

Using Video Sensor Networks to Command and Control Unmanned Ground Vehicles

Greg Kogut, Dr. Mike Blackburn, H.R. Everett
Space and Naval Warfare Systems Center San Diego
53406 Woodward Road, San Diego, CA 92152
www.spawar.navy.mil/robots
greg.kogut@navy.mil

ABSTRACT

Most unmanned ground vehicles (UGVs) used in defense and physical security applications have a limited onboard sensor range, and are only “aware” of their immediate surroundings. This limited range hinders the ability of UGVs to detect or respond to remote events, and limits their cost-effectiveness in securing large areas.

The Distributed Interactive Video Array (DIVA) system is a network of wireless, man-portable vision sensors under development at Spawar Systems Center (SSC) San Diego. DIVA provides two primary capabilities: 1) the detection, tracking, and classification of moving targets in a variety of tactical environments, and 2) the autonomous coordination of a UGV response to detected events.

The prototype network at SSC San Diego consists of three wireless, man-portable “smart” camera nodes, providing surveillance of approximately two square miles of area, including open space, roads, building clusters, and ocean. DIVA employs a communication protocol which allows it to communicate with a variety of resources, including other DIVA nodes and UGVs.

Collaborative communication between unmanned systems and stationary sensors provides a synergistic capability: the fixed sensor nodes achieve an active response capability, and the unmanned systems acquire an expanded “situational awareness,” along with a reduced response time to remotely detected events. These capabilities combine to increase the cost-effectiveness of UGVs in physical security applications.

1. Introduction

Since the birth of robotics, it has been a dream of robot builders to build a robot that can see like a human. While current state-of-the-art is still far short of human vision, most UGVs, particularly those involved in physical security or military applications, are highly dependent upon vision sensors. Vision sensors are an good sensor modality for medium to long-range sensing, because vision data is rich with information that is very easily accessible to human operators with little or no intermediate processing. The first and most common military uses of vision sensors in unmanned ground vehicles were for teleoperation of remote vehicles. Wireless transmission of visual data first allowed remote operation of unmanned vehicles over large distances, and removed human operators from dangerous situations. This capability remains the primary capability of unmanned systems in military and physical security applications. A good example of the value of teleoperation through remote transmission of video is in the tunnel and cave exploring UGVs such as the SSC San Diego Urbot [1, 2].

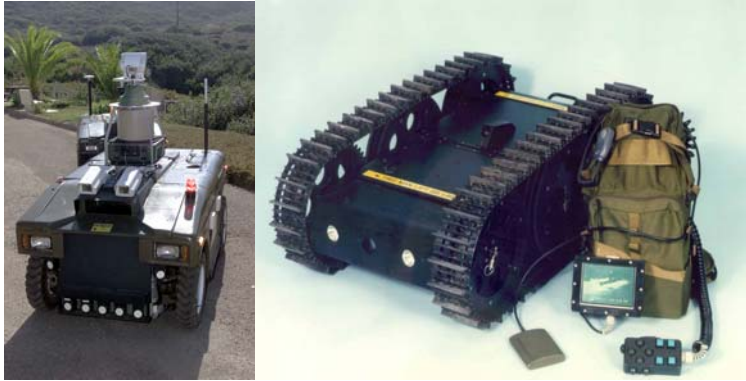


Figure 1 *Left: An SSC San Diego Mobile Detection Assessment Response System UGV, equipped with FLIR, video, and radar sensors. Right: SSC San Diego Urbot with teleoperation controls. The Urbot is equipped with four video cameras.*

More recently, advances in computer vision have allowed UGVs to achieve greater autonomy. Capabilities such as obstacle detection and avoidance and scene modeling are allowing unmanned ground vehicles to achieve greater degrees of autonomy, reducing the control burden on human operators. An example of this increased use of computer vision in many UGV tasks is DARPA's Tactical Mobile Robot program, which produced a small UGV capable of fully autonomous navigation of stairs, a task often difficult by teleoperation [3, 4].

However, there are limitations to a UGV's visual sensors. A UGV's visual sensors may give a limited perspective of the surrounding environment. For example, the range of forward-looking cameras on small robots diminishes as they approach any object taller than the height of the sensor mount. For example, common obstacles such as stairs and hills provide great challenges to small teleoperated and autonomous UGVs because of the visual perspective available from mounted sensors.

