

# 多摄像头协同感知系统的设计与实现

胡伏原 张艳宁 李映 杨兴 袁金刚

(西北工业大学计算机学院, 西安 710072)

**摘要** 为了提供连续的实时监控,提高区域安全性,利用多个活动像机,设计了一个对多个运动目标进行无缝检测和跟踪的协同感知系统,并提供给用户一个可视的3维场景。最后给出了实验室的应用实例。

**关键词** 视频网络 协同感知 目标检测 目标跟踪

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2006)12-1849-05

## Design and Implementation of Multi-sensor Cooperative Perception System

HU Fu-yuan, ZHANG Yan-ning, LI Ying, YANG Xing, YUAN Jing-gang

(Computer and Science College of Northwestern Polytechnical University, Xi'an 710072)

**Abstract** To provide continuous 24-hour monitoring of surveillance video, and improve the situational awareness of security providers and decision makers, this paper designs the semiautonomous surveillance system to seamlessly track objects using a network of cooperating active sensors, and presents a visual 3D scene. A instance of lab is also given in the end.

**Keywords** video network, cooperative perception, object detection, object tracking

## 1 引言

随着信息获取技术的发展,获取目标图像已经形成多种传感器、高空间分辨率、实时/准实时的对目标观测能力,使得协同利用各种图像源进行环境监控和目标实时分析成为必要。目前对视频网络的研究吸引了大量的研究机构,比如CMU于1997年开始的VSAM项目<sup>[1]</sup>,欧洲的ESPRIT PASSWORDS<sup>[2]</sup>、AVS-PV<sup>[3]</sup>和VIEWS<sup>[4,5]</sup>项目,日本的CDV项目<sup>[6]</sup>,国内上海交通大学、中国科学院自动化研究所等也有相关的研究。但是国内研究水平较世界先进国家还有一定的差距,应用方面总体上远没达到国民经济建设的需求。

我们正在研究的多摄像头协同感知系统(MSCPS),采用分布式的结构,每个传感器处理单元(VPU)分别进行目标的检测、跟踪、识别和定位,尽量减轻中央处理单元(CCU)的负载,然后通过

CCU进行时间虚拟同步、数据关联、摄像机的调度和分配以及可视化和虚拟漫游,最终给用户提供一个友好的交互界面。

## 2 系统概述

在现实的视频监控系统中,一个摄像头监控一次看到整个监控场景几乎是不可能的。一个好的解决方案是通过多个可以转动的摄像头形成一个视频监控网络,协同地进行连续无缝监控,如图1所示。然而,其带来了很大的挑战(1)多个摄像机信息的融合(2)针对目标和任务,摄像机的任务分配(3)场景的可视化。在整个系统中,基于分布式体系结构,主要分为CCU、VPU和GUI(图形用户接口)3个部分,总体设计框图如图2所示。

(1)VPU由单个摄像机和计算机组成,能够实时获取视频数据,并进行目标的检测、跟踪、识别、行为分析等。同时,利用摄像机的关系实现空间的虚拟

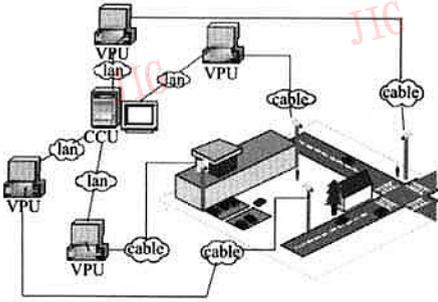


图 1 视频监控网络示意图

Fig. 1 Network of video surveillance

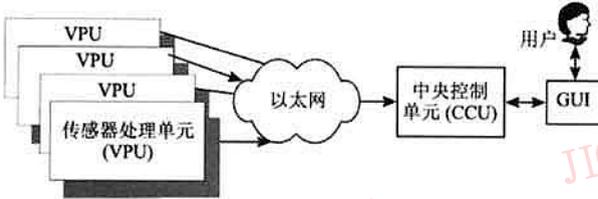


图 2 视频监控网络系统基本框图

Fig. 2 Framework of video surveillance

同步。

(2)CCU 主要完成每个 VPU 的信息融合,实现摄像机的分配、调度及可视化功能。

用户通过 GUI 可以获得目标在 3 维场景、目前地图中的信息;也可以通过设置感兴趣区域(ROI)对特定的区域进行严密的监控。

### 3 系统构成

MSCPS 主要是利用多个传感器对多个目标进行协同的检测、跟踪和识别,为单个用户提供大区域的综合监控结果。力求达到下面 5 个目标:

