

## 一种基于双目 PTZ 相机的主从跟踪方法

崔智高<sup>\*①②</sup> 李艾华<sup>②</sup> 姜柯<sup>②</sup> 周杰<sup>①</sup>

<sup>①</sup>(清华大学自动化系 北京 100084)

<sup>②</sup>(第二炮兵工程大学 502 教研室 西安 710025)

**摘要:** 借鉴变色龙视觉的高度独立性、对称性、全局性与选择性兼顾等特点, 该文提出一种基于双 PTZ (Pan-Tilt-Zoom) 相机的主从跟踪方法。由于两个相机的对称性和参数可变性、可控性, 这种方法相对于静止加主动相机的主从跟踪系统, 可以增大监控范围; 相对于多静止加主动相机的系统, 可减小硬件开销; 相对于全向加主动相机的系统, 更有利于信息融合。该文设计了基于球面坐标模型的主从控制方法, 可方便实现两相机在任意 pan-tilt-zoom 参数下的主从模式跟踪, 实现对目标的多尺度视觉关注。在室外场景中进行的多组实验验证了所提方法的有效性。

**关键词:** 目标跟踪; 主从跟踪; PTZ(Pan-Tilt-Zoom)相机; 变色龙视觉; 球面坐标模型

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2013)04-0777-07

DOI: 10.3724/SP.J.1146.2012.01023

## A Master-slave Tracking Algorithm Using Two PTZ Cameras

Cui Zhi-gao<sup>①②</sup> Li Ai-hua<sup>②</sup> Jiang Ke<sup>②</sup> Zhou Jie<sup>①</sup>

<sup>①</sup>(Department of Automation, Tsinghua University, Beijing 100084, China)

<sup>②</sup>(Faculty 502, The Second Artillery Engineering University, Xi'an 710025, China)

**Abstract:** This paper presents a master-slave tracking framework for wide area video surveillance with two Pan-Tilt-Zoom (PTZ) cameras, which is inspired by the chameleon vision of highly independent and symmetric, attention to both overall and local visual events. Because of the symmetry of two cameras and variability and controllability of camera parameters, compared with master-slave tracking system consisting of one static and one PTZ camera, the proposed system increases the surveillance coverage. Compared with system configuration using multiple static cameras and PTZ cameras, the proposed system decreases hardware cost. Compared with surveillance system composed of one omnidirectional and one PTZ camera, the proposed system is easy to fuse to exact more useful information. A master-slave control method is also designed using sphere coordinate model. Each camera can achieve cooperative master-slave tracking at arbitrary pan-tilt-zoom values. Therefore, visual attention of multi-resolution can be obtained. Quantitative results in the outdoor scene show good performance of the proposed approach.

**Key words:** Target tracking; Master-slave tracking; Pan-Tilt-Zoom (PTZ) camera; Chameleon vision; Sphere coordinate model

### 1 引言

随着世界范围内对公共安全以及军事安全的迫切需要, 视觉监控成为计算机视觉领域的重要研究课题。一般来说视觉监控包括以下几个方面的研究内容: 感兴趣目标检测、目标分割与跟踪、行为分析与理解等<sup>[1-3]</sup>。由于目标跟踪是其中的中心环节, 因此得到了国内外学者的广泛关注。传统的跟踪系统<sup>[4]</sup>多采用静止相机, 由于相机视场固定, 分辨率单一, 无法获得跟踪目标的高分辨率图像, 为日后的

查询、举证等工作带来困难; 随着硬件水平的提高, 基于单目 PTZ(Pan-Tilt-Zoom)相机的主动跟踪系统得到了广泛研究与应用<sup>[5]</sup>, 这些系统可使目标以较大尺度出现在图像中心, 但由于视场狭小丢失了全景信息, 难以直观获得目标在场景中的位置。

针对上述缺陷, 包含 PTZ 相机的多目视觉系统成为视觉监控中的研究热点, 这些系统一般工作在主从模式下, 主相机实现目标在全景下的跟踪, 并控制 PTZ 相机对目标进行主动跟踪。如文献[6,7]提出了一个静止相机与一个 PTZ 相机相结合的系统构架, 该系统主要问题在于监控视场有限, 仅局限于静止摄像机的监控视场, 无法适应于大场景下的

2012-08-10 收到, 2012-12-28 改回

国家自然科学基金(61020106004, 61021063, 61225008)资助课题

\*通信作者: 崔智高 cuizg10@mails.tsinghua.edu.cn

监控；为了扩大系统的监控范围，文献[8]提出了多静止相机与 PTZ 相机相结合的系统配置，这无疑增大了系统的硬件开销；此外，文献[9,10]提出了全向相机与 PTZ 相机相结合的跟踪系统配置，但该系统一般应用于室内场景，且由于全向相机分辨率低，较难实现两相机之间的信息融合。