However, these limitations can be overcome when the UGV is viewed as a part of a larger network of sensors, instead of as an independent, standalone system. Surveillance camera networks are a ubiquitous component of most physical security systems. As high bandwidth, digital wireless networking systems become widely available so does the capability to make the data available from these networks available to UGVs. Making surveillance network data available to UGVs is effectively a UGV force multiplier. Access to video streams from any camera in a nearby surveillance network both provides UGVs with visual perspectives that would be impossible to acquire with onboard sensors, and with constant awareness of much larger areas than is possible with onboard sensors.

The Distributed Interactive Video Array (DIVA) project at SSC San Diego has the goal of developing a wireless network of video cameras specifically geared toward communicating with UGVs as well as providing conventional physical security functionality. A network of man-portable, self-calibrating “smart” cameras distributed across a tactical environment could provide valuable capabilities to robotic and manned forces, including: 1) the continuous tracking of targets in a global coordinate system, and 2) the command and control of robotic forces, such as the autonomous deployment of robotic platforms to intercept detected targets. These two capabilities add an automated response capability to a physical security system. In addition to this core capability, DIVA could also employ other vision capabilities, such as license plate reading, face recognition, and 3D scene modeling.

The DIVA project is currently in its first phase of development at SSC San Diego, and consists of a three-node prototype network. Each node contains an omnidirectional camera and a pan-tilt-zoom camera. Each node also contains an embedded computer, wireless networking capability, and battery/solar power subsystem. The prototype network covers approximately one square mile of area, including coverage of roads, building clusters, bunkers, open space, and ocean. The covered area also includes test ranges for unmanned ground, surface, and air vehicles, allowing for close interaction with SSC San Diego's array of unmanned vehicle programs. The existing Multiple Resource Host Architecture (MRHA) allows for simultaneous communication with multiple heterogeneous robotic platforms [5].

DIVA is designed to be a robust, lightweight visual surveillance sensor network that can be easily deployed and configured without special technical knowledge. It will perform common surveillance tasks with less manpower than conventional surveillance networks. The sensor architecture is designed to be modular and expandable, allowing easy integration of maturing vision technologies.

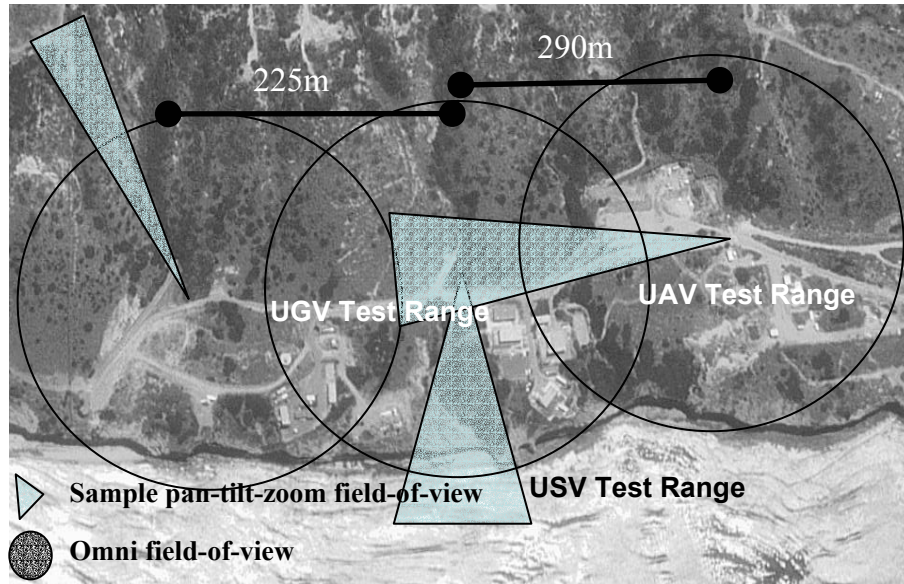


Figure 2 *Iconic overview of the area covered by the prototype DIVA sensor network, including UGV, USV, and UAV test sites.*

1. Sensor Network Hardware

Diva hardware consists of a sensor module, and a processor module. The processor module may also be used with most conventional video cameras, making it possible to give much DIVA functionality to existing video camera infrastructure.

