

双 PTZ 摄像机系统的标定

万定锐 周杰

(清华大学自动化系,北京 100084)

摘要 在视频监控领域,PTZ(Pan/Tilt/Zoom)摄像机因为其具有可变视角和可变焦能力,已经越来越多的被采用,而且用两个 PTZ 摄像机组成一个视觉系统具有可以同时获取全景图像、近景高分辨率图像以及场景深度信息等优势,虽然目前关于双 PTZ 视觉系统的研究还比较少,但其应用潜力很大。对双 PTZ 摄像机系统进行标定的目的之一是辅助两个镜头之间的配准,进而方便立体匹配等。两个摄像机之间公共坐标系的选择是标定问题种的一个关键。针对这一问题提出了一种球面校正公共坐标系,并研究了其标定问题。众所周知,传统的多目镜头标定主要针对静止镜头,而且一般是通过事先估计摄像机的架设位置或者利用标定物来计算摄像机之间的关系,但此类标定方法无法直接解决上述问题。为了更好地进行标定,在此提出了一种利用多组基础矩阵的方法进行自动估计的标定方法,该方法不需要标定物,而只与镜头架设方式有关,且不受摄像机 PTZ 参数改变影响。实验结果表明,该方法具有方便和实用性。

关键词 PTZ 摄像机 标定 极线几何 基础矩阵 立体校正

中图法分类号:TP242.62 文献标识码:A 文章编号:1006-8961(2008)04-0786-08

Calibration of Dual-PTZ-camera System

WAN Ding-ruì, ZHOU Jie

(Department of Automation, Tsinghua University, Beijing 100084)

Abstract PTZ (Pan/Tilt/Zoom) cameras have been widely used in visual surveillance domain because of its capability of changing both view-angle and image resolution. Dual-PTZ-camera system is one of the simplest systems that multi-resolution information and stereo information can be obtained synchronously. Although there are few related researches about dual-PTZ-camera system in the state of the art, we believe it has wide application prospects. Building the common coordinate system is an important problem in dual-PTZ-camera calibration, which can be used for image alignment and stereo matching. From this point, we adopt the longitude-latitude coordinate system and use a multi-fundamental matrix approach to deal with the calibration problem. Traditional multi-camera calibration approaches mainly deal with the static cameras, and which always take advantage of the known camera locations or use calibration tools. The proposed calibration is independent to specific PTZ parameters, which is very convenient in application. The experimental results show the practicality of our approach.

Keywords PTZ camera, calibration, epipolar geometry, fundamental matrix, stereo rectification

1 引言

计算机视觉研究开始于 20 世纪 60 年代初,发展至今,它已经成为一个涉及到计算机科学和工程、信号处理、物理学、应用数学和统计学、神经生理学

和认知科学等众多领域的一门综合性的学科,而且它被认为是工程领域,也是科学领域中的一个富有挑战性的研究领域^[1],而立体视觉研究的重要性更是不言而喻^[2]。

随着计算机软硬件技术的不断发展,视觉监控系统的应用趋势主要由单目到多目发展,并且由静

基金项目:国家自然科学基金项目(60673106,60573062 和 60721003)

收稿日期:2007-07-18;改回日期:2008-01-02

第一作者简介:万定锐(1981 ~),男,2004 年获取清华大学学士学位,现为清华大学自动化系博士研究生(硕士连续)。研究方向为计算机视觉、视频监控。E-mail:wandingrui00@mails.tsinghua.edu.cn

止镜头向主动镜头过渡。本文提出的双目 PTZ (Pan/Tilt/Zoom) 视觉系统就是其中一种最简单的配置,由于它可以同时获取多视角、多分辨率图像信息以及立体信息等,因此相对于传统的双目静止镜头视觉系统,它具备了更多的优势。在视频监控领域,PTZ 摄像机因为其具有可变视角和可变焦能力,已经越来越多地被采用。用两个 PTZ 摄像机组成一个视觉系统具有可以同时获取全景图像、近景高分辨率图像以及场景深度信息等优势。

从仿生学的角度看,该系统类似于变色龙的眼睛系统,既具有可以同时独立进行大范围巡视,又能协同实现立体视觉的优势。众所周知,变色龙的两只眼睛可以独立运动,各自负责 180° 左右的空间视场,互不牵制,当发现猎物或者敌害时,两眼则会自动会聚视场,获取立体信息,以确定食物或者敌害的准确距离,从而方便捕食或者逃逸。由此可见该系统的优越性。

