The ATON Infrastructure

Mohan Trivedi, Greg Kogut, Brett Hall, Kim Ng, Ivana Mikic

Computer Vision and Robotics Research Laboratory, UCSD

1. System Specifications

The ambitious goals of the ATON project require the real-time processing of large volumes of data from multiple sensor modalities, as well as from multiple sensor clusters. This processing includes algorithms for communication and coordination among nodes in sensor networks, as well as the algorithms for monitoring and analyzing traffic.

Such processing will require the tuning of known algorithms for traffic scenes, as well as the development of new algorithms. This development requires the availability of the types of data that will be used during later implementations of the ATON project. There is no current infrastructure that provides this type of data. Current traffic monitoring systems typically include single views from single rectilinear cameras. Also, the data cannot be accurately simulated, since no simulation could accurately model the complexity of the proposed environment or the unpredictable content of a traffic scene. There is a necessity for the development of a testbed that can provide the necessary data. A non-trivial amount of research will be involved in developing and fine-tuning such a testbed. The unique, dynamic infrastructure proposed in the ATON project is an integral part of the development of the system as a whole.

Having a local infrastructure will provide easy availability to the necessary data. This increases the efficiency of research, and provides valuable firsthand experience with the development of all aspects of the ATON project. The design philosophy of the implementation of the ATON testbed includes incorporating modularity, scalability, and generality. These ideas will allow design changes, evolution of the design based on new hardware or algorithms, and possible application to many other applications or environments.

2. Architecture

The ATON hardware infrastructure is based upon the sensor cluster. This distinguishes the ATON infrastructure apart from previous efforts in intelligent transportation research, which have been largely based upon single sensors, and single types of sensors. By combining multiple sensors and multiple sensor modalities at each sensor sites, there is more information available to resolve many of the classical problems in computer vision applications. For example, fusing information from multiple sensor modalities can help with resolving occlusion and shadow detection.

The types of sensors used in the ATON infrastructure include standard rectilinear cameras, omniview (ODVS) cameras, and pan-tilt cameras. Other types of sensors being researched include infrared and audio sensors. A typical sensor cluster includes an ODVS camera, providing a medium resolution view of the surrounding traffic scene. The video from an ODVS camera alone is sufficient for many traffic applications, including segmenting moving cars, tracking, and pattern recognition applications. The sensor cluster also supports a high-speed pan-tilt camera, capable of providing high-resolution views within the scene. The pan-tilt camera is capable of providing views of license plates, or providing a close-up view of debris on the highway.

The volume of data produced by having a network of sensor clusters, each with two or more sensors, is enormous. To date, little research has been done involving such a large volume of sensor data, with multiple views and locations. This makes the coordination and effective design of the infrastructure a rich area of research in itself. Research issues include:

- Developing effective sensor hand-off techniques
- Researching the capabilities of sensor fusion among a cluster's sensors
- Communicating information among sensor clusters
- Transmitting data to the Base Unit and Visualization Units

Though there can be some local processing at each sensor clusters, the bulk of processing will occur at the Base Unit, a central processing location containing the hardware necessary for video digitization,