Sensor Module

The sensor module consists of two cameras in a single lightweight, weatherproof housing. One camera is a Canon VCC4-R pan-tilt zoom camera, which provides a high quality analog video stream, and a fast pan-tilt head. The second camera is a Sony FCB-IX10A camera block fixed with an omnidirectional mirror. The combination covers a field-of-view of slightly greater than a hemisphere, and provides both constant low-resolution awareness of the entire field-of-view, and the ability to acquire high resolution views anywhere within the hemisphere.

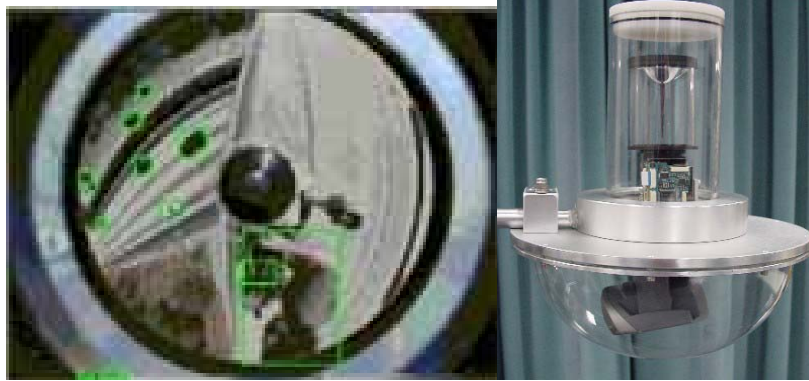


Figure 3 *The image to the left shows an omnidirectional image. Vehicles on a freeway are being tracked in the image. To the right is the DIVA sensor module showing durable aluminum and acrylic housing containing the pan-tilt zoom (bottom) and omnidirectional (top) cameras. All wiring exits the waterproof housing through the tube extending to the left.*

The omnidirectional camera provides constant 360-degree awareness throughout a hemisphere field of view. Motion detected in the omnidirectional field-of-view can be viewed in greater detail by cueing the pan-tilt-zoom camera.

Computer Module

The video and control cables from the sensor module extend to the computer module. The computer module consists of a PC/104+ computer stack housed in a ruggedized, weatherproof housing. The stack includes an 800 MHz Transmeta Crusoe processor, two video digitizers, 512MB RAM, 512MB CompactFlash disk, and a PCMCIA 802.11b wireless Network Interface Card. The computer module performs many roles, including video capture, digitization, compression, and transmission, as well as computer vision, and communication with other DIVA nodes and multiple UGVs.

Each module can simultaneously transmit two video streams at a resolution of 320x240 at 30 frames per second. Higher resolutions are possible if one of the streams is

dropped, or the frame rate is reduced. Video can be transmitted in a variety of formats, including H.263, MJPEG, and MPEG-2/4.

2. Software Architecture

The processing architecture consists of three separate, but interdependent modules.

The first module is the video compression and transmission module. This module simply provides raw video streams to any user, DIVA node, or UGV that requests the data. The video stream can be provided in a variety of formats, depending on the need of the user and the capacity of the available network connection. The modules can serve two video streams simultaneously. Requests for video are moderated by the networking modules, described below.

The second module is the computer vision module. This module performs computer vision algorithms on the raw video available from the attached sensor module. This currently implemented vision algorithms include moving object segmentation (motion detection), and tracking of moving objects. Plans include adding other algorithms such as license plate and face recognition capability.

The third module, the networking module, serves as a link to other resources available to the node. This includes UGVs, other DIVA nodes, as well as any other resource capable of communicating via the Multiple Resource Host Architecture. This module advertises, and, when necessary, distributes available raw and processed information acquired via a DIVA node. The modules can also acquire data made available by other DIVA nodes or UGVs. The networking module, for example, can

provide a UGV with the tracking coordinates of a detected intruder as reported by DIVA's computer vision module. The DIVA network is self-configuring. New DIVA nodes join the network as they are turned on, and advertise their position and information to all other nodes and resources.

It should also be noted that both the software architecture and computer processing modules can be easily used with existing video surveillance infrastructure, and don't require the specialized DIVA sensor modules. This allows adding much of DIVA's functionality to the enormous existing video surveillance infrastructure used in physical security systems across the world.