变色龙具有特殊的视觉系统，两只眼睛既可以独立伸缩、转动以获得较大视场，又可以协同实现立体视觉。通过有效地利用自身独特的视觉系统，变色龙可以及时发现敌害并准确定位猎物，这在很大程度上提高了其生存能力。针对变色龙视觉特性，周杰等人<sup>[11-15]</sup>利用两个 PTZ 相机构建了仿变色龙视觉系统，并研究了其中的标定、立体视觉、场景理解、视频稳定化等问题。

受变色龙视觉系统的启发，本文提出一种基于双 PTZ 相机的主从跟踪方法。由于双目 PTZ 相机的对称性和参数的可变性、可控性，相比于传统主从跟踪系统，具有灵活性强、监控范围广、便于信息融合等优势。针对传统主从控制方法的局限性，本文设计了一种基于球面坐标模型的主从控制方法，统一了两相机不同参数下的情况，可以实现两摄像机任意 pan-tilt-zoom 参数下的主从模式跟踪，适应于大范围场景下的监控。实际监控场景中的多组实验结果验证了本文方法的有效性。

## 2 系统结构

系统总体结构如图 1 所示。可分为图像捕获模块，两个摄像机的控制模块以及两个摄像机的跟踪模块。摄像机控制模块一方面的用途是改变摄像机 pan-tilt-zoom 参数以切换监控视场，另一方面的用途是主动跟踪时受另一相机(主相机)控制，从而获得目标的高分辨率图像。两摄像机的跟踪模块具有对称的结构，某一相机发现目标后即成为主相机，另一相机为从相机，主相机在每一帧图像中将跟踪目标的预测值传送给从相机，从相机根据预测值以及主相机 pan-tilt-zoom 参数利用球面坐标模型计算其 pan-tilt-zoom 参数。

## 3 基于球面坐标模型的主从跟踪方法

### 3.1 问题描述

视觉巡视过程中，发现感兴趣目标的相机为主相机，记为 Cam-M，另一相机为从相机，记为 Cam-S。本质上，本文问题可描述为主从跟踪过程中的从相机参数估计问题，即主相机 Cam-M 在任意 pan-tilt-zoom 参数( $P_M, T_M, Z_M$ )下，由跟踪目标在 Cam-M 观测图像  $I_M$  上的运动轨迹  $p_M^t (t = 1, \dots, n)$ ，估计从相机 pan-tilt-zoom 参数( $P_S^t, T_S^t, Z_S^t$ )

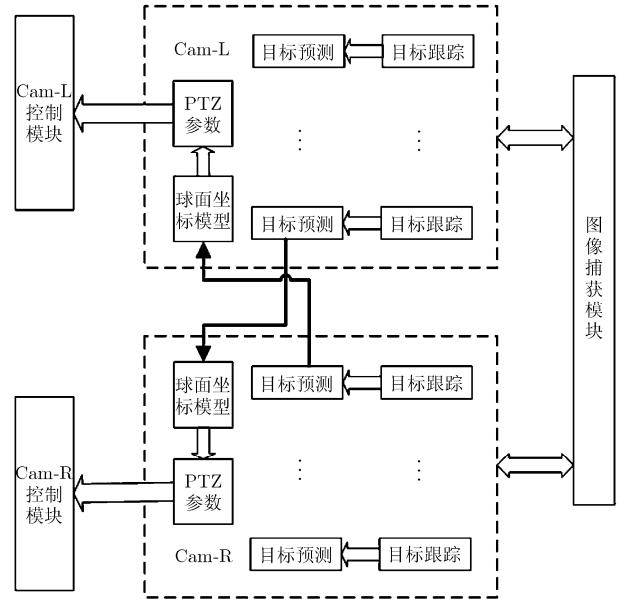


图 1 系统结构图

( $t = 1, \dots, n$ )，使得跟踪目标始终处于 Cam-S 图像中心位置。本文之所以强调 Cam-M 在任意 pan-tilt-zoom 参数下，是因为实际应用中不同的监控场景或监控任务对应的 pan-tilt-zoom 参数不同，实现任意参数下的主从模式跟踪，具有更大的实际应用价值。

传统静止相机加 PTZ 相机的系统配置中<sup>[6,7]</sup>，两摄像机的协同控制采用离散采样的标定方法。设  $[x_s, y_s, 1]^T$  为静止摄像机图像中目标的齐次坐标， $[p, t]^T$  为使该图像坐标位于 PTZ 相机图像中心的参数，通过手动采集一系列图像点及其对应的参数，可获得两者之间的匹配函数  $F$ 。

针对本文系统，如果采用上文提到的离散采样映射的标定方法，当 Cam-M 参数变化时，需要重新建立 Cam-M 图像坐标与 Cam-S 的 pan-tilt-zoom 参数之间的映射关系，这无疑会影响系统的应用。本文设计了一种基于球面坐标模型的主从控制方法，只要两个相机安装固定，不受 Cam-M 参数改变的影响，可实现任意 pan-tilt-zoom 参数下的主从模式跟踪。