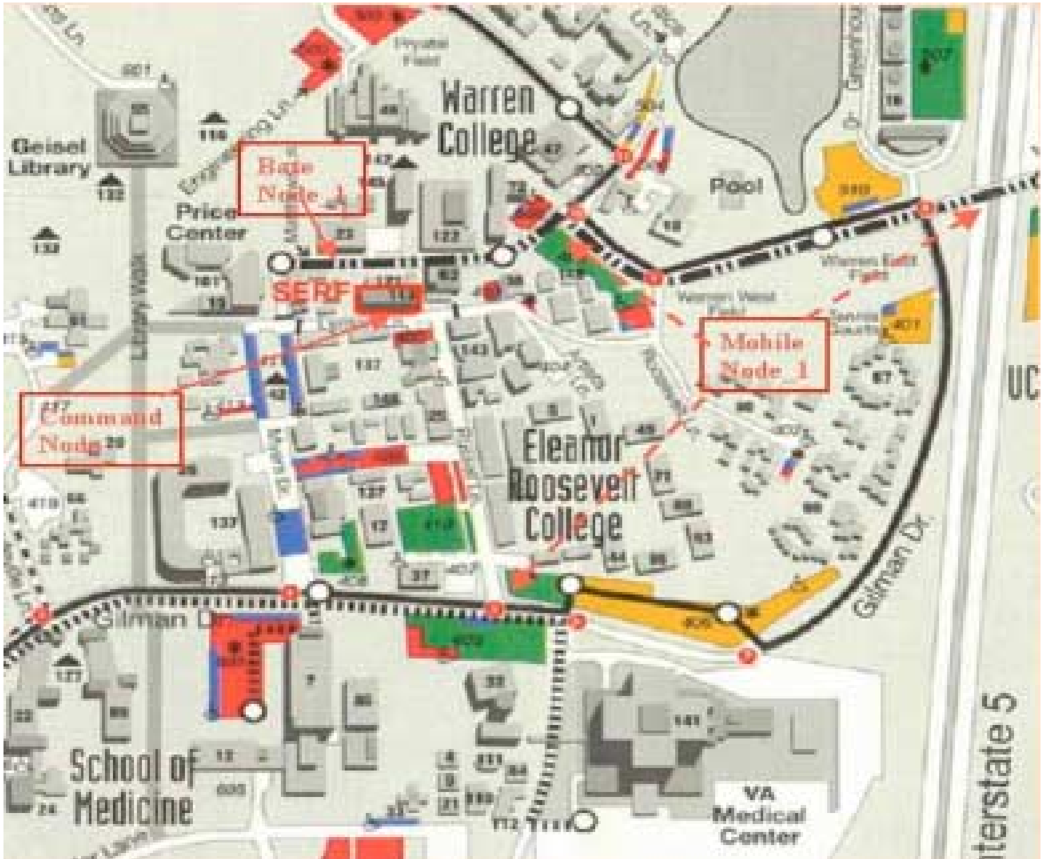

Figure 1Location of the Base Node, Command Node, and possible location for the Mobile Nodes on the UCSD main campus.

processing, storage, and distribution. The Base Unit also maintains a database that will coordinate information. The Base Unit also provides control information to each sensor. For example, the Base Unit will command the movement of each pan-tilt camera. An effective communication network is necessary to provide this two-way communication between the sensor clusters and the Base Unit.

