y - y') + c_y(x' - x) = x'y - xy' \quad (6)$$

式中, 理论上只需要 2 组匹配点即可估计主点坐标, 一般情况下, 匹配点数目远多于两组, 因此可以利用最小二乘方法求解主点的最优估计值。

(2) 估计离散 zoom 下的焦距 $f(z)$ 。对某个固定 z 值, 保持相机 zoom 参数不变, 少量改变 pan-tilt 参数, 获得参数为 (p, t, z) 和 (p', t', z) 的两幅图像。对其进行 SIFT 特征匹配, 对每一组匹配特征点满足:

$$x = \kappa \mathbf{K}(z) \mathbf{R}(p, t) \mathbf{R}(p', t')^{-1} \mathbf{K}(z)^{-1} x' \quad (7)$$

由于外参矩阵可通过已知的 pan-tilt 参数进行

计算, 主点已求得, 因此, 对每一组匹配特征点求解上式可获得焦距的一个估计值, 为了提高鲁棒性, 可以利用多组匹配特征点求取焦距的平均值。为进一步对焦距进行精确求解, 文中以匹配特征点矩阵变换前后坐标值欧式距离的平方和建立优化目标函数:

$$\begin{cases} E = \sum_{i=1}^n \|x - x''\|^2 \\ x'' = \kappa \mathbf{K}(z) \mathbf{R}(p, t) \mathbf{R}(p', t')^{-1} \mathbf{K}(z)^{-1} x' \end{cases} \quad (8)$$

以文中估计的焦距平均值作为初值, 利用 LM 优化算法^[13]计算得到焦距的最优估计值。

(3) 估计非线性模型系数 a_f, b_f, c_f, d_f 。通过步骤(2), 可获得多组离散对应关系 $\{z, f(z)\}$, 因此可通过模型拟合方法估计非线性模型系数。

1.2 基于目标运动的双目 PTZ 共面约束自标定

双目 PTZ 相机标定的目的是建立两相机之间的坐标对应, 进而方便实现协同跟踪等应用。经典的对极几何理论满足相机间点对线的约束, 无法直接解决本文问题。考虑到中远距离场景中, 运动目标往往约束在某一参考平面上, 而该平面上的点 X 在两相机的成像点 (x, y) 和 (x', y') 满足单应约束关系 H , 如公式(9)所示。因此文中利用单应约束建立两相机之间的坐标关联, 并提出了一种基于目标运动信息的单应矩阵自标定方法。

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{cases} x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + 1} \\ y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + 1} \end{cases} \quad (9)$$

如图 1 所示, 文中在目标质心确定的平行于地平面的平面上建立参考平面。设两相机拍摄参考平面上的同一块区域, 令此时两相机参数分别为 (p_1, t_1, z_1) 和 (p_2, t_2, z_2) 。文中方法的具体步骤如下:

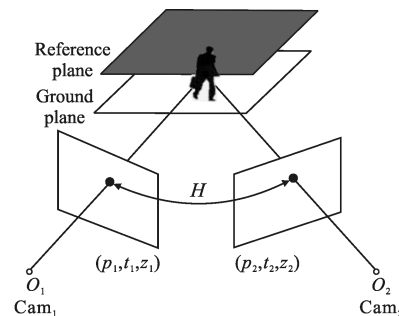


图 1 单应约束示意图

Fig.1 Illustration for the homography

(1) 目标轨迹获取。采集两相机同步视频,利用静止相机下的跟踪方法^[14]逐帧获取运动目标轨迹(两相机在固定参数 (p_1, t_1, z_1) 和 (p_2, t_2, z_2) 时,等价于静止相机),不同运动物体赋予不同的标号,并记录帧序号、目标数目及质心坐标。

(2) 样本帧筛选。用于训练的视频片段中可能有多个目标,为保证一一对应性,首先将目标数目不相等的同步帧图像滤除;其次,在两相机相对位置未知的情况下,只能通过组合方式获取目标对应关系,而随着目标数目的增加,错误对应点所占比例也逐渐增大(如某一同步帧图像中包含两个跟踪目标,通过组合方式可以获得 4 组对应点,而其中只有 2 组是正确的),因此文中将目标数目大于 3 的同步帧图像滤除,以减少错误对应点所占比例。