### 3.2 具体实现

**3.2.1 主相机目标跟踪** 主相机 Cam-M 发现目标后，采用静止相机下的跟踪方法逐帧跟踪目标。基于系统的实时性要求，本文采用基于颜色直方图的 Mean Shift 跟踪算法<sup>[16]</sup>。颜色模型选用受光照影响较小的 H 分量作为目标的颜色特征向量，并把 H 分量量化为 16 个色度空间，利用该分量核函数加权直方图作为目标模型，通过 Mean Shift 的迭代运算，可在当前帧中搜索与目标模型最相似的潜在目标，目标模型与候选模型的相似度用巴氏系数度量。

Mean Shift 跟踪算法一般采用前一帧目标的中心作为当前帧跟踪窗口的初值进行迭代，当目标运动速度较快时，有可能丢失跟踪目标，为此，本文引入 Kalman 滤波器。本文中 Kalman 滤波器有两个作用：一是通过先前帧的观测预测目标在当前帧的中心位置  $p_M^t$ ，并在当前帧的预测邻域内进行目标搜索与匹配，这样既可提高迭代收敛速度，又可以减小目标丢失的可能性；二是将跟踪目标中心位置的预测值  $p_M^t$  传送给从相机，从而可以抵消图像处理 and 从相机机械运动的延时误差，保证目标处于从相机中心位置。如图 2 所示。

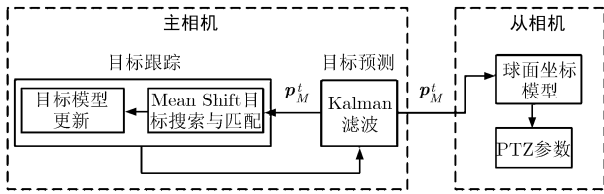


图2 目标跟踪与目标预测示意图

**3.2.2 从相机参数估计** 主相机能够及时捕捉目标运动的全景信息，但分辨率较低，目标的精确定位与跟踪需要由从相机完成，从相机需要不断调整自身参数以适应目标位置的变化。本文设计了基于球面坐标模型的从相机参数估计方法，其基本思想是对两个摄像机坐标系分别求取一个变换矩阵  $R_M$  和  $R_S$ ，建立球面坐标系，并使两摄像机坐标系上对应点在球面坐标系上的经度值保持一致，纬度差可用来表达视角差异。这样以两个摄像机的球面坐标系为桥梁，通过一系列坐标变换，可准确地估计从相机参数。

为了估计变换矩阵  $R_M$  和  $R_S$ ，采用文献[11,13]的方法：手动采集监控场景中的  $N$  对图像，并使每对图像的视场基本保持一致以获得更多的匹配特征点对，然后采用多组基础矩阵的方法计算变换矩阵  $R_M$  和  $R_S$ ，由于此方法是基于统计意义上的，理论上讲  $N$  越大， $R_M$  和  $R_S$  估计越准确，本文中取  $N = 20$ 。当摄像机安装固定后，即使两摄像机 pan-tilt-zoom 参数变化， $R_M$  和  $R_S$  也不会改变，因此可以方便处理相机参数变化的情形。

从相机参数估计过程具体可以分为 4 个步骤，如图 3 所示。

**步骤 1** Cam-M 图像坐标到摄像机坐标的变换。设  $p_M^t = [u_M^t, v_M^t]^T$  为  $t$  时刻跟踪目标的预测值对应的图像坐标，根据单目 PTZ 摄像机模型，可以计算  $p_M^t$  所确定的摄像机坐标系下的一条射线，如式(1)所示。其中  $\hat{p}_M^t$  代表齐次坐标， $\kappa$  为某个尺度因子，

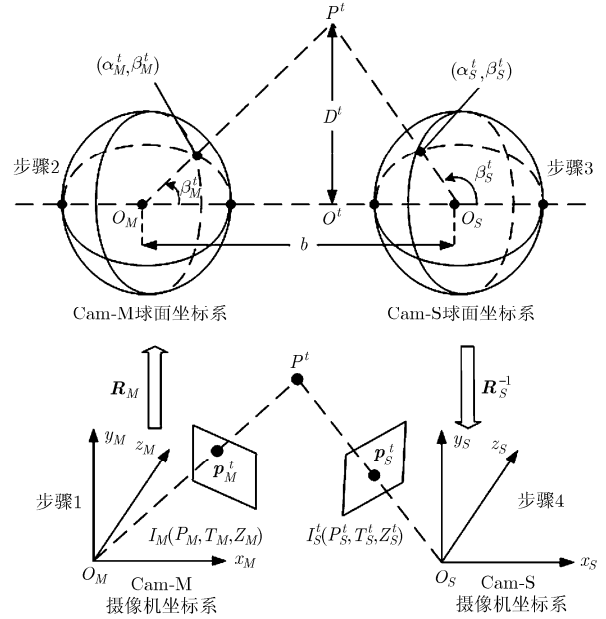


图3 从相机参数估计流程

以满足  $\|Y_M^t\| = 1$ 。

$$Y_M^t = \kappa R^{-1}(P_M, T_M) K^{-1}(Z_M) \hat{p}_M^t \quad (1)$$

本文采用匹配特征点[17]的方法对 PTZ 相机进行标定，摄像机参数矩阵  $R$  和  $K$  可直接通过摄像机 pan-tilt-zoom 参数进行计算。