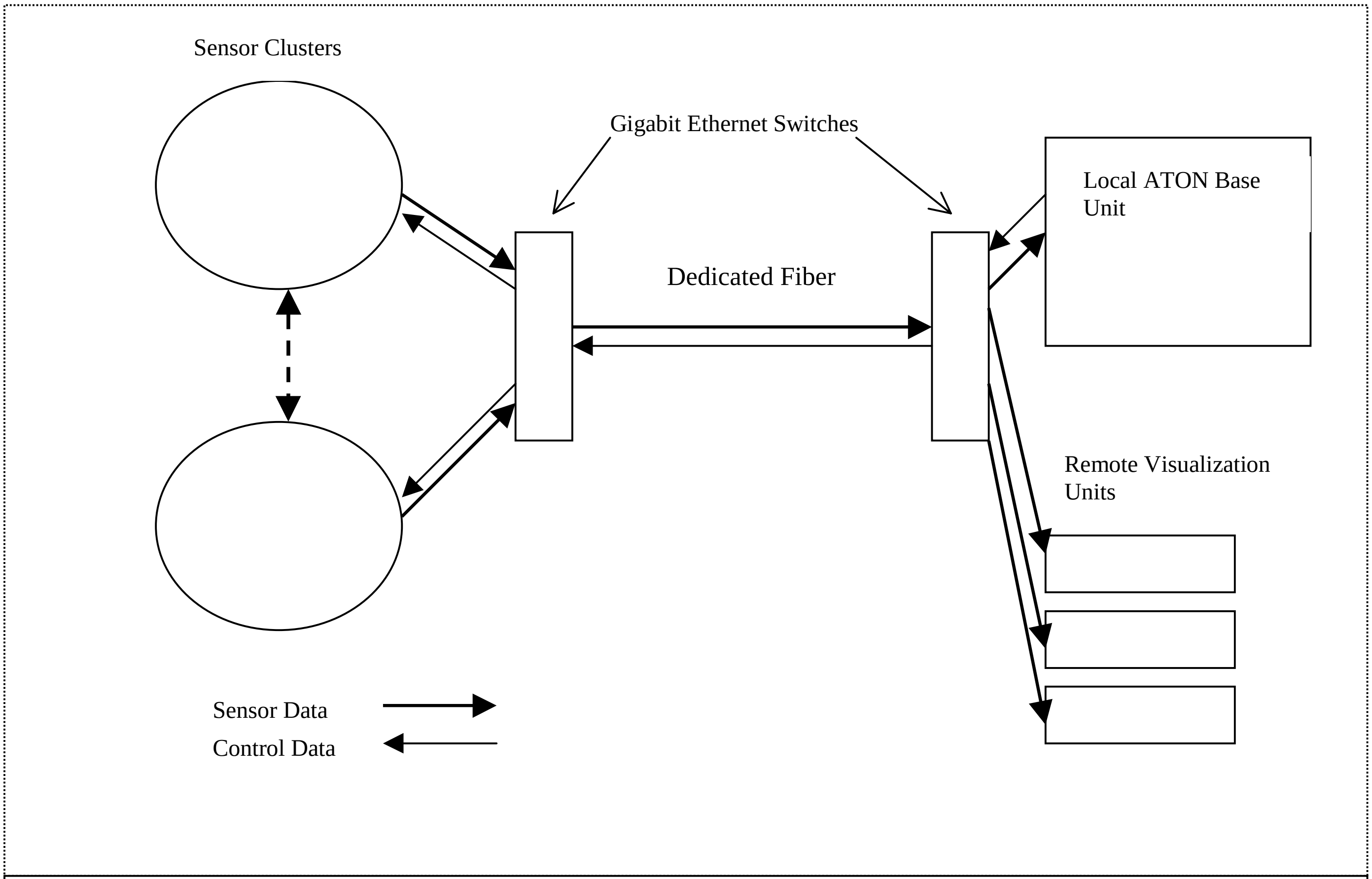

Figure 2 Data flow diagram for the ATON Infrastructure, highlighting the scalability and modularity of the infrastructure design

3. Video Network

3.1. Rectilinear Cameras

Each sensor cluster contains one Pelco Spectra Lite™ integrated pan-tilt camera system. Each camera is capable of 360 degrees of pan rotation and 16X optical zoom. The pan-tilt mechanism may be remotely controlled, and contains variable speed motors to allow tracking of moving objects. The camera can pan at up to 80 degrees per second, and scan at up to 40 degrees per second. The camera is sensitive to light down to 1 lux. The camera and receiver are self-contained in a small, discreet dome that may be mounted in a variety of configurations including poles, walls, and rooftops.

Figure 3 A Pelco Pan-Tilt-Zoom Integrated Positioning Camera

3.2. ODVS

Each sensor cluster will include at least one omni-directional visual sensor (ODVS. The ODVS consists of a conventional imaging system in combination with a mirror. The mirror is used to project a 360-degree view of the surrounding area onto the image plane. A conventional perspective view can be recovered from the projection. A wider range of area may be covered by relatively few ODVS than with other types of sensors. ODVS can also be used for N-ocular vision, and they can self-localize by locating other ODVS within a scene.

Figure 4A 360-degree ODI along with its "unwrapped" perspective view image

Omni-directional images (ODI) can provide 360 degree virtual walkthroughs of an environment, as well as a rich visual information useful in many vision-based applications. They also provide a constant real-time view of a large area without mechanical movement, making them ideal for surveillance applications or real-time scene analysis.

The ODVS used in the ATON project are based on hyperboloidal mirrors, which allow for easy transformation of ODI into perspective images, and are more cost-effective than other types of mirrors, not requiring special lenses.

The ODVS will be used in the ATON infrastructure both to provide medium-resolution visual data for wide area, and to guide the movement of the high-resolution pan-tilt-zoom rectilinear camera which can be used to focus on specific activities occurring within a scene.

(a) C-ODVS with a hyperboloidal mirror (b) Ultra C-ODVS with a spherical mirror

Figure 5 Several types of ODVS. The sensors on the left are complete with mirror and imaging system, while the right group shows only the mirrors.

3.3. Video Capture

Raw video from the sensor clusters is received in the lab as a set of standard NTSC signals. These signals may be digitized in full, or combined with a video matrix unit. The video matrix unit can combine up to four NTSC signals into a single signal by reducing the resolution of each incoming signal by half, to 320x240. The currently used video matrix unit is the Robot MV87 Color Quad.