研究双目 PTZ 摄像机系统需要的一个核心技术就是两个镜头之间的信息融合,而其中最基本的就是建立一个公共坐标系,本文称之为双目 PTZ 摄像机系统的标定。

2 系统设计

双目 PTZ 摄像机系统的标定主要包含以下两个部分:单目 PTZ 摄像机标定和双目 PTZ 摄像机公共坐标系的建立(如图 1 所示)。

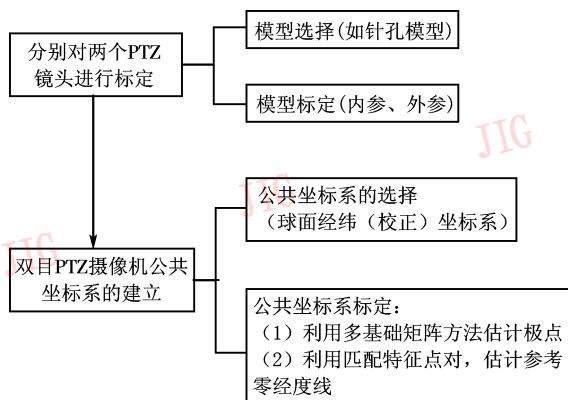


图1 双目 PTZ 摄像机系统标定流程图

Fig. 1 Flow-chart of calibration of dual-PTZ-camera system

关于单目 PTZ 摄像机标定,首先需要建立当 PTZ 参数固定时的摄像机模型(本文选用的是针孔模型);然后需要考虑 PTZ 参数变化对模型的影响。这部分内容将在第 3 节详细介绍。

对于双目 PTZ 摄像机公共坐标系,传统的多目镜头标定主要是通过一个世界坐标(一般采用笛卡尔坐标系)来分别估计每个摄像机在该坐标系下的相对位置关系,以建立单个摄像机坐标系到公共坐标系之间的联系,这种标定方法目前已经非常成熟。考虑到双目 PTZ 摄像机的特殊性,本文采用了球面经纬坐标系。相对传统笛卡尔公共坐标系而言,本文提出的系统有如下优势:首先,球面坐标系可以方便地处理镜头 PT 参数变化的情形;其次,由于这种坐标系的选择可以方便地辅助立体校正,进而可在很大程度上提高立体视觉的效率,因此这种坐标系也可称为球面校正坐标系。本文将在第 4、5 节详细介绍其标定方法。

传统的多目镜头标定主要针对静止镜头,而且一般要通过事先估计摄像机的架设位置或者利用标定物来计算摄像机之间的关系,但此类标定方法无法直接解决本文提出的问题。本文采用经纬坐标系作为公共坐标系,并采用多组基础矩阵的方法进行自动估计,其不需要标定物。由于本文提出的标定方法不受摄像头架设朝向影响,只有当镜头架设方式改变时,才需要重新进行标定,而且不受摄像机 PTZ 参数改变的影响,从而提高了系统的方便性。实验结果验证了本文方法的实用性。

3 单目 PTZ 摄像机标定

在使用 PTZ 镜头的视觉系统中,对 PTZ 镜头进行标定是一件很有意义的事情,特别对于与 PTZ 镜头相关的视频图像处理有很大的帮助。文献[3]综述了几种典型的传统摄像机标定方法,其中包括使用标定物和 LED(light emitting diode)进行标定的方法,以及通过物理连接的方式获取关于摄像机的参数,或者通过其他传感器获得摄像机位置信息等的标定方法。本文系统采用了基于特征的标定方法。和传统方法相比,基于 PTZ 镜头本身的特点,本文的方法还结合了直接从主动镜头获取的 PTZ 参数信息,相对于一般主动镜头的标定,本文所采用的方法更简单实用。为简单起见,本系统采用简化的摄像机模型,即不考虑焦距变化以及图像畸变。

3.1 摄像机模型

本文系统中使用的 PTZ 镜头的型号为 SONY EVI D70 系列,并采用如下摄像机模型:

$$\tilde{\mathbf{x}} = \kappa \mathbf{K} \mathbf{R} \mathbf{X}, \mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

其中, \mathbf{x} 和 \mathbf{X} 分别为点的图像坐标和世界坐标; 符号“ \sim ”表示广义坐标, 如 $\tilde{\mathbf{x}} = [\mathbf{x}^T \ 1]^T$ 。 f 为等效焦距, (u_0, v_0) 为图像主点 (principal point)。为简化模型, 可进行以下假设:

- (1) 摄像机的旋转中心保持固定不变;
- (2) 镜头像素的纵横比 (aspect ratio) 设为 1, 倾斜度 (skew) 设为 0;

(3) 主点 (u_0, v_0) 用 Zoom 中心近似代替^[4,5]。所谓 Zoom 中心是指当镜头的 Zoom 变化时, 视场中每一个给定点在图像中成像的位置处于某一条过某个固定点的直线上, 该固定点即为 Zoom 中心。Zoom 中心也称为膨胀中心 (center of expansion)^[4];

- (4) 不考虑镜头畸变。

3.2 模型的标定

3.2.1 Zoom 中心的估计

关于 Zoom 中心的估计, 本文是采用基于图像数据的估计方法^[5], 即首先保持摄像机的 Pan 和 Tilt 参数不变, 只改变 Zoom 参数, 例如从 z_{\min} 到 z_{\max} 变化; 然后采集一个图像序列, 再对其中的某些特征点进行跟踪; 最后用最小二乘估计即可得到 Zoom 中心的估计。

实验中, 两个镜头测得的 Zoom 中心分别为 $(150.1, 127.2)$ 和 $(159.4, 123.3)$, 其中图像的大小为 320×240 。实验结果见图 2。