(3) 外点滤除。设 $\{x_i, y_i, t\}_{i=1}^N$ 和 $\{x'_i, y'_i, t\}_{i=1}^N$ 为经过上述步骤获得的对应点集,其中 (x, y) 为目标质心坐标, t 代表帧序号。点集中包含众多不符合单应约束的错误对应点,称之为外点 **Outliers**。文中采用 RANSAC 方法^[15]剔除外点,该方法对于错误超过 50% 的数据仍然能够处理,是一种鲁棒的外点去除方法。

(4) 单应矩阵估计。设通过步骤(3)获得的正确对应点数目为 M ,以该 M 组对应点的重投影误差最小建立目标函数:

$$E = \sum_{i=1}^M \left(\left(x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + 1} \right)^2 + \left(y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + 1} \right)^2 \right) \quad (10)$$

利用 LM 方法求解此非线性优化问题,得到单应矩阵的最优解。

1.3 系统所具有的两个重要性质

通过上述标定,文中系统具有如下性质。

性质 1: 对于通过 1.1 节标定的单目 PTZ 相机,给定参数 (p, t, z) 及图像平面或其延伸平面上的某一点 x ,可计算对应的 pan-tilt 参数 (p', t') ,使得当相机运动到该参数时,点 x 位于图像主点位置。

如图 2 所示, (c_x, c_y) 和 $f(z)$ 分别为相机主点坐标和等效焦距。设 x 点坐标为 (x, y) ,则由针孔相机模型,可计算 pan-tilt 参数需改变的绝对夹角 Δp 和 Δt ,并由点 x 在图像平面上的位置确定最终的 pan-tilt 参数 (p', t') :

$$\begin{cases} \Delta p = \arctan \frac{|x - c_x|}{f(z)} \\ \Delta t = \arctan \frac{|y - c_y|}{f(z)} \end{cases} \quad (11)$$

$$\begin{cases} p' = p + \Delta p, t' = t - \Delta t & \text{if } x \geq c_x \wedge y \geq c_y \\ p' = p + \Delta p, t' = t + \Delta t & \text{if } x \geq c_x \wedge y < c_y \\ p' = p - \Delta p, t' = t - \Delta t & \text{if } x < c_x \wedge y \geq c_y \\ p' = p - \Delta p, t' = t + \Delta t & \text{otherwise} \end{cases} \quad (12)$$