Video capture takes place on Matrox Meteor II video digitizer cards installed in dual Pentium III Windows PCs. Each matrox card is capable of capturing one full stream of NTSC video at 30 fps. Each dual-Pentium PC contains two Matrox cards, enabling the digitization of two full NTSC streams, or up to 8 half-resolution streams. Once digitized, the video can either be stored on local disk, or passed on to the real-time video processing architecture.

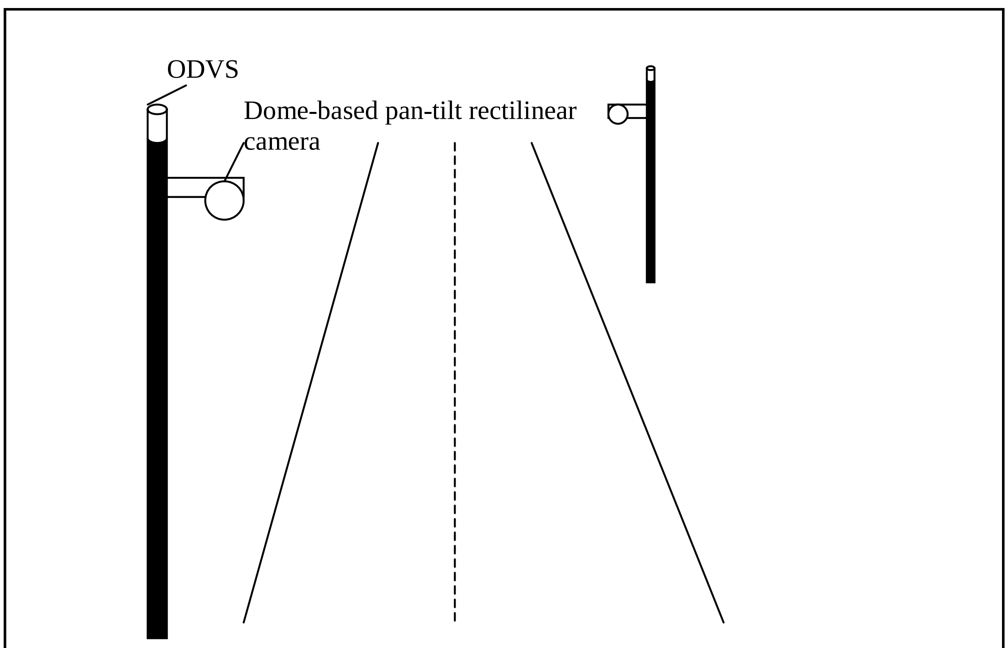

Figure 6 Two sensor clusters covering a section of highway

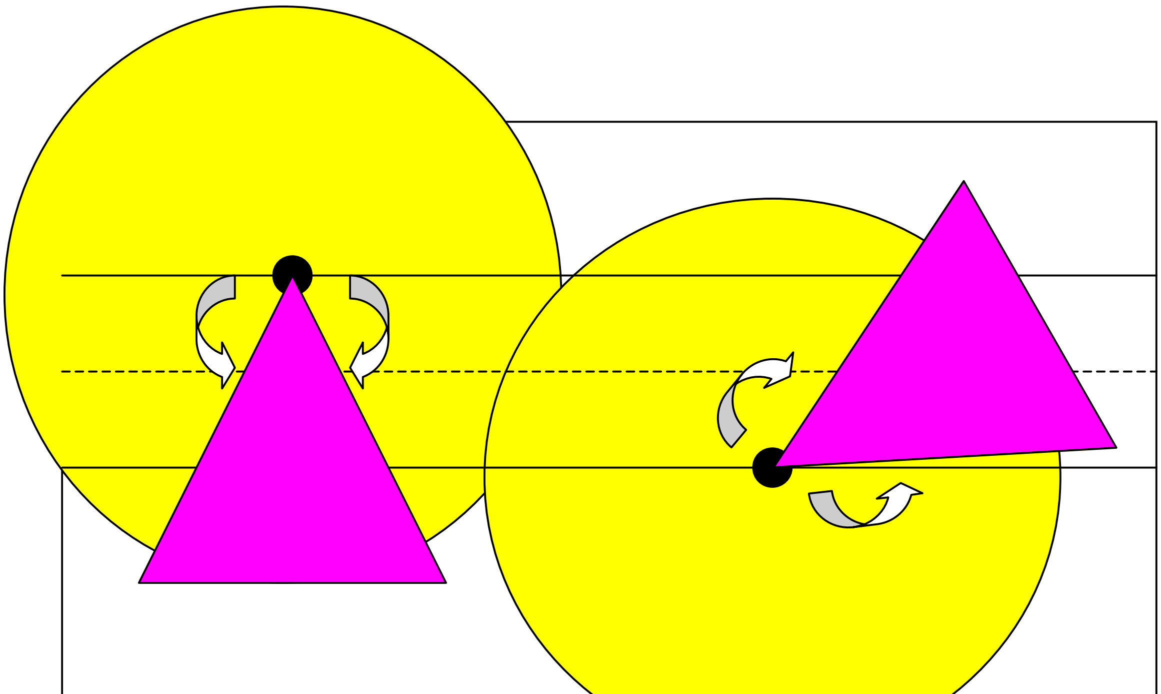

Figure 7Illustration of the view range of the ODVS camera overlaid with that of a conventional rectilinear PTZ camera

4. Communication System

The analog NTSC signals from the ODVS and rectilinear cameras must be digitized and packetized before being sent over the digital communications network. Digitization is preformed using AXIS 2400 video servers. A single AXIS 2400 video server serves one full NTSC video stream over a network using the TCP/IP protocol. The server employs motion-JPEG compression with an ARTPEC-1 compression chip, which allows user-selectable compression levels.

Data is transmitted to the Base Unit using a Gigabit Ethernet line. All data information from each sensor cluster is multiplexed then sent over a fiber-optic network via a Cisco Catalyst 3500 XL Series Gigabit

Ethernet Switch. This makes the data available to any Internet-capable computer, provided the necessary bandwidth is available between the sensor cluster and the receiving computer. It is also possible to provide lower-quality data with much lower bandwidth requirements by using commonly available video and audio compression techniques. However, the necessity for high quality data during the research phase of the ATON project dictates the need for a high-speed local network.

5. Mobile Unit

To support the ATON project goals, significant amounts of video data need to be taken from a variety of situations at different times on highways, around people & cars, in town congestion, etc. As a piece of the ATON project infrastructure, we determined there was a need for a mobile unit with the capability of taking video data from heights similar to roadside streetlights to enable us to acquire the necessary data on demand *easily*. The resulting system has been quite successful in achieving it's basic goals, however future expansion on the unit will ensure it's capability will fully match the requirements of the ATON project as it progresses to the next stage.