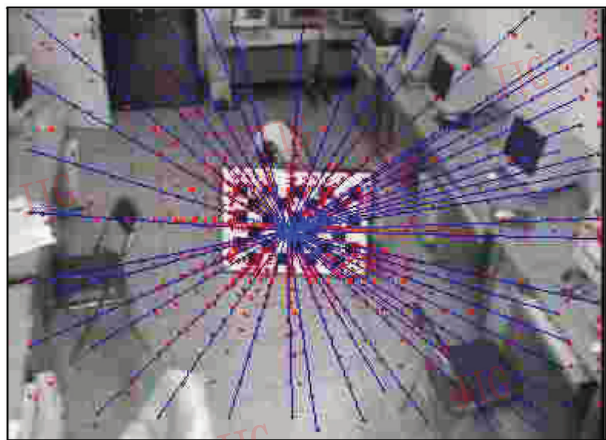


图 2 Zoom 中心估计的实验示意图

Fig. 2 Zoom center estimation

如果首先将图像进行平移, 使得图像坐标原点

与主点 (Zoom 中心) 重合, 那么内参矩阵 \mathbf{K} 可以写成对角阵的形式, 并且只包含一个待定参数, 该参数的标定会在下一小节说明。

3.2.2 内参矩阵 \mathbf{K} 的估计

由上所述, 内参矩阵 $\mathbf{K} = \text{diag}\{k_z, k_z, 1\}$ 仅有一个自由度 k_z , 且该变量仅与当前 Zoom 值有关。本文采用类似文献[3]的方法, 即在几个离散的 Zoom 值下, 首先估计 k_z , 然后利用这些离散采样值构造一个模型来近似表达 $k_z = k(z)$, 本实验是选择下面的模型:

$$k(z) = p_1 e^{p_2 z} + p_3 e^{p_4 z} \quad (2)$$

其中, p_1, p_2, p_3, p_4 为待定参数。实验发现, 该模型可以拟合得很好, 图 3 显示了拟合结果。其中参数 $p_1 = 374.8$, $p_2 = 0.08216$, $p_3 = 7.641$, $p_4 = 0.3924$ 。

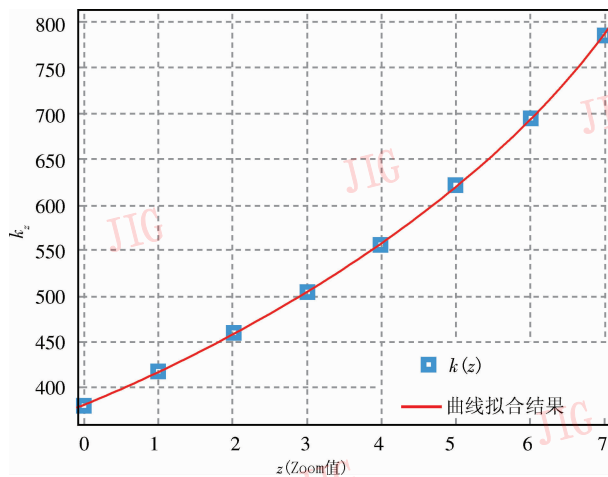


图 3 $k_z = k(z)$ 的拟合曲线

Fig. 3 Curve fitting result for $k_z = k(z)$

3.2.3 旋转矩阵 \mathbf{R} 的计算

旋转矩阵 \mathbf{R} 可以直接通过已知的 Pan 和 Tilt 参数进行计算。计算公式为

$$\mathbf{R} = \begin{bmatrix} \cos\theta_p & 0 & \sin\theta_p \\ -\sin\theta_p \sin\theta_T & \cos\theta_T & \cos\theta_p \sin\theta_T \\ -\sin\theta_p \cos\theta_p & -\sin\theta_T & \cos\theta_p \cos\theta_T \end{bmatrix} \quad (3)$$

其中, θ_p 和 θ_T 分别为对应的 Pan 和 Tilt 参数值。需要说明的是, 该公式包含以下两个重要的假设: (1) 旋转中心和光心重合; (2) Pan 和 Tilt 的旋转轴正交。对于比较成熟的商用一体 PTZ 镜头, 这两个假设都基本满足。对于更一般的情形, 即两个假设都不成立, 文献[6]在这一方面做了比较深入的研究, 这里不再赘述。

4 建立双目 PTZ 摄像机校正坐标系

4.1 建立校正坐标系的目的

当已经建立好单个 PTZ 摄像机坐标系后, 如果没有两个坐标系之间的关系, 则很难对其各自的信息进行交互。举一个具体的例子说明, 如立体匹配, 如果已经在某一个镜头图像中确定了一点, 那么为了获取立体信息, 就需要在另一个镜头的图像中找到其对应点, 如果两个镜头坐标系之间的关系未知, 那么只能在图像中进行 2 维搜索, 其计算量是非常大的。如果建立一种公共坐标系, 那么就有可能在很大程度上降低搜索量, 例如由 2 维降低到 1 维。事实上, 传统的双目立体视觉中的立体校正就已实现了这样的目标, 即已将 2 维图像搜索降低到了 1 维的直线搜索。因此也可称这样的公共坐标系为(立体)校正坐标系。

该坐标系不光是对立体视觉有用, 如果系统是一个主从式系统, 人们总是希望通过其中一个 PTZ 镜头的图像中某个目标的位置来指导另一个镜头动作到某个包含该目标的位置, 例如使目标位于另一个镜头的图像中心, 那么也需要建立这样的公共坐标系。

