Image Understanding Workshop

Proceedings of a Workshop held in Monterey, California

November 20–23, 1998

Volume I

Sponsored by:
Defense Advanced Research Projects Agency
Information Systems Office

This document contains copies of reports prepared for the DARPA Image Understanding Workshop. Included are Principal Investigator reports and technical results from the basic and strategic computing programs within DARPA/ISO-sponsored projects and certain technical reports from selected scientists from other organizations.

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

Distributed by:
Morgan Kaufmann Publishers Inc.
340 Pine Street, 6th Floor
San Francisco, Calif. 94104-3205
ISBN: 1–55860–583–5
Printed in the United States of America

Table of Contents

Volume I

Section I — Video Surveillance and Monitoring (VSAM)

Video Surveillance and Monitoring – Principal Investigator Reports

Video Surveillance and Monitoring – Technical Papers

Event Recognition and Reliability Improvements for the Autonomous Video Surveillance System

Frank Z. Brill, Thomas J. Olson, and Christopher Tserng
Texas Instruments
P.O. Box 655303, MS 8374, Dallas, TX 75265
brill@csc.ti.com, olson@csc.ti.com, tserng@csc.ti.com

Abstract

This report describes recent progress in the development of the Autonomous Video Surveillance (AVS) system, a general-purpose system for moving object detection and event recognition. AVS analyses live video of a scene and builds a description of the activity in that scene. The recent enhancements to AVS described in this report are: (1) use of collateral information sources, (2) camera hand-off, (3) vehicle event recognition, and (4) complex-event recognition. Also described is a new segmentation and tracking technique and an evaluation of AVS performing the best-view selection task.

1. Introduction

The Autonomous Video Surveillance (AVS) system processes live video streams from surveillance cameras to automatically produce a real-time map-based display of the locations of people, objects and events in a monitored region. The system allows a user to specify alarm conditions interactively, based on the locations of people and objects in the scene, the types of objects in the scene, the events in which the people and objects are involved, and the times at which the events occur. Furthermore, the user can specify the action to take when an alarm is triggered, e.g., to generate an audio alarm or write a log file. For example, the user can specify that an audio alarm should be triggered if a person deposits a briefcase on a given table between 5:00pm and 7:00am on a weeknight. Section 2 below describes recent enhancements to the AVS system. Section 3 describes progress in improving the reliability of segmentation and tracking. Section 4 describes an experiment that quantifies the performance of the AVS "best view selection" capability.

This research was sponsored in part by the DARPA Image Understanding Program.

2. New AVS functionality

The structure and function of the AVS system is described in detail in a previous IUW paper [Olson and Brill, 1997]. The primary purpose of the current paper is to describe recent enhancements to the AVS system. These enhancements are described in four sections below: (1) collateral information sources, (2) camera hand-off, (3) vehicle event recognition, and (4) complex-event recognition.

2.1. Collateral information sources

Figure 1 shows a diagram of the AVS system. One or more "smart" cameras process the video stream to recognize events. The resulting event streams are sent to a Video Surveillance Shell (VSS), which integrates the information and displays it on a map. The VSS can also generate alarms based on the information in the event streams. In recent work, the VSS was enhanced to accept information from other sources, or "recognition devices" which can identify the objects being reported on by the cameras. For example, a camera may report that there is a person near a door. A recognition device may report that the person near the door is Joe Smith. The recognition device may be a badge reader, a keypad in which a person types their PIN, a face recognition system, or other recognition system.

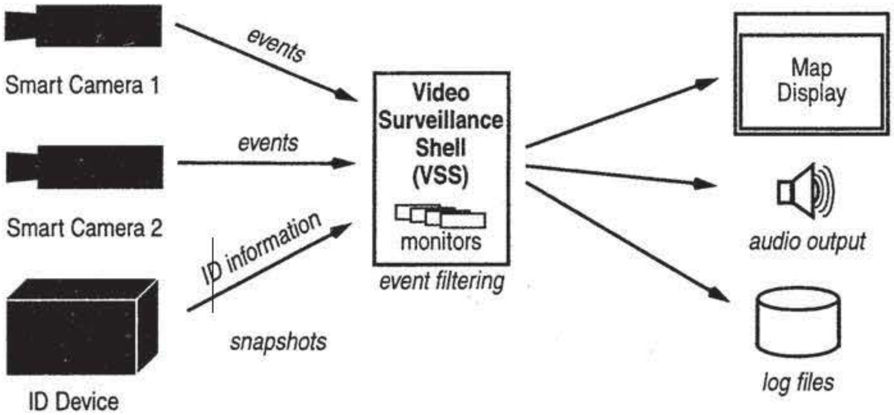

Figure 1: AVS system diagram

The recognition device we have incorporated is a voice verification system. The user stands in a predefined location in the room, and speaks his or her name. The system matches the utterance to previously captured examples of the person speaking their name, and reports to the VSS if there is a match. The VSS now knows the identity of the person being observed, and can customize alarms based on the person's identity.

A recognition device could identify things other than people, and could classify actions instead of objects. For example, the MIT Action Recognition System (MARS) recognizes actions of people in the scene, such as raising their arms or bending over. MARS is trained by observing examples of the action to be recognized and forming "temporal templates" that briefly describe the action [Davis and Bobick, 1997]. At run time, MARS observes the motion in the scene and determines when the motion matches one of the stored temporal templates. TI has obtained an evaluation copy of the MARS software and used it as an recognition device which identifies actions, and sends the result to the AVS VSS. We successfully trained MARS to recognize the actions of opening a door, and opening the drawer of a file cabinet. When MARS recognizes these actions, it sends a message to the AVS VSS, which can generate an appropriate alarm.

2.2. Camera hand-off

As depicted in Figure 1, the AVS system incorporates multiple cameras to enable surveillance of a wider area than can be monitored via a single camera. If the fields of view of these cameras are adjacent, a person can be tracked from one monitored area to another. When the person leaves the field of view of one camera and enters another, the process of maintaining the track from one camera view to another is termed *camera hand-off*. Figure 2 shows an area monitored by two cameras. Cam-

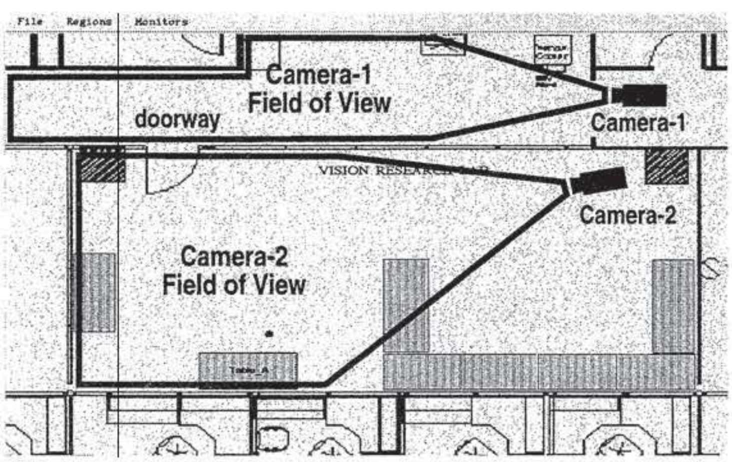

Figure 2: Multiple cameras with adjacent fields of view

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase® Smarter legal research.