The established mobile system is comprised of a 12' aluminum pole that is mounted on an electric golf-cart and stabilized using chains attached to the cart frame. A camera mount is attached on top for attachment of the omni-directional video sensor (ODVS) and rectilinear camera. There is a rugged PC onboard the cart that will utilize an external array of fast hard drives to store the video data. An inverter that draws power from the main cart batteries powers this PC and any other necessary peripherals.

Camera support hardware:

The resulting system has a 12', 2-3/4"OD aluminum pole mounted on the back of an electric golf-cart. The pole is attached to the cart whenever it is needed, it can be detached and stored elsewhere. The bottom end of the pole has a ~1" OD solid brass rod bolted into it for mounting purposes. The end of this rod extends from the pole a couple of inches, enough to fit snugly into a universal joint attached to the cart. This bottom mounting using a universal joint allows for angular freedom of the pole enabling it to be leveled/adjusted for a given terrain. When erected, the pole is stabilized by 3 chains fastened to hook eyes installed in the pole approximately ~120 degrees from eachother a few feet up from the universal joint. The other ends of the chains are attached to the steel frame around the back of the flatbed of the EZ-GO cart. The use of a "fine adjustment" device attached to the end of the chains allows for more precise adjustment of the amount of tension & length of the chains. On the top of the pole the camera mount is attached by bolting an aluminum cylinder of slightly smaller diameter than the ID of the pole to the mount. This mount is then set in place with 6 machine screws in two layers. The resulting fixture for this is quite strong and sturdy. Since the aluminum block is somewhat soft relative to the steel screws, the screws deform it somewhat when tightened – resulting in a stronger mount.