4.2 选择合适的校正坐标系

前面已经提到, 由于该校正坐标系主要是为立体校正服务, 因此该坐标系的选择应尽可能有利于立体校正的实现。下面简单介绍立体校正的原理。

传统的立体校正就是希望通过图像变换使得两幅图像中对应物体的像处于同一条水平扫描线上^[7,8]。其主要采用的方法就是利用经典的极线几何约束来将极线映射到校正图像的水平扫描线上。由于本文采用的是 PTZ 镜头, 如果采用传统的立体校正方法, 那么当镜头 PTZ 参数发生变化时, 就需要重新计算所有极线, 并重新映射, 因此本文提出了球面立体校正的方法^[9]。其主要思想就是先对每个 PTZ 镜头建立单位球面坐标系, 然后将每一对图像对分别映射到各自球面坐标系中, 最后对相应的球面区域进行校正。这种方法的重大优势就在于统一了各种 PTZ 参数情况下的校正方法, 即只要两个镜头安装固定, 即使 PTZ 参数发生变化, 也不需要重新计算校正坐标系, 而且每次校正只需要计算少量的参数, 因此非常有利于实时的立体匹配等后续处理。

为方便起见, 本文选择经纬坐标系作为公共坐标系(如图 4 所示)。I₁, I₂ 为两个镜头的图像平面,

φ_1, φ_2 为与两个镜头对应的摄像机单位球面坐标系。如果采用经纬坐标系, 则关键是要确定极点(epipole)和参考零经度线(zero longitude)。公共坐标系要求两个坐标系的极点 $E_i, E'_i (i = 1, 2)$ 共线, 并且要求各自的参考零经度线 $\overline{E_1 m_1 E'_1}$ 和 $\overline{E_2 m_2 E'_2}$ 共面。选择这种坐标系后, 就可以直接将传统的极线几何扩展到球面:

- (1) 极平面 $\Pi PO_1 O_2$ 的定义不变;
- (2) 极点定义为基线 $\overline{O_1 O_2}$ 与单位球面的交点, 且正好与经纬坐标系的极点 $E_i, E'_i (i = 1, 2)$ 重合;
- (3) 极线的定义修改为极线圆 c_1, c_2 , 它是单位球面与极平面的交线。

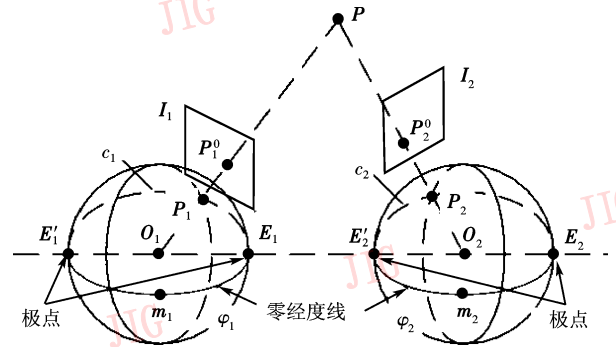


图 4 公共经纬坐标系

Fig. 4 Common longitude-latitude coordinates

立体校正就是把相同经度的经线映射到校正平面的水平扫描线上, 具体的校正方法见文献[9]。下面详细介绍校正坐标系的建立。

本文采用多组图像对来估计坐标系参数, 要求每组图像对的视场应尽可能接近, 这样就可以用传统的 RANSAC^[11] 方法来估计对应的基础矩阵。设第 j 组图像对, 校正时, 首先分别对两幅图像进行特征点检测(例如 Harris 角点), 然后用 RANSAC 方法估计其基础矩阵 F_j , 并记录匹配的特征点对 $\{x_{k,j}^{(1)}, x_{k,j}^{(2)}\}$ 。

4.2.1 利用多个基础矩阵估计极点

以第 j 组图像对为例, 设 $\{e_{j,1}, e_{j,2}\}$ 为图像中的极点对(传统极线几何中的定义), 则它们是摄像机坐标系下的目标极点对 $\{E_1, E_2\}$ 在图像中的投影, 因此, 由摄像机镜头模型知, $\tilde{e}_{j,i} = \kappa_{j,i} K_{j,i} R_{j,i} E_i$, ($i = 1, 2$), 其中符号“~”表示广义坐标, $\kappa_{j,i}$ 为比例因子。再由极线理论知, $F_j \tilde{e}_{j,1} = \tilde{e}_{j,2}^T F_j = 0$ 。两个镜头的基础矩阵定义为

$$\begin{cases} \mathbf{F}_j^{(1)} = \mathbf{F}_j \mathbf{K}_{j,1} \mathbf{R}_{j,1} \\ \mathbf{F}_j^{(2)} = \mathbf{R}_{j,2}^T \mathbf{K}_{j,2}^T \mathbf{F}_j \end{cases} \quad (4)$$