3. Camera Calibration

DIVA requires that individual sensor nodes both know where they are in a global coordinate system, and are also able to report the global position of anything within their field of view. For example, if two adjacent sensors "see" the same UGV from different locations, they should both report the same global position for the UGV, even though it may appear different in each field of view.

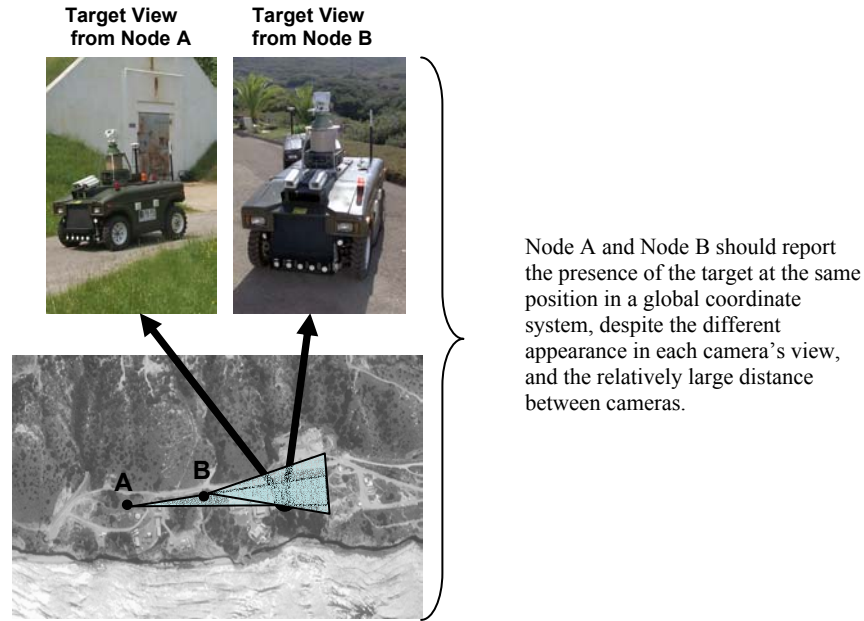


Figure 4 *DIVA will have the ability to consistently report the global coordinates of detected targets, a valuable capability of a physical security system.*

This ability requires that DIVA's cameras be calibrated. Most camera calibration schemes are both labor intensive and require the use of special props. DIVA however, is required to be easy to set-up, and use, minimizing the labor needed for its operation in the field. Therefore a new automated calibration scheme is employed, based on a calibration system for outdoor cameras developed at Carnegie Mellon University (CMU) [6]. The CMU method relies on a user being able to located landmarks of known global position in a camera's field of view, and then click on the landmarks with a mouse pointer to "tell" the system where the landmarks are. The DIVA calibration system takes the human out of the calibration loop by using UGVs instead of landmarks, and a UGV's GPS sensor to report the UGV's position to the DIVA computer. The usefulness and accuracy of UGV-mounted GPS sensors has been research and reported by SSC San

Diego [7]. This automated calibration method has several advantages over manual calibration methods: it is faster, doesn't require prior of scene geometry, requires no special props, and is potentially more accurate than landmark-based calibration schemes.

While the DIVA calibration system is still undergoing development, the goal of the system is to provide target position with worse-case accuracy of less than 3m error. This degree of worst-case error will allow DIVA to guide a UGV close enough to a target that the UGV's own sensors will be able to acquire a target more accurately.

4. Applications

Physical Security

The first demonstration of DIVA's capability will be to assist the MDARS (Figure 1) [8]. The MDARS program is charged with providing physical security to Department of Defense facilities and storage sites. A test-bed for MDARS development exists on site at SSC San Diego (Figure 2), and overlaps the DIVA prototype network.

The MDARS system works by sending MDARS-E UGVs on random patrols. An Intrusion Detection System (IDS) mounted on the MDARS UGV is used to detect intruders. In the event of an intrusion, an alarm is sent to an Operator Station which is capable of monitoring and controlling up to 255 UGVs and unmanned sensors. A human operator then directs a response.

DIVA closes the loop between the surveillance network and the MDARS UGV by automatically detecting events, and communicating directly with MDARS. While events would also be reported to the Operator Station, no manual intervention would be needed to produce a response from MDARS.