**步骤 2** Cam-M 摄像机坐标到球面坐标的变换。将  $Y_M^t$  变换到 Cam-M 对应的球面坐标系，并计算经度  $\alpha_M^t$  和纬度  $\beta_M^t$ ，计算公式如式(2)所示。

$$\left. \begin{aligned} Y_{M \rightarrow r}^t &= R_M Y_M^t \\ \alpha_M^t &= \arctan(Y_{M \rightarrow r}^t(3), Y_{M \rightarrow r}^t(2)) \\ \beta_M^t &= \arccos(Y_{M \rightarrow r}^t(1)) \end{aligned} \right\} \quad (2)$$

式中各变量定义如下： $R_M$  为 Cam-M 摄像机坐标系到球面坐标系的变换矩阵， $Y_{M \rightarrow r}^t$  为  $Y_M^t$  对应的球面坐标系上的笛卡尔坐标， $Y_{M \rightarrow r}^t(n)$  表示向量  $Y_{M \rightarrow r}^t$  的第  $n$  个元素。

**步骤 3** Cam-M 球面坐标到 Cam-S 球面坐标的映射。此步骤的目的是由 Cam-M 球面坐标系下的经纬度  $\alpha_M^t, \beta_M^t$  估计对应点在 Cam-S 球面坐标系下的经纬度  $\alpha_S^t, \beta_S^t$ 。由于建立球面坐标模型的目的是使两摄像机坐标系下对应点的经度值保持一致，因此  $\alpha_S^t = \alpha_M^t$ 。为了估计  $\beta_S^t$ ，假设  $t$  时刻目标深度已知，用  $D_t$  表示，深度定义为  $t$  时刻目标  $P$  到基线  $O_M O_S$  的距离，其与基线交于点  $O^t$ ， $b$  为两摄像机的基线长度，如图 3 所示。

设  $O_M O^t = x_M^t$ ， $O_S O^t = x_S^t$ ，则有如下几何关系：

$$\left. \begin{aligned} \tan \beta_M^t &= D^t / x_M^t \\ \tan(\pi - \beta_S^t) &= D^t / x_S^t \\ b &= x_M^t + x_S^t \end{aligned} \right\} \quad (3)$$

因此

$$\beta_S^t = \arctan \left( \frac{-D^t \cdot \tan \beta_M^t}{b \cdot \tan \beta_M^t - D^t} \right) \quad (4)$$

事实上场景中目标的深度  $D^t$  是未知的, 系统无法准确估计目标在 Cam-S 球面坐标系下的纬度  $\beta_S^t$ 。本质上, 计算目标深度  $D^t$  和计算目标在 Cam-S 球面坐标系下的纬度  $\beta_S^t$  是基本等价的, 它们可看作 chicken-and-egg 问题。

考虑到大场景监控中, 两相机的基线长度  $b$  相对场景深度非常小。因此本文给定场景深度范围  $D_{\min}$  和  $D_{\max}$ , 分别估计对应的纬度  $\beta_{S \rightarrow \max}^t$  和  $\beta_{S \rightarrow \min}^t$ , 如图4所示。

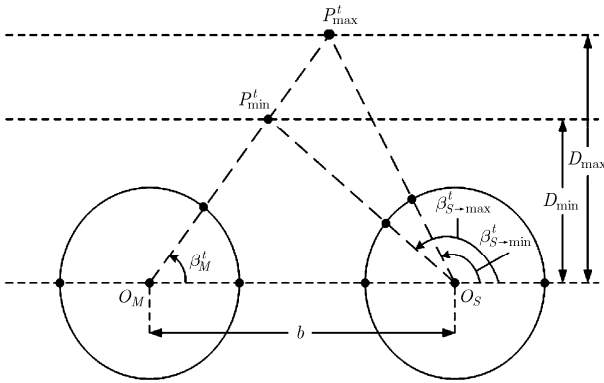


图4 根据深度范围确定从相机纬度范围

由于场景深度远大于两相机之间的距离,  $\beta_{S \rightarrow \max}^t$  和  $\beta_{S \rightarrow \min}^t$  之间的差异很小, 此外考虑到系统的实时性要求, 采用线性加权的方法估计从相机纬度值  $\beta_S^t$ , 计算公式如式(5)所示。

$$\left. \begin{aligned} \beta_{S \rightarrow \max}^t &= \arctan \left( \frac{-D_{\min} \cdot \tan \beta_M^t}{b \cdot \tan \beta_M^t - D_{\min}} \right) \\ \beta_{S \rightarrow \min}^t &= \arctan \left( \frac{-D_{\max} \cdot \tan \beta_M^t}{b \cdot \tan \beta_M^t - D_{\max}} \right) \\ \beta_S^t &= \lambda \cdot \beta_{S \rightarrow \max}^t + \gamma \cdot \beta_{S \rightarrow \min}^t \end{aligned} \right\} \quad (5)$$