对于 zoom 参数 z' ,可根据具体应用赋予相应值。

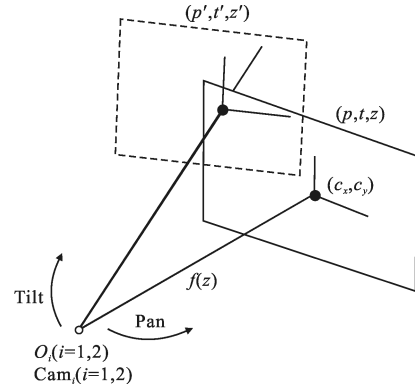


图 2 系统性质 1 示意图

Fig.2 Illustration for the first property of proposed system

性质 2: 对于通过 1.1 节和 1.2 节标定的双目 PTZ 相机,两相机在任意 pan-tilt-zoom 参数下,通过两个单应矩阵建立坐标关联。

以相机一(Cam₁)举例说明。如图 3 所示,Cam₁ 在参数 (p, t, z) 和 (p_1, t_1, z_1) 下的成像平面分别记为 I 和 I_1 ,

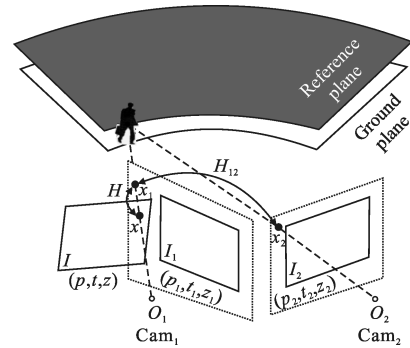


图 3 系统性质 2 示意图

Fig.3 Illustration for the second property of proposed system

Cam₂ 在参数 (p_2, t_2, z_2) 下的成像平面记为 I_2 ; (p_1, t_1, z_1) 和 (p_2, t_2, z_2) 为 1.2 节中估计参考平面诱导的单应矩阵时的两相机对应参数,设该单应矩阵的估计值为 H_{12} 。若已知参考平面上的某点在 Cam₁ 像平面上 I 的投影 x ,根据单目 PTZ 相机模型,可计算该点在 Cam₁ 像平面 I_1 或其延伸平面上的对应点 x_1 :

$$x_1 = \kappa \mathbf{K}(z_1) \mathbf{R}(p_1, t_1) \mathbf{R}(p, t)^{-1} \mathbf{K}(z)^{-1} x = \mathbf{H}x \quad (13)$$

式中,单应矩阵 \mathbf{H} 仅与参数 (p, t, z) 和 (p_1, t_1, z_1) 有关。

给定 x_1 , 根据 1.2 节单应矩阵的标定结果 H_{12} , 可计算该点在 Cam_2 像平面 I_2 或其延伸平面上的对应点 x_2 :

$$x_2 = H_{12}x_1 \quad (14)$$

由多视图几何理论, 在两相机安装固定及参考平面选定后, H_{12} 仅与参数 (p_1, t_1, z_1) 和 (p_2, t_2, z_2) 有关^[16]。因此, 综合等式(13)和(14)可知, Cam_1 在任意参数下均可通过两个单应矩阵与 Cam_2 建立坐标关联, 且两单应矩阵只与相机参数有关, 而与图像内容无关。

2 在线协同跟踪

基于上述标定过程, 文中设计了分段静止的协同跟踪策略, 以实现目标的分辨率视觉关注, 流程图如图 4 所示。

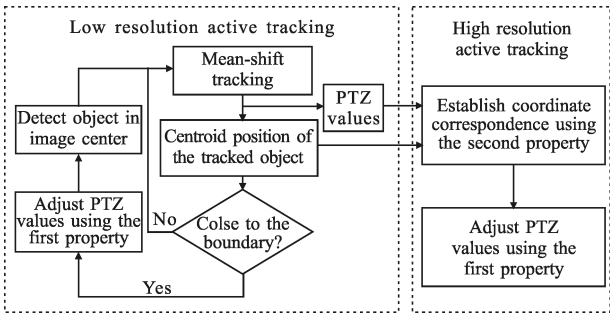


图 4 在线协同跟踪流程图

Fig.4 Flowchart of online cooperative tracking

低分辨率相机视场较大, 因此可以在目标距离图像边界较远时保持静止状态, 并利用 Mean-shift 方法^[17]进行跟踪; 在目标接近图像边界时利用系统性质 1 对参数进行快速调整, 使目标处于图像主点位置(近似为图像中心), 并在图像主点的某一局部区域内通过双向差分^[18]运算重新定位目标; 参数调整结束后, 相机重新处于静止状态, 继续利用 Mean-shift 方法对目标进行跟踪。

高分辨率相机视场较小, 需要在每一时刻进行参数调整, 以保持目标处于图像中心位置。由于在每一时刻可获得低分辨率相机参数及目标质心位置, 因此根据系统性质 2 可建立两相机之间的坐标关联, 并由系统性质 1 计算高分辨率相机参数, 使目标处于图像中心位置。

3 实验结果

文中实验所用双目 PTZ 视觉系统由两个 SONY

EVI D70 摄像机构成, 该相机的 pan 参数变化范围为 $-170^\circ \sim 170^\circ$, tilt 参数变化范围为 $-30^\circ \sim 90^\circ$, zoom 参数变化范围为 $0 \sim 18$, 图像分辨率为 320×240 。两相机通过视频线连接到一台 PC 机的视频采集卡上, 并通过 RS232 串口总线与摄像机交互, 包括参数获取和动作控制。为了验证方法的有效性, 在室外监控场景中进行多组实验, 实验结果可从 <http://ivg.au.tsinghua.edu.cn/uploads/Main/Data/results.zip> 下载。算法实现基于 Windows 操作系统, 采用 Visual Studio2008 和 Open CV2.1 作为软件平台。