因此有 $\mathbf{F}_j^{(1)} \mathbf{E}_1 = \mathbf{E}_2^T \mathbf{F}_j^{(2)} = 0$ 。

考虑到鲁棒性,本文采用多组图像进行极点对 $\{\mathbf{E}_1, \mathbf{E}_2\}$ 的估计。定义

$$\begin{cases} \mathbf{A}_1 = [(\mathbf{F}_1^{(1)})^T, (\mathbf{F}_2^{(1)})^T, \dots, (\mathbf{F}_n^{(1)})^T]^T \\ \mathbf{A}_2 = [\mathbf{F}_1^{(2)}, \mathbf{F}_2^{(2)}, \dots, \mathbf{F}_n^{(2)}] \end{cases} \quad (5)$$

然后利用 SVD 的方法,通过求解 $\mathbf{A}_1 \mathbf{E}_1 = 0$ 和 $\mathbf{E}_2^T \mathbf{A}_2 = 0$ 即可得到 $\{\mathbf{E}_1, \mathbf{E}_2\}$, $\{\mathbf{E}'_1, \mathbf{E}'_2\}$ 即为各自关于原点的对称点。

4.2.2 利用匹配特征点对估计参考零经度线

以第 j 组图像对为例,设 $\{\mathbf{x}_{k,j}^{(1)}, \mathbf{x}_{k,j}^{(2)}\}$ 为匹配的特征点对,其中 $k = 1, 2, \dots, N_j$ (N_j 为第 j 组图像对中检测到的总匹配特征点对数目),则由摄像机镜头模型就可以得到对应的摄像机坐标系下的 $\{\mathbf{X}_{k,j}^{(1)}, \mathbf{X}_{k,j}^{(2)}\}$ 。当极点已经确定后,则需在两个单位球面上任意分别选定一条等经度线作为参考零经度线(实际上为两条半圆弧),并设 \mathbf{m}_1 和 \mathbf{m}_2 为两条半圆弧的中点,然后就可在由 $\{\mathbf{E}_i, \mathbf{E}'_i, \mathbf{m}_i\}$ 确定的经纬坐标系下,分别将 $\{\mathbf{X}_{k,j}^{(1)}, \mathbf{X}_{k,j}^{(2)}\}$ 转化为经纬坐标,并且只考虑第 s 个经度分量 $\{\alpha_s^{(1)}, \alpha_s^{(2)}\}$,其中 $s = 1, 2, \dots, N$ (N 为所有图像对中检测到的总匹配特征点对数目)。剩下的就是一个简单的优化问题,即通过调整 \mathbf{m}_2 (实际是绕极轴旋转的某个角度 $\Delta\alpha$),使得某个指标尽可能小。这个指标可以定义为平均经度绝对偏差量 $\frac{1}{N} \sum_{s=1}^N |\alpha_s^{(1)} - \alpha_s^{(2)}|$ 。

5 球面的极线约束

事实上,假设两个球面 φ_1 和 φ_2 分别对应两个 PTZ 镜头的摄像机坐标系,并且 $\{X_s^{(1)}, X_s^{(2)}; s = 1, 2, \dots, N\}$ 为两个球面上对应的特征点。可以证明,一定存在这样的矩阵 \mathbf{M} (类似基础矩阵),使得 $(X_s^{(2)})^T \mathbf{M} X_s^{(1)} = 0$, 并且对于极点 $\{\mathbf{E}_1, \mathbf{E}_2\}$ 也有 $\mathbf{M} \mathbf{E}_1 = \mathbf{E}_2^T \mathbf{M} = 0$ 。简单证明如下:

假设两个摄像机固定 (PTZ 参数可变),那么必定存在旋转矩阵 \mathbf{R} 和平移向量 \mathbf{t} 使得两个摄像机坐标系坐标 \mathbf{X}_1 和 \mathbf{X}_2 之间满足关系 $\mathbf{X}_2 = \mathbf{R} \mathbf{X}_1 + \mathbf{t}$ 。令全局世界坐标与第 1 个摄像机坐标系重合,即 $\mathbf{X}_w = \mathbf{X}_1$, 并设当前两个摄像机模型参数分别为 $\mathbf{K}_1, \mathbf{R}_1$ 和 $\mathbf{K}_2, \mathbf{R}_2$ (实际上可以按照镜头模型标定的方法,通过

PTZ 参数直接获取); 设 $\mathbf{u}_1, \mathbf{u}_2$ 为两幅图像中一组对应点,相应各自的摄像机坐标为 \mathbf{X}_1 和 \mathbf{X}_2 , 全局世界坐标为 $\mathbf{X}_w = \mathbf{X}_1$ 。

传统的基础矩阵 \mathbf{F} 由当前两个图像坐标和全局世界坐标之间的投影矩阵确定,并满足关系 $\tilde{\mathbf{u}}_2^T \mathbf{F} \tilde{\mathbf{u}}_1 = 0$, 如果可以由两个摄像机坐标系和全局世界坐标系之间的关系 (\mathbf{R}, \mathbf{t}) 确定新的矩阵 \mathbf{M} , 那么就可以避开 PTZ 参数对 \mathbf{M} 的影响,并且满足关系 $\mathbf{X}_2^T \mathbf{M} \mathbf{X}_1 = 0$ 。