An example scenario would be:

- 1) A DIVA node detects an unidentified moving vehicle.
- 2) DIVA reports the event to an MDARS-E vehicle located 400m away, and directs the vehicle to respond to the intruder.
- 3) While the MDARS UGV moves towards the intruding vehicle, DIVA constantly provides target locations to MDARS so that an efficient path plan can be followed.
- 4) DIVA directs MDARS close enough to the target so that MDARS' own sensors can take over.

There are two primary advantages to this approach: reduced cost and decreased response time. While a DIVA network is more expensive than a conventional CCTV network, a DIVA network requires much less manpower to monitor, reducing labor costs which tend to be a large component of the costs of physical security. DIVA could also reduce the cost of a UGV-patrolled site by allowing fewer UGVs to cover a given area. The quality of the physical security provided would also be increased by reducing the response time to detected intrusions or events.

Force Protection

A DIVA network can also be set up in areas where no existing surveillance or UGV infrastructure exists. This allows the construction of temporary, ad-hoc surveillance zones, perhaps to protect temporary camps. Functionality could be similar to that provided by DARPA's Combat Zones That See program [9]. In addition to providing video surveillance capability, each DIVA node is also an independent digital networking access point, capable of acting as a network relay for any type of digital traffic. DIVA nodes could extend the range of a UGV by extending the range of the wireless communication. This could allow DIVA to be a part of the future vision of a C4ISR network (command, control, communications, computers, intelligence, and

reconnaissance) network which integrates battlespace resources with a common, pervasive, and secure digital network.

5. Conclusion

DIVA improves the quality and decreases the cost of physical security systems by reducing the manpower and response necessary to detect, assess, and respond to intruders. DIVA interacts closely with UGVs to remove human intervention in the process of assessing and responding to detected events. This interaction between fixed sensors and UGVs greatly increases the capabilities of both systems. DIVA is inexpensive, and easily transported and set-up, consisting almost entirely of lightweight, off-the-shelf components.

6. References

1. Laird, R.T., Bruch, M.H., West, M.B., Ciccimaro, D.A., and H.R. Everett, "Issues in Vehicle Teleoperation for Tunnel and Sewer Reconnaissance," Proceedings, Workshop 7: Vehicle Teleoperation Interfaces, *IEEE International Conference on Robotics and Automation, ICRA2000*, San Francisco, CA, 28 April, 2000.
2. Bruch, M.H., Laird, R.T., and H.R. Everett, "Challenges for deploying man-portable robots into hostile environments," *SPIE Proc. 4195: Mobile Robots XV*, Boston, MA, November 5-8, 2000.
3. *A Portable, Autonomous, Urban Reconnaissance Robot*
L. Matthies, Y. Xiong, R. Hogg, D. Zhu, A. Rankin, B. Kennedy, M. Hebert, R. Maclachlan, C. Won, T. Frost, G. Sukhatme, M. McHenry, S. Goldberg
The 6th International Conference on Intelligent Autonomous Systems
Venice, Italy, July 2000
4. *Vision-Guided Autonomous Stair Climbing*, Y. Xiong, L. Matthies
IEEE International Conference on Robotics and Automation
San Francisco, California, April 2000
5. Everett, H.R., Laird, R.T., Carroll, D.M., Gilbreath, G., Heath-Pastore, T.A., Inderieden, R.S., Tran, T., Grant, K., and D.M. Jaffee, "Multiple Resource Host Architecture (MRHA) for the Mobile Detection Assessment Response System (MDARS)," *Technical Document 3026, Revision A*, Space and Naval Warfare Systems Center, San Diego, CA, September 2000

6. Collins and Tsin, "Calibration of an Outdoor Active Camera System" *IEEE Computer Vision and Pattern Recognition (CVPR 99)*, Fort Collins, CO, June 23-25, 1999, pp. 528-534
7. Bruch, M.H., Gilbreath, G.A., Muelhauser, J.W., and J.Q. Lum, "Accurate Waypoint Navigation Using Non-differential GPS," AUVSI Unmanned Systems 2002, Lake Buena Vista, FL, July 9-11, 2002
8. Carroll, D., Gilbreath, G.A., and H.R. Everett, "Extending Mobile Security Robots to Force Protection Missions," AUVSI Unmanned Systems 2002, Lake Buena Vista, FL, July 9-11, 2002
9. <http://dtsn.darpa.mil/ixo/solicitations/CTS/index.htm>