3.1 离线标定实验

利用文中的单目相机标定方法, 对两个 PTZ 相机分别进行了标定。对于主点估计, 每个相机在特征点较丰富的场景, 分别获取 19 幅图像进行估计, 实验测得的两相机的主点坐标分别为 $(151.63, 126.67)$ 和 $(155.82, 122.79)$ 。对于焦距随 zoom 参数的模型估计, 每个相机在固定 zoom 参数下, 分别获取 2 幅图像进行估计(每个相机共获取 38 幅图像), 并通过模型拟合方式获得最终的模型系数。为了保证匹配所需要的特征点数目, 固定 zoom 参数下获取的两幅图像至少保持 1/2 重叠视场。图 5 给出了两相机在离散

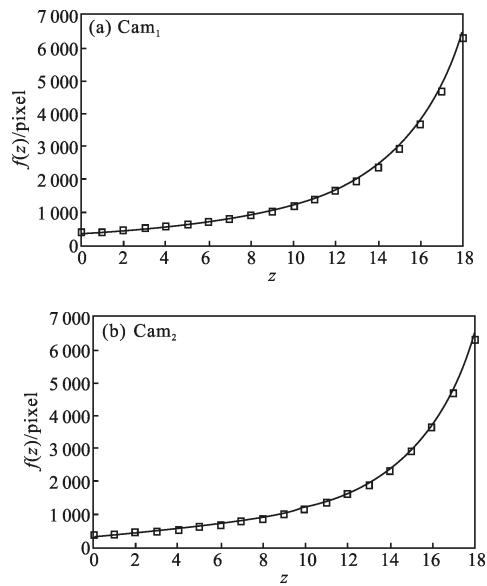


图 5 焦距随 zoom 参数变化模型的标定结果

Fig.5 Calibration result of focal length as a function of zoom parameter

zoom 下的焦距估计值以及模型拟合结果。从实验结果可以看出各离散估计值与模型很接近, 表明模型估计的准确性。最终估计的两相机焦距随 zoom 参数

的变化关系为:

$$\begin{aligned} \text{Cam}_1: f(z) &= 345.30 \times e^{0.116 \times z} + 1.78 \times e^{0.424 \times z} \\ \text{Cam}_2: f(z) &= 346.80 \times e^{0.113 \times z} + 1.69 \times e^{0.423 \times z} \end{aligned} \quad (15)$$

如公式(13)所示,单目相机在不同参数拍摄图像上的对应点满足单应矩阵关系,且该单应矩阵与相机内参标定结果相关,因此可以采用图像配准的方法评测标定结果的准确性。文中采用两幅图像配准重合区域对应像素灰度的平均绝对差异 (Mean Absolute Difference, MAD) 作为评价准则, MAD 值越小,表明两幅图像之间的单应矩阵估计的越准确,即内参标定的越准确。表 1 为文中方法与参考文献[10]方法的实验对比,其中每个相机各 5 组实验。从实验结果可以看出,由于文中方法在各离散 zoom 下并行计算焦距值,保证了特征匹配的稳定性,避免了误差累积,进而提高了标定的准确性。图 6 为其中一组实验的配准结果,图 6(a)和 6(b)两幅图像的 pan-tilt-zoom 参数分别为 (-74.84, -12.67, 9.80) 和 (-71.76, -14.77, 13.00), 图 6(c)为参考文献[10]配准结果,图 6(d)为文中方法配准结果。

表 1 文中标定方法与其他文献的性能对比

Tab.1 Comparison between the proposed and literature method

Method	Cam ₁					Cam ₂				
	1	2	3	4	5	1	2	3	4	5
Ref. [10]	13.81	8.92	9.43	10.67	10.35	8.83	7.94	9.21	8.02	9.32
Proposed	11.70	7.83	7.67	8.89	9.03	6.69	6.48	7.61	7.17	7.78