设 \mathbf{t} 为一个 3 维向量,定义 $[\mathbf{t}]_{\times}$ 为由 \mathbf{t} 确定的反对称矩阵^[11]。根据 $[\mathbf{t}]_{\times}$ 的性质以及 $\mathbf{X}_2 = \mathbf{R} \mathbf{X}_1 + \mathbf{t}$ 可知, $[\mathbf{t}]_{\times} \mathbf{X}_2 = [\mathbf{t}]_{\times} \mathbf{R} \mathbf{X}_1$, 因此

$$\mathbf{X}_2^T \mathbf{M} \mathbf{X}_1 = 0, \quad \mathbf{M} = [\mathbf{t}]_{\times} \mathbf{R} \quad (6)$$

下面讨论 \mathbf{F} 和 \mathbf{M} 之间的关系: 由摄像机镜头模型可得

$$\begin{cases} \rho_1 \tilde{\mathbf{u}}_1 = \mathbf{K}_1 \mathbf{R}_1 \mathbf{X} \\ \rho_2 \tilde{\mathbf{u}}_2 = \mathbf{K}_2 \mathbf{R}_2 \mathbf{R} \mathbf{X} + \mathbf{K}_2 \mathbf{R}_2 \mathbf{t} \end{cases} \quad (7)$$

基础矩阵 \mathbf{F} 的定义为

$$\mathbf{F} = [\mathbf{K}_2 \mathbf{R}_2 \mathbf{t}]_{\times} (\mathbf{K}_2 \mathbf{R}_2 \mathbf{R}) (\mathbf{K}_1 \mathbf{R}_1)^{-1} \quad (8)$$

令 $\mathbf{A} = \mathbf{K}_2 \mathbf{R}_2 = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ (其中 $\mathbf{a}_i (i = 1, 2, 3)$ 为 \mathbf{A} 的列向量), 利用 $[\mathbf{t}]_{\times} \mathbf{r} = \mathbf{t} \times \mathbf{r}$ 以及 $\mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3) = \mathbf{b}_2 \cdot (\mathbf{b}_3 \times \mathbf{b}_1) = -\mathbf{b}_2 \cdot (\mathbf{b}_1 \times \mathbf{b}_3)$ 的性质(其中 $\mathbf{b}_i (i = 1, 2, 3)$ 为 3×1 大小的向量), 不难证明, $\mathbf{A}^T [\mathbf{A} \mathbf{t}]_{\times} \mathbf{A} = \rho [\mathbf{t}]_{\times}$, 因此

$$\mathbf{F} = \rho (\mathbf{K}_2 \mathbf{R}_2)^{-T} \mathbf{M} (\mathbf{K}_1 \mathbf{R}_1)^{-1} \quad (9)$$

其中, $\rho = \mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)$ 为一个常比例因子。

由此可见,传统的基础矩阵 \mathbf{F} 被分解为独立的两部分,即与摄像机位置相关的 \mathbf{M} 以及各自 PTZ 参数相关的 $\mathbf{K}_1, \mathbf{R}_1$ 和 $\mathbf{K}_2, \mathbf{R}_2$ 。类比极线几何中的理论, $\mathbf{M} \mathbf{X}_1$ 代表第 2 个摄像机坐标系下某一个过原点和极点的平面(实际上就是 \mathbf{X}_1 所在的极平面)的法向量,而 $\mathbf{X}_2^T \mathbf{M} \mathbf{X}_1 = 0$ 则表示 \mathbf{X}_2 处于该平面上。因此对极点而言,有 $\mathbf{M} \mathbf{E}_1 = \mathbf{E}_2^T \mathbf{M} = 0$ 。

上面的证明也提供了另外一种估计极点的方法,即先通过找到两个摄像机坐标下的对应点对,然后通过估计矩阵 \mathbf{M} , 进而得到极点 $\{\mathbf{E}_1, \mathbf{E}_2\}$ 的值。另外,这个结论也同时说明,这种标定方法与两个镜头之间的相对位置无关,即只要两个镜头相对位置固定(即 (\mathbf{R}, \mathbf{t}) 参数固定),则标定结果不发生改变。

6 实验结果

实验流程如下:

首先选定 N 组标定图像对,然后分别进行特征点的提取、匹配,并估计其基础矩阵。由于此步骤为离线进行,因此为了提高基础矩阵以及特征点匹配的准确度,要求图像对尽可能满足以下条件:

- (1) 每组图像对的 Zoom 参数设定为一致,以减少不同尺度对标定的影响;
- (2) 每组图像对的两图像重叠的视场尽可能

大,以保证匹配所需要的特征点数目;

(3) 与所有图像对对应的视角尽可能均匀地分布在两镜头中心连线的四周,并使样本尽量均匀,以减少误差。

图 5 给出了其中的两组标定图像对(视角大致相差 180°)及其特征点的匹配结果。其中左右两图分别表示来自摄像机 1 和 2 的图像,图上标记的黑圆点表示所有检测到的 Harris 角点,线段表示左右两图已匹配上的特征点的连线。