(1)功能强大,满足用户需求 MSCPS 实现了视频理解、目标定位、时间和空间虚拟同步、目标和航迹关联、摄像机分配和调度、3 维重建、图像配准、虚拟漫游等功能,同时用户也可以通过交互,设定感兴趣区域(ROI)和感兴趣目标(TOI),实现对特定情况的监控。

(2)与设备无关 对于任何非压缩的数字和模拟设备都能实现视频的捕获。

(3)界面友好 用户能够很方便的观看和分析目标、摄像机在场景中的信息。

(4)可扩展性好 目前系统主要研究摄像头的处理,以后在加入声音信号、红外信号也能很容易扩展。

(5)故障检测和容错设计 在某个 SPU 崩溃,可以进行自动恢复和运行,不对其他 SPU 和 CCU 产生影响。

为此, MSCPS 提出采用 3 层结构,第 1 层 GUI 接口,为用户提供支持;第 2 层为数据管理和调度层,主要包括对系统各个模块之间的调度和资源管理,并为 GUI 提供数据支持;第 3 层为各自独立模块的实现,比如视频处理、目标关联等,主要用于实现各个部分的功能。MSCPS 的结构如图 3 所示。

#### 3.1 GUI

对于一个人监控大范围的场景,采用多个显示结果很不合适。本文通过提供一个 3 维界面和一个 2 维地图进行目标的显示和摄像机区域的显示。实现的主要功能为用户对 ROI、感兴趣目标(TOI)和传感器的设置等。

#### 3.2 调度和管理

它是 GUI 和数据处理部分的枢纽,并对各个独

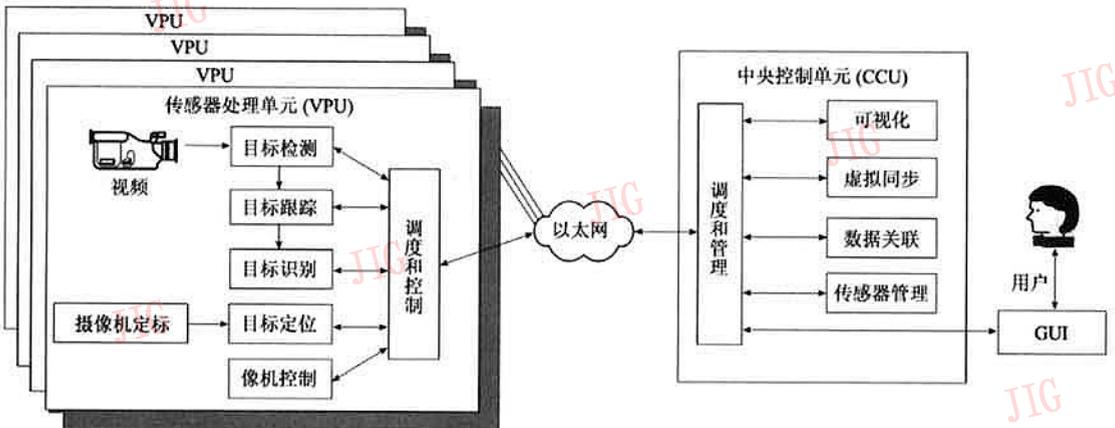


图 3 MSCPS 结构图

Fig. 3 Structure of MSCPS

立模块实现管理。在 VPU 端主要对视频处理模块、目标定位、摄像机控制 3 个模块进行管理,同时负责对处理后的数据发送到 CCU,并接收来自 CCU 的传感器信息和用户设置信息。在 CCU 端,主要管理 3 维场景更新和虚拟漫游、虚拟同步、数据关联和传感器的调度和分配,同时负责和各个 VPU、GUI 通讯。