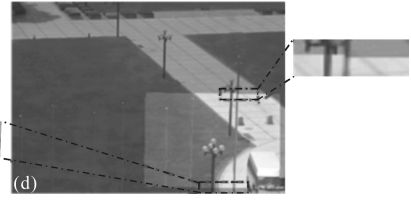
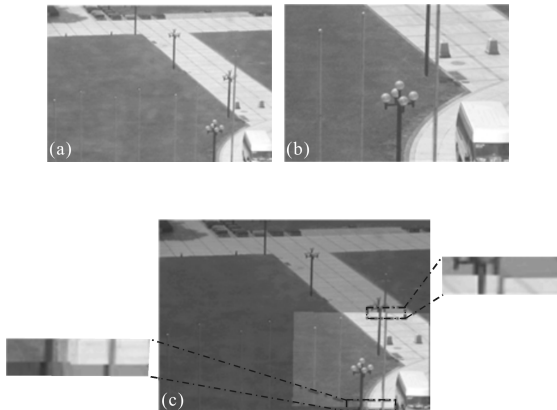


图 6 配准比较

Fig.6 Registration comparison

对于两相机共面约束自标定,文中设置两相机参数分别为 (-71.54, 14.25, 12.00) 和 (-73.67, -11.85, 12.00), 即对应于 1.2 节和 1.3 节中的 (p₁, t₁, z₁) 和 (p₂, t₂, z₂)。捕获两相机同步视频,利用 1.2 节方法估计单应矩阵 H₁₂。图 7 给出了训练视频中其中一帧的跟踪结果,图 7(a)和图 7(b)分别为 Cam₁ 和 Cam₂ 的捕获图像。矩阵 H₁₂ 的最终标定结果为:

$$H_{12} = \begin{bmatrix} 0.960 0 & 0.036 8 & 24.998 8 \\ -0.008 5 & 1.004 4 & 11.048 6 \\ -0.000 1 & -0.000 1 & 1 \end{bmatrix} \quad (16)$$

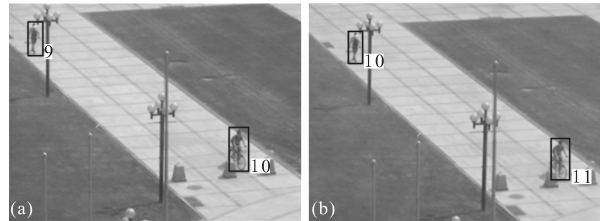


图 7 训练视频中其中一帧同步图像对

Fig.7 A synchronized image pair of training video

由于 H₁₂ 的真实值难以获知,文中采用间接方式评测标定结果的精度:保持两相机参数不变,重新获取同步视频片段,对每一帧图像手动获取目标质心坐标及对应关系,并计算这些对应点通过矩阵 H₁₂ 变换前后坐标值的平均欧式距离,统计得到平均误差为 5.2425,即水平方向和垂直方向误差平均不超过 4 个像素。文中认为该标定结果满足要求。

3.2 在线协同跟踪实验

图 8、图 9 给出了协同跟踪的两组实验结果,对应两相机分别作为低分辨率相机的情形。实验中,低分辨率相机 pan-tilt 参数改变了三次,高分辨率相机 zoom 参数设置为最大值 18。由实验结果可以看出,文中方法可以在大视场下有效地主动跟踪目标,并获得目标的多分辨率图像。

图 10 以图 8 跟踪结果为例,直观解释了协同跟

踪的实现过程。图 10(a)为低分辨率相机 Cam_1 在三个静止阶段的目标运动轨迹,在每一阶段,当目标即将离开其观测视场时(文中设置为目标质心距离图像边界小于 50 个像素),利用性质 1 快速调整 pan-tilt 参数,该参数调整过程很快,因此参数调整后,可基本保持目标位于图像的中心位置;图 10(b)为利

用性质 2 建立两相机坐标关联的实验结果,即依次将目标定位到 Cam_1 参数 (p_1, t_1, z_1) 像平面或其延伸平面上,以及 Cam_2 参数 (p_2, t_2, z_2) 像平面或其延伸平面上, (p_1, t_1, z_1) 与 (p_2, t_2, z_2) 的具体取值见 3.1 节;最后,在每一时刻 Cam_2 利用性质 1 计算自身参数,从而使目标处于图像中心位置,如图 10(c)所示。