然后,利用估计出的多个基础矩阵以及式(5)进行极点对 $\{E_1, E_2\}$ 的估计。

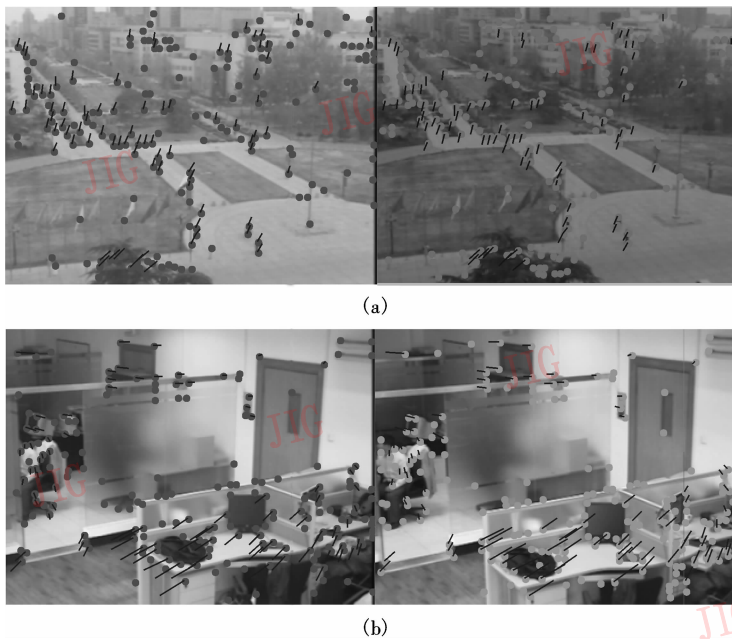


图 5 用于标定公共坐标系的图像对中随机选择的两组

Fig. 5 Two image pairs used for common coordinate's establishment

最后,利用匹配上的特征点对来估计参考零经度线。

为评价本文方法的效果,本文用立体校正结果来检验公共坐标系标定结果的好坏。为了量化标定结果,本文将标定误差定义为匹配的特征点在经度分量(对应的就是校正图像中的垂直坐标,它们之间相差 1 rad 与像素之间的尺度因子,变换关系参考文献[10])上的平均偏差,即

$$\delta = \frac{1}{S} \sum_s |\alpha_s^1 - \alpha_s^2| \quad (10)$$

其中, S 为所有特征点个数。 δ 越小,表明校正结果越好。需要说明的是,这种量化只有在数据量(标定图像对的数目 N)较大,并且与标定图像对应的

视角在全空间分布相对均匀的情况下有意义,因为事实上,虽然利用很少的图像对(甚至 $N = 1$)就可以给出标定结果,并且总可以找到合适的零经度参考向量,使得平均偏差 δ 较小,但是由于镜头模型的误差和每幅图像从摄像机获取的 PTZ 参数存在一定的误差,因此求出的极点对 $\{E_1, E_2\}$ 可能存在较大的误差,这样,即使在当前的匹配特征点集上,平均偏差 δ 很小,可是对于新的图像对,特别是视角接近极点的情况,校正后的平均偏差则可能会很大。一般来说,极点估计的偏差越大,平均偏差 δ 在较大数据上统计不会太小。

为了对比单组标定和多组标定的性能差异,实验中选择 $N = 6$ 组标定图像对进行标定,这 6 组数

据对应在室内和室外(视角相差 180° 左右)各 3 组,视角相对均匀分布。

在本实验中,两个 PTZ 镜头基本水平,并平行放置,实验估计得到的极点坐标为

$$\begin{cases} E_1 = [0.028\ 6, & -0.023\ 5, & 0.999\ 31]^T \\ E_2 = [0.042\ 1, & -0.005\ 2, & 0.999\ 11]^T \end{cases} \quad (11)$$

立体校正在这 6 组图像对中所有特征点对(524 对)上的平均偏差为 $\delta = 0.001\ 4\ \text{rad}$ 。

表 1 给出了单组标定和多组标定的一个结果比较,实验结果表明,当采用的标定图像对数目比较少的时候,所求出的极点偏差可能比较大;表 2 给出了采用 6 组标定数据和 10 组标定数据(包含前 6 组数据)的结果比较,由表 2 可以看出,多组标定的偏差 δ 和极点 E_1 的相对偏差明显小于单组标定结果。而且由于采用了更多数据,10 组数据标定的偏差 δ 小于 6 组数据标定的结果。由此可以得出一般实验结论:标定图像对的数目越多,标定的稳定性、准确性越高。

表 1 单组标定结果与 6 组标定结果的偏差

Tab. 1 Comparison between calibration result from single image pair and that from 6 image pairs

参与标定的图像对序号	偏差 δ (rad) (524 组特征点对)	与采用 6 组数据标定结果的极点 E_1^* 的偏差角度 (rad)
1	0.011 5	0.197 1
2	0.005 8	0.054 9
3	0.004 4	0.262 3
4	0.003 9	0.248 8
5	0.008 8	0.086 6
6	0.003 6	0.099 5
1~6	0.001 4	—

表 2 6 组标定结果与 10 组标定结果的偏差

Tab. 2 Comparison between calibration result from 6 image pair and that from 10 image pairs

参与标定的图像对数目	偏差 δ (rad) (917 组特征点对)	极点 E_1 的相对偏差角度 (rad)
6	0.002 7	0.005 4
10	0.001 9	