Cameras:

The camera located at the top of the pole currently is a SONY handicam with attached ODVS. This handicam is mounted to the camera mount bolted to the pole, and oriented such that the ODVS is vertical. The video signals can be recorded by the camcorder for simple operations, or exported down to the PC on the cart. This configuration will be expanded soon into two mounted ODVS sensors with their own camera, and a pan/tilt/zoom rectilinear camera for high resolution shots.

Mobile unit power:
The industrial flatbed EZ-GO electric golf-cart is powered by six 6V lead-acid rechargeable batteries in series that drive the rear wheels with a powerful 36V electric motor. It has 2 seats in the front, headlights, taillights, and a storage area underneath the front of the wood topped flatbed. These same batteries also allow us to operate the PC and video equipment while out at the remote site. A small 36V inverter that is wired to the batteries directly provides the power for the computer and other necessary peripherals. As a safety precaution, we will soon install a cut-off switch for the cart drive system (i.e. motor) when the computer is attached and running to avoid the costly mistake of creating voltage spikes/drops if the accelerator is accidentally hit during an acquisition.

On-board computer:

The PC system which will do the video capture contains a dual-Pentium III 700MHz system with 256M Ram, SuperMicro P6DBU motherboard (onboard Ultra2Wide SCSI bus and RAID capable), ATI all-in-wonder 128 32M video board, Matrox Millenia video capture card, 13GB system hard drive, CDROM, Floppy. Up to 6 external 18 GB Seagate cheetah 10000-RPM U2W SCSI hard drives mounted in a portable chassis are used to store the video data. These can be configured in a RAID format if necessary, but we predict that striping the data across multiple hard drives should allow us to achieve the necessary transfer rates to capture high quality video in real time. This PC is housed in a Trans 2000 “Titan-T9” ruggedized case, which will ensure vibration resistance and a solid industrial package for use out in the field. It contains a 15” LCD screen, waterproof keyboard, and built in mouse. Windows NT is run on the PC, although we are considering upgrading Windows 2000 professional if drivers are available for all our hardware.

Near-future enhancements:
- Pan/Tilt camera & mount for higher resolution rectilinear images
- Implementing software which aids in leveling the ODVS image at acquisition location
- Lifts for holding the vehicle off it’s normal suspension for increased stability during acquisition
- Wireless communication capability for live video transfer back to the CVRR lab (range is critical)
- Cart-mounted GPS unit for accurate position information and path determination.

6. Base Unit, Servers, Computing Resources, and Interfaces

All the video information from the sensor cluster is received at the Base Unit. The Base Unit consists of the computing hardware necessary for digitizing, processing, storing, and distributing all the information generated by the ATON sensors.

As described above, the video digitizing computers consists of dual-Pentium III PCs containing two Matrox Meteor II video capture cards. Similar machines will provide services for audio digitization. These machines are also capable of archiving segments of raw digital video onto high-speed local disk arrays.

From here, the data can be passed onto other machines via a distributed software architecture. Other “layers” in the information processing pipeline include such things as segmentation, tracking, pattern recognition, and behavior recognition. The distributed nature of the architecture allows individual processing steps and research issues to be encapsulated, so that changes in the design decisions of one area minimally effect other areas. All Base Unit computers are linked via a high-speed network to the real-time transfer of high-bandwidth video, audio, and processed data.

The distributed software infrastructure requires the need for a flexible communication interface, allowing for changes in design decisions, and the scaling of the project size, such as the addition of more sensor clusters or data from remote sites. The dependence upon single platforms or compression algorithms should be minimized.

7. Summary: The nature of experimental investigations; expandability

The ATON infrastructure is an active research testbed, with many areas of the design subject to changes in design decision as investigations into the nature of large-scale, real-time data collection and analysis progress. There is no currently operational real-time computer vision system of the nature and scale of a fully implemented final vision for the infrastructure, involving multiple locations, sensor modalities, and intensive use of sensor fusion.

Therefore, the early design decisions are directed towards scalability and modularity to allow for substantial modification in design decisions and expansion of the local testbed to arbitrary size. For example, new types of sensor might become available which may contribute significantly to the accuracy of analysis. Also, the basic design must be able handle hundreds or thousands of sensor clusters.

The infrastructure is also designed with generality, and need not be restricted to traffic applications. Many applications, indoors or outdoors, might benefit from the ideas incorporated into the design. Examples include remote learning, advanced video-conferencing, sports event viewing, and even planetary exploration.