其中  $\lambda$  和  $\gamma$  为加权系数, 满足  $\lambda + \gamma = 1$ 。

步骤4 Cam-S 球面坐标到摄像机坐标的变换及  $(P_S^t, T_S^t, Z_S^t)$  估计。估计出目标在 Cam-S 球面坐标系下的经纬度坐标  $\alpha_S^t$  和  $\beta_S^t$  后, 需要首先将其变换到 Cam-S 摄像机坐标系下, 此过程类似于步骤2的逆过程, 如式(6)。其中  $\mathbf{Y}_{S \rightarrow r}^t$  为目标在 Cam-S 球面坐标系下的笛卡尔坐标,  $\mathbf{R}_S^{-1}$  为 Cam-S 摄像机坐标系

到球面坐标系变换矩阵  $\mathbf{R}_S$  的逆矩阵,  $\mathbf{Y}_S^t$  为目标在 Cam-S 摄像机坐标系下的坐标。

$$\left. \begin{aligned} \mathbf{Y}_{S \rightarrow r}^t(1) &= \cos \beta_S^t \\ \mathbf{Y}_{S \rightarrow r}^t(2) &= \sin \beta_S^t \cos \alpha_S^t \\ \mathbf{Y}_{S \rightarrow r}^t(3) &= \sin \beta_S^t \sin \alpha_S^t \\ \mathbf{Y}_S^t &= \mathbf{R}_S^{-1} \mathbf{Y}_{S \rightarrow r}^t \end{aligned} \right\} \quad (6)$$

根据 pan-tilt 参数在摄像机坐标系下的物理意义, 给定 Cam-S 摄像机坐标系下的观测点  $\mathbf{Y}_S^t$ , 可计算其对应的 pan-tilt 参数  $P_S^t, T_S^t$ , 使得 Cam-S 运动到该参数值时, 点  $\mathbf{Y}_S^t$  在图像上的观测位于主点位置(即光轴穿过  $\mathbf{Y}_S^t$ )。对于 zoom 参数  $Z_S^t$  可以给定一较大值, 如  $Z_S^t = 18$ 。

**3.2.3 从相机控制** 估计出从相机 pan-tilt-zoom 参数后, 即可通过串口发送命令控制相机运动到指定参数。PTZ 控制包括参数控制和速度控制两部分。对于参数控制, 由于本文在主相机目标跟踪环节引入了预测机制, 因此可以部分抵消指令发出的通讯延迟以及相机机械运动的延时误差, 保证目标处于从相机中心位置。此外, 相机 pan-tilt 运动的速度应与目标运动的速度成正比, 因此本文在主相机中度量前一帧目标中心位置和当前帧目标预测位置的差异, 如果某个方向坐标偏移较大, 则给定一较大速度, 反之, 则给定一较小速度( $x$  方向坐标差异对应 pan 控制速度,  $y$  方向坐标差异对应 tilt 控制速度), 这样可保证跟踪的平滑性, 并可降低相机运动带来的图像模糊。

**3.2.4 高分辨率结果图** 跟踪结束后, 系统以高分辨率全景形式输出, 其作用在于可以提高可视化效果, 并且有利于后期进行行为分析、犯罪取证等应用。

首先, 采用基于特征点和直接像素相结合的方法<sup>[14]</sup>对两相机的同步帧图像进行配准, 由于基线长度相对于监控场景深度可忽略不计, 本文采用仿射配准模型。采用 running average<sup>[18]</sup>方法估计主相机背景区域并分离前景, 通过两相机同步帧不同分辨率图像间的配准模型, 估计出从相机对应帧图像中的前景区域和背景区域。

然后, 将高分辨率全景图  $I_H$  的大小设为原始图像大小的  $k$  倍, 设  $\mathbf{A}_{LH}^n$  代表第  $n$  帧主相机低分辨率图像  $I_L^n$  和从相机高分辨率图像  $I_H^n$  的配准模型, 则  $I_H$  与  $I_H^n$  之间的配准模型为

$$\mathbf{A}^n = \mathbf{A}_{LH}^n \times \begin{bmatrix} 1/k & 0 & 1-1/k \\ 0 & 1/k & 1-1/k \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

依次将从相机每一帧高分辨率图像的背景区域

映射到  $I_H$  中, 并采用 running average<sup>[18]</sup> 的方法进行更新。

最后, 依次将从相机每一帧高分辨率图像的前景区域映射到  $I_H$  中, 生成每一帧的高分辨率全景图。

### 4 实验结果与分析

#### 4.1 系统配置

本文系统使用 SONY EVI D70 系列 PTZ 相机, 两个 PTZ 相机固定在室内屋顶用来监控室外大场景。两个 PTZ 相机通过 RS232 串口总线控制, 捕获图像的大小为  $320 \times 240$ , 整个系统在 C++ 环境执行, 运行环境为 3.0 G CPU 和 1 G 内存的 PC 机。