图 6 为利用 6 组标定数据的标定结果进行立体校正的结果图,其中的两幅图片分别来自两个摄像机,并且 PTZ 参数均有一定差异。在准确的标定下,两幅校正图像中的对应点应该处于同一水平扫描线上。从图 6 中可以看出,该标定结果比较理想。如果标定结果偏差较大,则对应点在校正图像中的垂直坐标会存在一定的偏差。

需要说明的是,极点 E_1 的偏差直接影响校正图像的形变程度,在有些视角的校正图像上,校正偏差可能比较小,但是当视角越接近极点附近时,图像的形变会非常严重,如果极点不正确,那么这种形变可能导致标定偏差很大。这也印证了前面给出平均偏差 δ 定义时的条件,即只有当极点估计的偏差比较小时, δ 才有意义。

实际标定中,可以首先选出几组标定数据(一般不能太少,否则如前分析,错误的极点估计也可能得到较小的偏差 δ 。本文在实验中选择 $N=6$,并且要求视角相对均匀分布,该要求可以使极点估计不会有较大的误差),同时估计极点;然后添加新的标定图像数据直到估计出的新的极点偏差足够小,最后确认极点位置,并用当前所有匹配的特征点对来估计零经度参考线以及平均偏差 δ ,如果 δ 已经达到要求,则标定停止,否则可以考虑添加新的标定数据。



图 6 根据标定结果进行立体校正的结果图

Fig. 6 Stereo rectification result by using the calibration result

另外,关于镜头架设的几点说明:本文提出的标定方法和摄像机的架设姿态无关,但是因为标定中需要对两个镜头图像中的特征点进行匹配,所以笔者建议在实用中两个镜头之间的距离不要太大(本实验中设定为0.5~1m,在这一范围内实验结果都比较理想),并且两个镜头尽量以相同的姿态放置,这样,在估计图像对的基础矩阵时,精度会更有保证。

7 结 论

本文讨论了一种双目PTZ摄像机系统的标定方法。关于公共坐标系,该方法采用的是球面校正坐标系,其和传统的笛卡尔公共坐标系相比,本文提出的系统的最大优点就在于不仅可以方便地进行立体校正,且可以进一步实现双目立体视觉;而且对于PTZ镜头而言,采用球面坐标能更加方便处理PT参数的变化。这种坐标系最主要的缺点在于不容易将该系统推广到包含更多镜头的系统中,因为对于多个镜头,需要两两建立球面校正坐标系,这也是这种标定方法的一个局限。

在标定方法上,本文提出了一种通过多组基础矩阵进行双PTZ摄像机系统的自动标定算法。实验表明,该标定方法具有较高的鲁棒性;另一方面,由于标定结果只有当镜头改变架设方式时才需要重新进行标定,而不受摄像机PTZ参数改变影响,从而提高了系统的方便性,最后通过立体校正实验验证了本文方法的实用性。

后续工作将引入畸变模型,并分析PTZ参数的不确定性对标定结果的影响,以进一步改善系统性能。

参考文献 (References)

1 Xu Guang-you. Computer Vision [M]. Beijing: Tsinghua University

Press, 1999. [徐光佑. 计算机视觉[M]. 北京: 清华大学出版社. 1999.]

2 You Su-ya, Xu Guang-you. State of the art and Future the development of stereo vision [J]. Journal of Image and Graphics, 1997, 2(1): 17~24. [游素亚, 徐光佑. 立体视觉研究的现状与进展[J]. 中国图象图形学报, 1997, 2(1): 17~24.]

3 Sinha S, Pollefeys M. Towards calibrating a pan-tilt-zoom cameras network [A]. In: Fifth Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras (OMNIVIS) [C], Prague, Czech Republic, 2004.

4 Li Meng-xiang, Lavest Jean-marc. Some aspects of zoom lens camera calibration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(11): 1105~1110.

5 Collins Robert, Tsin Yang-hai. Calibration of an outdoor active camera system [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Ft. Collins, CO, USA, 1999: 528~534.

6 Hayman Eric, Murray David W. The effects of translational misalignment when self-calibrating rotating and zooming cameras [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(8): 1015~1020.

7 Forsyth David A, Ponce Jean. Computer Vision: A Modern Approach [M]. Englewood Cliffs, NJ, USA: Prentice Hall, 2002. [Forsyth D A, Ponce J. 原著, 林学, 王宏等译, 计算机视觉——一种现代方法[M]. 北京: 电子工业出版社, 2004.]

8 Brown Myron Z, Burschka Darius, Hager Gregory D. Advances in computational stereo [J]. IEEE Transactions on Pattern Analysis and Machine. Intelligena, 2003, 25(8): 993~1008.

9 Wan D, Zhou J, Zhang D. A spherical rectification for dual-ptz-camera system [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [C], Honolulu, Hawaii, USA, 2007, 1(1): 777~780.

10 Fischler Martin A, Bolles Robert 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.

11 Ma Song-de, Zhang Zheng-you. Computer Vision-Computing Theory and Algorithm Fundamental [M]. Beijing: Science Press, 1998 (in Chinese) [马颂德, 张正友. 计算机视觉——计算理论与算法基础[M]. 北京: 科学出版社, 1998.]