图 8 协同跟踪实验结果(Cam_1 为低分辨率相机)

Fig.8 Cooperative active tracking results(Cam_1 acting as the low resolution camera)

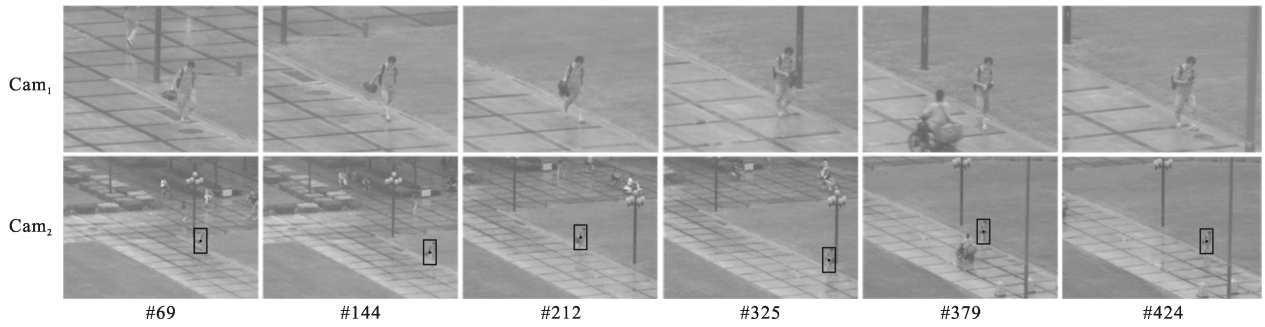


图 9 协同跟踪实验结果(Cam_2 为低分辨率相机)

Fig.9 Cooperative active tracking results(Cam_2 acting as the low resolution camera)