实验中给定场景深度范围为 50~200 m, 两相机基线长度为 0.78 m, 由于监控场景较大, 跟踪目标尺度变化较小, 本文对从相机给定固定的 zoom 值, 令  $zoom = 18$ 。加权系数  $\lambda$  和  $\gamma$  取 0.5, 生成的高分辨率全景图大小为  $1600 \times 1200$ , 即  $k = 5$ 。在室外场景中进行了多组实验, 验证了方法的有效性, 实验结果可从 <http://ivg.au.tsinghua.edu.cn/index.php?n=Main.Data> 下载。下文中给出了其中的两个实验结果。

#### 4.2 实验结果

**实验 1** Cam-L 发现感兴趣目标成为主相机 (Cam-M), 对应的 pan-tilt-zoom 参数为  $p = -63.52, t = -11.32, z = 11.50$ 。实验结果如图 5 所示。图 5(a) 给出了其中几帧的跟踪结果, 在第 1 帧图像中 Cam-L 发现目标并开始跟踪, Cam-R 从初始位置开始调整 PTZ 参数, 在第 48 帧获得目标的高分辨率

图像, 中间图像为第 195 帧、310 帧以及 460 帧的跟踪结果, 第 565 帧为目标离开 Cam-L 视场, 跟踪结束, 两相机恢复到预先设定位置。图 5(b) 为高分辨率全景拼接结果(为了可视效果及节省空间, 本文将 48 帧、195 帧、310 帧以及 460 帧的前景映射到一张图像中)。

**实验 2** Cam-R 发现感兴趣目标成为主相机 (Cam-M), 对应的 pan-tilt-zoom 参数为  $p = -90.66, t = -13.50, z = 10.00$ 。实验结果如图 6 所示(说明可参照实验 1)。

为了验证从相机主动跟踪过程中参数估计的准确性, 本文以实验 1 为例, 根据目标在主相机的运动轨迹, 手动调整从相机参数, 使目标在主相机每一帧图像上轨迹点坐标处于从相机中心位置, 并记录从相机 pan-tilt 参数, 该参数可以看作真实值, 将其与本文从相机参数估计结果进行比较, 如图 7 所示。可以看出, 本文方法能够比较精确地估计从相机参数, 其中 pan 参数的平均估计误差为  $0.243^\circ$ , tilt 参数的平均估计误差为  $0.299^\circ$ 。

### 5 结束语

本文借鉴变色龙视觉系统的机制, 利用两个 PTZ 相机构建了能够实现大场景下广域监视和高分辨率精确关注的主从跟踪系统。该系统相对于其它系统具有灵活性强、监控范围广、便于信息融合等优势。针对传统主从控制方法的局限性, 我们设计了一种基于球面坐标模型的主从控制策略, 可方便处理两相机参数变化的情形, 能够实现两摄像机在任意参数下的主从模式跟踪。此外, 为了直观获得