### 3.3 独立模块构成

在 MSCPS 系统主要包括 8 个独立模块,各个模块的数据通讯主要通过调度和管理模块进行协调。

(1)视频处理模块 主要负责对目标进行检测、跟踪和识别。目标行为的行为分析有待进一步考虑。

(2)目标定位模块 主要通过视频处理的结果,并结合 DEM 数据进行目标在 3 维空间的位置确定。

(3)摄像机控制模块 CCU 的摄像机分配信息,利用控制协议和云台实现对摄像机的状态的控制。

(4)通讯模块 在 SPU 和 CCU 之间都需传送对应的数据。采用 TCP/IP 协议、UDP 协议进行 VPU 和 CCU 之间的数据传输。

(5)虚拟同步 为了构造一个可靠的系统,需要对已有时间点上目标信息进行分析,来估计相同时刻的目标信息。我们称估计得到的目标信息为虚拟目标信息,MSCPS 通过时间、特征的两次同步来完成虚拟目标信息的估计,也即虚拟同步。

(6)数据关联模块 建立适当的匹配模型,实现目标关联和航迹关联,从而形成统一的目标信息和航迹。

(7)摄像头的调度和分配模块 根据目标信息和用户信息,实现摄像机的自动跟踪。主要分为传感器仲裁和协同控制两个部分。

(8)场景的可视化模块 在一个综合的“场景表示”中显示所有传感器的视野范围,形象地表示出对目标的检测、跟踪和识别。这个综合场景表示监控区域的 3 维场景和 2 维地图场景,对于 3 维场景用户可以从不同的角度进行观察运动目标。同时,用户也可以进行特定视频流的查看。

## 4 主要模块实现机制

### 4.1 视频处理实现机制

对目标进行跟踪,首先得保证此目标信息真实

可靠。通常 Kalman 滤波跟踪通过运动检测建立目标和更新目标,普遍采用的是两层数据结构<sup>[7]</sup>:运动检测结果和跟踪处理结果。这种数据结构很容易建立虚假目标和引起目标更新的偏差。MSCPS 提出了 3 层区域链交互结构如图 4 所示,很好地衔接了运动检测与运动跟踪,消除虚假目标。

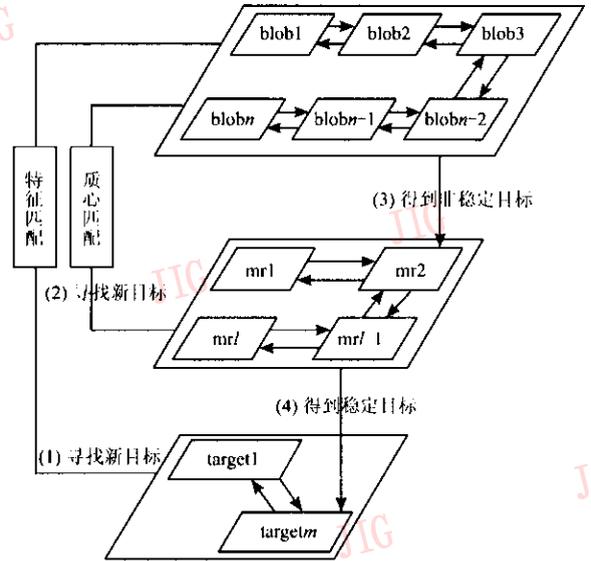


图 4 3 层区域链表结构

Fig. 4 Three-level structure of region list

该结构由 3 层链表组成:第 1 层为 blob 链表,管理运动区域检测得到的 blob;第 2 层为 mr 链表,与第 1 层的 blob 元通过目标质心位置的最近邻匹配进行交互动作,从而完成对图像中新出现且没有确定其稳定性的目标的跟踪。mr 链表中已达到稳定跟踪条件的目标,传送给第 3 层管理器。第 3 层为 target 链表,与第 1 层 blob 链表通过特征匹配交互,跟踪稳定目标和得到非稳定新目标传送给中间目标管理层,同时与中间目标管理层交互,得到稳定的待跟踪新目标。target 元中包含 mr 元的所有参数以及目标的跟踪标志。

### 4.2 通讯模块实现机制

目前, MSCPS 采用单个 CCU 和多个 VPU 相连,当 VPU 发展到一定规模,将采用多级结构,即每个区域由单个 CCU 和 VPU 相连,而在区域之间 CCU 通过上级的 CCU 相连,形成 CCU 级联,实现可重构管理。