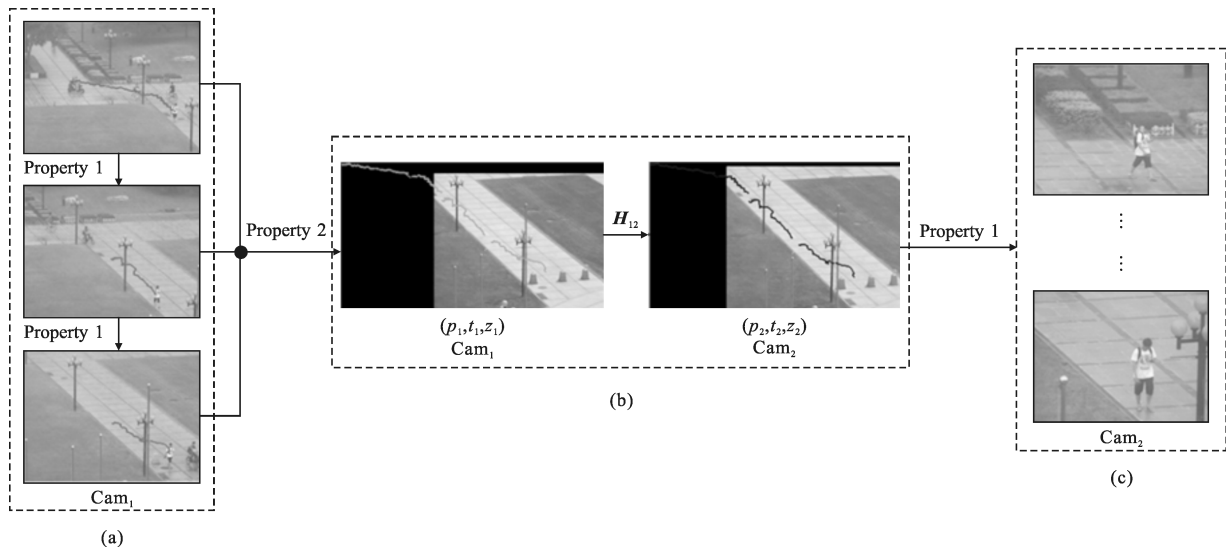


图 10 协同跟踪实现过程的解释

Fig.10 Illustration for the procedure of cooperative active tracking

4 结 论

文中研究了一种基于双目 PTZ 相机的多分辨率主动跟踪方法,该方法可以实现两相机在任意视角和任意尺度下的协同跟踪,获得目标的多分辨率信息,有利于后续的行为分析和场景理解等应用。文中方法的实现包括离线标定和在线协同跟踪两部分。离线标定部分,提出了一种改进的基于特征匹配的单目 PTZ 自标定和基于目标运动信息的双目 PTZ 共面约束自标定方法,并总结了通过此标定系统具有的两个性质;基于上述标定方法,采用分段静止的策略实现目标的协同跟踪。文中最后通过实际监控场景的实验验证了方法的有效性,结果表明文中方法可以有效地主动跟踪目标,为解决主动跟踪中监控视场和目标分辨率之间的矛盾提供了一条有价值的途径。

参考文献:

- [1] Zhao Qian, Yuan Jianquan, Lu Xinping, et al. Moving object tracking algorithm based on object estimation and mean shift theory [J]. *Infrared and Laser Engineering*, 2010, 39(6): 1152–1156. (in Chinese)
赵倩, 袁健全, 鲁新平, 等. 结合目标预估计与 Mean Shift 理论的运动目标跟踪算法[J]. *红外与激光工程*, 2010, 39(6): 1152–1156.
- [2] Meng Gang, Jiang Zhiguo, Zhao Danpei. Histogram of oriented gradients and submanifold applications in object tracking [J]. *Infrared and Laser Engineering*, 2012, 41(6): 1664–1668. (in Chinese)
孟钢, 姜志国, 赵丹培. 梯度方向直方图和子流形在目标跟踪中的应用[J]. *红外与激光工程*, 2012, 41(6): 1664–1668.
- [3] Hu W M, Tan T N, Wang L, et al. A survey on visual surveillance of object motion and behaviors [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 2004, 34(3): 334–352.
- [4] Yilmaz A, Javed O, Shah M. Object tracking: a survey [J]. *ACM Computing Surveys*, 2006, 38(4): 229–240.
- [5] Xue K, Liu Y, Ogunmakin G, et al. Panoramic gaussian mixture model and large-scale range background subtraction method for PTZ camera-based surveillance systems [J]. *Machine Vision and Application*, 2012, 11(4): 1–16.
- [6] Kumar P, Dick A, Sheng T S. Real time target tracking with pan tilt zoom camera [C]//*Digital Image Computing: Techniques and Applications*, 2009: 492–497.
- [7] Varcheie P D Z, Bilodeau G A. People tracking using a network-based PTZ camera[J]. *Machine Vision and Application*, 2011, 22(4): 671–690.
- [8] Bernardin K, Camp F V D, Stiefelhagen R. Automatic person detection and tracking using fuzzy controlled active cameras [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1–8.
- [9] Zhang Z Y. A flexible new technique for camera calibration [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 11(11): 1330–1334.
- [10] Sinha S, Pollefeys M. Towards calibrating a pan-tilt-zoom cameras network [C]//*European Conference on Computer Vision Workshop*, 2004: 1–11.
- [11] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91–110.
- [12] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. *Computer Vision and Image Understanding*, 2008, 110(3): 346–359.
- [13] Levenberg K. A method for the solution of certain nonlinear problems in least squares [J]. *Quarterly of Applied Mathematics*, 1944, 2(2): 164–168.
- [14] Stauffer C, Grimson W. Learning patterns of activity using real-time tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 747–757.
- [15] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. *Communications of the ACM*, 1981, 24(6): 381–395.
- [16] Stein G, Romano R, Lee L. Monitoring activities from multiple video streams: establishing a common coordinate frame [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 758–767.
- [17] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2000: 142–149.
- [18] Dubuisson M P, Jain A K. Contour extraction of moving objects in complex outdoor scenes [J]. *International Journal of Computer Vision*, 1995, 14(1): 83–105.