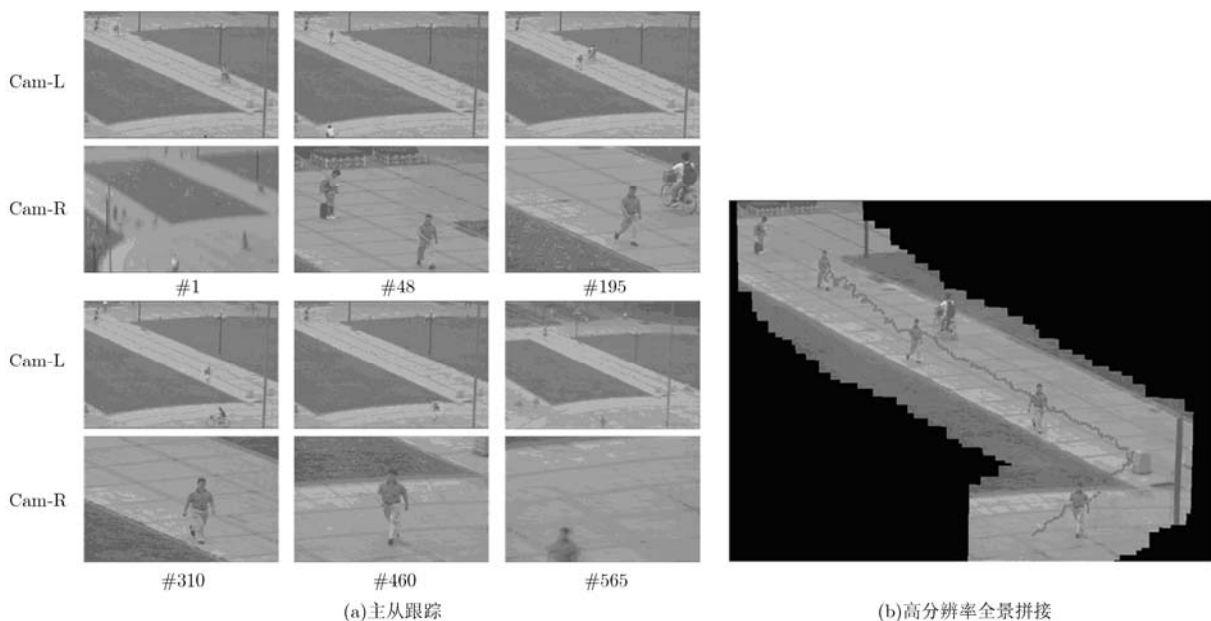


图5 主从跟踪与高分辨率全景拼接结果(Cam-L 为主相机)

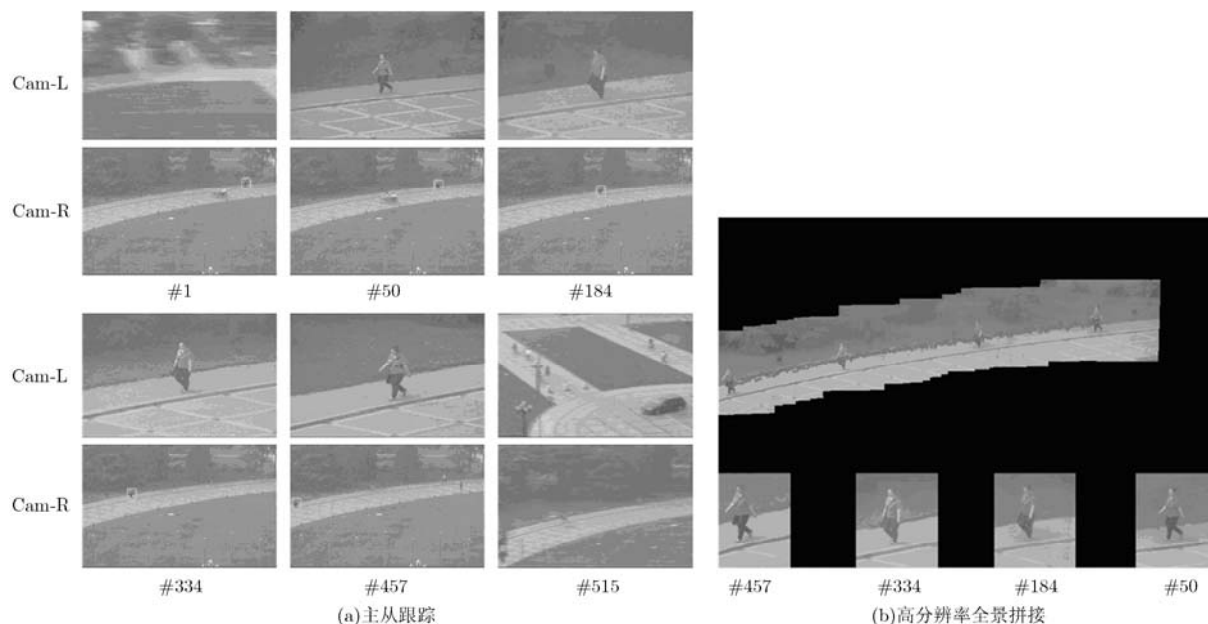


图6 主从跟踪与高分辨率全景拼接结果(Cam-R 为主相机)

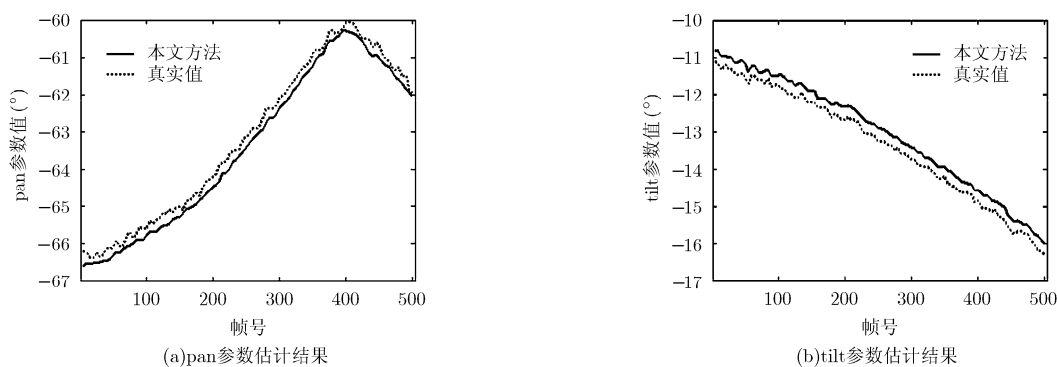


图7 pan 参数和 tilt 参数估计结果

目标运动的高分辨率信息,系统以高分辨率全景作为结果输出,增强了系统的可视性和实用性。在室外场景中进行了多组实验,验证了方法的有效性,为大场景下可变视角监控以及多分辨率视觉跟踪提供了一条有价值的途径。