MSCPS 主要采用没有冗余头信息的包结构。主要包括 8 个部分的数据段:感知器类型、感知器 ID、感知器状态、目标类型、目标状态、ROI 事件。包头信息

主要用于描述目前的包结构所包括的数据段。所有的通讯在 VPU ,OCU 和 GUI 之间是兼容的。

### 4.3 数据关联模块

视频监控网络系统中存在多个 VPU ,每个节点获取的结果都是相对于整个系统的局部信息 ,且相互之间存在冗余性和互补性 ,所以必须将这些局部处理结果送到 CCU 进行数据融合操作 ,减少观测的不确定性和模糊性 ,提高系统对目标检测、识别和跟踪的能力。MSCPS 主要通过设计接受数据的广义表存储结构进行数据接收 ,利用全局目标轨迹链表结构存储最终的关联结果 ,如图 5 和图 6 所示。

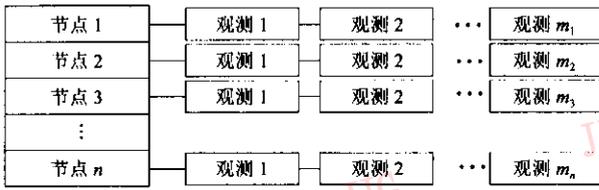


图 5 接收数据的广义表式存储结构

Fig. 5 General lists structure of receiving data

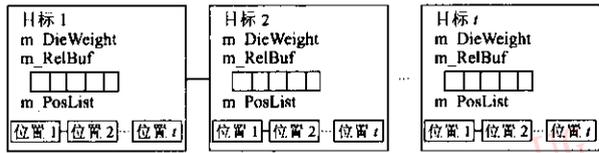


图 6 系统全局目标轨迹链表结构

Fig. 6 Global object track list

### 4.4 虚拟同步模块

各个传感器独立获取的目标信息在时间上不完全相同 ,并且由于网络延时 ,在进行关联计算时往往得不到所需时间点上的数据。时间上的不同步性不仅对于关联算法提出了更高的要求 ,而且也影响了系统的整体性能。因此 ,考虑到系统的可靠性 ,在关联计算前对接收到的数据进行同步处理。本文主要对时间和特征进行同步。时间同步上 ,MSCPS 采用串口进行时间同步 ,其精度在 1ms 左右。特征同步方面 ,通过采用目标缓冲池链 ,其结构如图 7 所示。

### 4.5 摄像头的调度和分配模块

这是协同的核心部分 ,需要解决如何协调相对有限的资源与可能存在的无限任务之间的关系以及如何将抽象的命令转化为具体的控制序列两个问题。即如何将“用摄像机 C 跟踪目标 T”这样的抽象行为转化为“将摄像机 C 从 A 点移动到 B 点”这样一系列能够直接作用于摄像机云台的命令。对于调度模块 ,主要解决目标的可见性以及像机的分配 ,

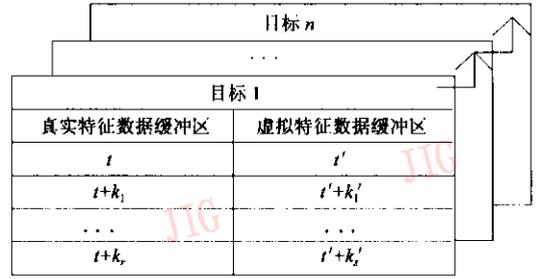


图 7 数据缓冲池链表结构

Fig. 7 Data buffer lists

目标可见性实际是目标是否在一个多边形内部的问题 ,像机的分配实际上是一个优化问题。MSCPS 系统使用二次分配算法 ,其设计原则是尽可能多地跟踪场景中的目标 ,并在此基础上 ,优先跟踪较重要的目标(可人工设定)。第 2 个问题是像机的调度 ,主要是串口通讯问题。

### 4.6 场景可视化模块

场景可视化主要通过 3 维、地图和单路视频流进行显示。其中 3 维的显示、更新和虚拟现实主要通过 OPGl 实现 ,地图主要通过镶嵌技术形成 ,视频流的显示主要利用 DirectShow。目前 3 维重建的精度还不高 ,正在研究和测试阶段。

### 4.7 并发控制

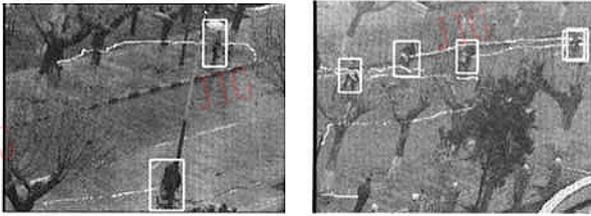
该程序是一个多线程程序 ,为了保证数据的一致性和正确性 ,MSCPS 并发控制以封锁机制实现 ,有共享锁和互斥锁。读操作可以共享资源 ,而写操作相互排斥。封锁粒度主要是设置临界区锁。

## 5 实验结果与分析

本文通过实验室的两个固定的摄像头进行了仿真实验。OCU 端运行在 P4 2.6G 512M 内存上 ,VPU 分别运行在 P4 2.6G 512M 和赛扬 2.6G 512M 内存上 ,VPU 目标检测和跟踪速度均在 22fps 左右。

利用 3 帧差分检测运动目标 ;用 Kalman 和 MeanShift 相结合跟踪运动目标 ;利用与摄像机角度无关信息 ,和最小距离实现目标的航迹融合 ,最后在地图上进行目标轨迹的显示。

图 8 为对实际场景利用一个摄像机对目标进行检测和跟踪的结果。通过设计 3 层结构很好地实现了目标的遮挡处理。在图 8(a)最左方的人被树木遮挡住 ,此时遮掩处理对其运动轨迹做了有效的预测分析。图 8(b)中汽车被图像中的大树完全遮



( a ) 雨天

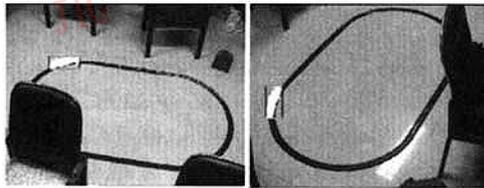
( b ) 雾天

图 8 室外检测和跟踪结果

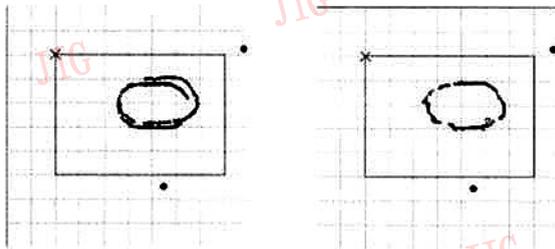
Fig. 8 The result of detection and tracking in outdoor scenes

掩,目标没有消失,其运动轨迹仍然延续。

图 9 ( a ) 为摄像机静止的拍摄环境,并展示了两个像机在某个时间段的跟踪结果。图 9 ( b ) 为目标在 3 维坐标系中观测结果,图 9 ( c ) 为关联结果。从图 9 ( a ) 可以看出,单个像机仅检测和跟踪部分区域,而通过像机间的合作,能够实现对整个实验室的监控,并通过关联处理,得到小车的完整航迹。



( a ) 实验场景



( b ) 观测结果

( c ) 航迹关联结果

图 9 航迹关联结果

Fig. 9 Correlation results

## 6 结 论

本系统正处于研究完善阶段,通过简单的实验进行平台的验证,系统的功能强大,具有灵活、高效、实用、界面友好的特点。整个系统的管理机制和系统构建简捷、易于扩充,图像采集设备的独立性强,易于在不同机器上移植。目前,在摄像头的调度分配方面还正在研究,3 维重建的精度还有待进一步提高。

### 参考文献 (References)

- 1 Collins R, Lipton A, Kanade T, et al. A System for Video Surveillance and Monitoring: VSAM Final Report[R]. CMU-RI-TR-00-12, UA: Carnegie Mellon University, 2000.
- 2 Bogaert M, Chleq N, Cornez P, et al. The passwords project[A]. In: Proceedings of Image Processing[C]. Lausanne, Switzerland, 1996, 3(3):675~678.
- 3 Rota M, Thonnat M. Video sequence interpretation for visual surveillance[A]. In: Proceedings of 3rd Workshop on Visual Surveillance[C], Dublin, Ireland, 2000:59~68.
- 4 Corral D. Views: computer vision for surveillance applications[A]. In: IEEE Colloquium on Active and Passive Techniques for 3-D Vision[C], London, UK, 1992:8/1~8/3.
- 5 Buxton H, Gong S G. Visual surveillance in a dynamic and uncertain world[J]. Artificial Intelligence, 1995, 78(1):431~459.
- 6 Matsuyama T. Cooperative distributed vision-dynamic integration of visual perception[A]. In: Proceedings of Image Understanding Workshop[C], Monterey CA, Mexico, 1998:365~384.
- 7 Stauffer C, Grimson, W E L. Learning patterns of activity using real-time tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8):747~757.