本文下一步将构建准确的场景深度模型以实现从相机的精确控制,此外,考虑到一般监控场景下,地平面总是可见的,研究如何利用地平面约束实现主从模式跟踪也是本文下一步的研究内容。

### 参考文献

- [1] Salti S, Cavallaro A, and Stefano L D. Adaptive appearance modeling for video tracking: survey and evaluation[J]. *IEEE Transactions on Image Processing*, 2012, 21(10): 4334-4348.
- [2] 李拥军, 曾标, 徐克付, 等. 复杂背景下基于贝叶斯-全概率联合估计的前景检测[J]. *电子与信息学报*, 2012, 34(2): 388-392. Li Yong-jun, Zeng Biao, Xu Ke-fu, *et al.*. Foreground object detection in complex background based on Bayes-total probability joint estimation[J]. *Journal of Electronics & Information Technology*, 2012, 34(2): 388-392.
- [3] Jiang F, Yuan J S, Tsaftaris S, *et al.*. Anomalous video event detection using spatiotemporal context[J]. *Computer Vision and Image Understanding*, 2011, 115(3): 323-333.
- [4] 赵高鹏, 薄煜明, 尹明锋. 一种红外和可见光双通道视频目标跟踪方法[J]. *电子与信息学报*, 2012, 34(3): 529-534. Zhao Gao-peng, Bo Yu-ming, and Yin Ming-feng. An object tracking method based on infrared and visible dual-channel video[J]. *Journal of Electronics & Information Technology*, 2012, 34(3): 529-534.
- [5] Varcheie P D Z and Bilodeau G A. People tracking using a network-based PTZ camera[J]. *Machine Vision and Application*, 2011, 22(4): 671-690.
- [6] Zhou X H, Collins R T, Kanade T, *et al.*. A master-slave system to acquire biometric imagery of humans at a distance[C]. *ACM SIGMM International Workshop on Video Surveillance*, Berkeley, USA, Nov. 7, 2003: 113-120.

- [7] Choi H C, Park U, and Jain A K. PTZ camera assisted face acquisition, tracking & recognition[C]. IEEE International Conference on Biometrics: Theory, Application and Systems, Washington, USA, Sept. 27-29, 2010: 1-6.
- [8] Krahnstoever N O, Yu T, Lim S N, *et al.*. Collaborative real-time control of active cameras in large-scale surveillance systems[C]. Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications, Marseille, France, Oct. 18, 2008: 1-12.
- [9] Chen C H, Yao Y, Page D, *et al.*. Heterogeneous fusion of omnidirectional and PTZ cameras for multiple object tracking[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2008, 18(8): 1052-1063.
- [10] 毛晓波, 陈铁军. 基于人类视觉特性的机器视觉系统[J]. 仪器仪表学报, 2010, 31(4): 832-836.  
Mao Xiao-bo and Chen Tie-jun. Machine vision system based on human vision features[J]. *Chinese Journal of Scientific Instrument*, 2010, 31(4): 832-836.
- [11] Wan D R and Zhou J. Stereo vision using two PTZ cameras[J]. *Computer Vision and Image Understanding*, 2008, 112(2): 184-194.
- [12] Wan D R and Zhou J. Multi-resolution and wide-scope depth estimation using a dual-PTZ-camera system[J]. *IEEE Transactions on Image Processing*, 2009, 18(3): 677-682.
- [13] Wan D R and Zhou J. Self-calibration of spherical rectification for a PTZ-stereo system[J]. *Image and Vision Computing*, 2010, 28(3): 367-375.
- [14] Zhou J, Hu H, and Wan D R. Video stabilization and completion using two cameras[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, 21(12): 1879-1889.
- [15] Zhou J, Wan D R, and Wu Y. The chameleon-like vision system[J]. *IEEE Signal Processing Magazine*, 2010, 27(5): 91-101.
- [16] Comaniciu D, Ramesh V, and Meer P. Kernel-based object tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564-577.
- [17] Sinha S N and Pollefeys M. Pan-tilt-zoom camera calibration and high-resolution mosaic generation[J]. *Computer Vision and Image Understanding*, 2006, 103(3): 170-183.
- [18] Cucchiara R, Grana C, Piccardi M, *et al.*. Detecting moving objects, ghosts, and shadows in video streams[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(10): 1337-1342.
- 崔智高: 男, 1984年生, 博士, 研究方向为双目视频监控。  
李艾华: 男, 1966年生, 教授, 博士生导师, 研究方向为视频监控、信号处理。  
姜柯: 男, 1985年生, 博士, 研究方向为视频监控。  
周杰: 男, 1968年生, 教授, 博士生导师, 研究方向为双目视频监控、模式识别、图像处理等。