REDUCED BANDWIDTH VIDEO FOR REMOTE VEHICLE OPERATIONS*

Timothy E. Noell and Frederick W. DePiero
Robotics & Process Systems Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, Tennessee 37831-6304
Telephone 615-576-3640
Facsimile 615-576-2081

Submitted to
AUVS-93 Unmanned Systems: Expanding Government and Commercial Applications
Washington, D.C.
June 28-30, 1993

*Research sponsored by the Army Research Laboratory, Office of Technology Development, and the U.S. Department of Energy under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

REDUCED BANDWIDTH VIDEO FOR REMOTE VEHICLE OPERATIONS*

Timothy E. Noell and Frederick W. DePiero

Robotics & Process Systems Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, Tennessee 37831-6304

ABSTRACT

Oak Ridge National Laboratory staff have developed a video compression system for low-bandwidth remote operations. The objective is to provide real-time video at data rates comparable to available tactical radio links, typically 16 to 64 thousand bits per second (kbps), while maintaining sufficient quality to achieve mission objectives. The system supports both continuous lossy transmission of black and white (gray scale) video for remote driving and progressive lossless transmission of black and white images for remote automatic target acquisition. The average data rate of the resulting bit stream is 64 kbps. This system has been demonstrated to provide video of sufficient quality to allow remote driving of a High-Mobility Multipurpose Wheeled Vehicle at speeds up to 15 mph (24.1 kph) on a moguled dirt track. The nominal driving configuration provides a frame rate of 4 Hz, a compression per frame of 125:1, and a resulting latency of ~1s. This paper reviews the system approach and implementation, and further describes some of our experiences when using the system to support remote driving.

BACKGROUND

The use of untethered teleoperated vehicles for many remote operations is greatly limited because of a need to use low-bandwidth communication links. For example, tactical operations that require a low signature give rise to low-bandwidth requirements. Low-bandwidth channels are also encountered in underwater operations and in space applications. The most notable difficulty in using low-bandwidth channels for vehicle control and mission package support is the problem of video transmissions from the vehicle to the driver's station. Namely, standard black and white video requires

*Research sponsored by the Army Research Laboratory, Office of Technology Development, and the U.S. Department of Energy under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

~60 million bits per second. Currently available tactical communication links support only 16 to 64 thousand bits per second (kbps). This equates to a minimum video compression requirement of ~1000:1 for remote operations via these types of low-bandwidth channels.

SUMMARY OF APPROACH

The Oak Ridge National Laboratory (ORNL) system [1] achieves video compression by combining a multi-resolution decomposition (a Laplacian pyramid), simplified motion estimation, scalar quantization, foveation, zero-run-length coding, Huffman entropy coding, and frame rate reduction. A more complete description and discussion of this approach can be found in reference [2].

The Laplacian pyramid approach to image coding [3] decomposes an image into a set of subbands, each containing a separate spatial frequency band. The generation of subbands is a process of recursive bandpass filtering A uniform quantizer is then applied to each pixel of each subband. Subbands can be quantized more or less aggressively to provide a frequency selective distribution of the loss of image information.

Foveation [4] provides a region of higher image quality in the resulting image. This is achieved by simply masking data from each subband as desired. The areas where all subbands are included is the area of highest quality for a given set of quantizers. Various effects can be achieved with different subsets of the subbands present in any given area of the image. In general, the system is set up to provide the highest quality in the center of the field of view, with decreasing resolution at points removed from the center region. The system also provides dynamic adjustments that allow an operator to position the foveal area on the output image. This effectively allows the operator to allocate bandwidth to the operationally significant areas of the scene. Figure 1 shows a quantized and foveated Laplacian pyramid together with a compressed and an original image. This image is taken from a 64-kbps compressed video bit stream during a demonstration of the system at the U.S. Army's Churchville test facility near Churchville, Maryland, in May of 1992.

Simple motion estimation is implemented by predicting the next instance of a subband with the current instance. When the next subband is calculated, the prediction error is determined by subtraction, and only the error is transmitted. The decompression algorithm does this same prediction and adds the error to reconstruct the subband. The system also implements a periodic transmission of an intact subband that serves to ensure a temporal limitation of channel errors. Experiments were performed to determine which subbands would provide a net compression gain if predicted in this manner and to determine a suitable refresh period. The final system used subband prediction on only the two lowest frequency bands because their image content varied slowly enough to make compression gains realizable. A four-frame interval was chosen so that the temporal duration of a channel error was limited to 1 s (given ~4-Hz nominal frame rate).

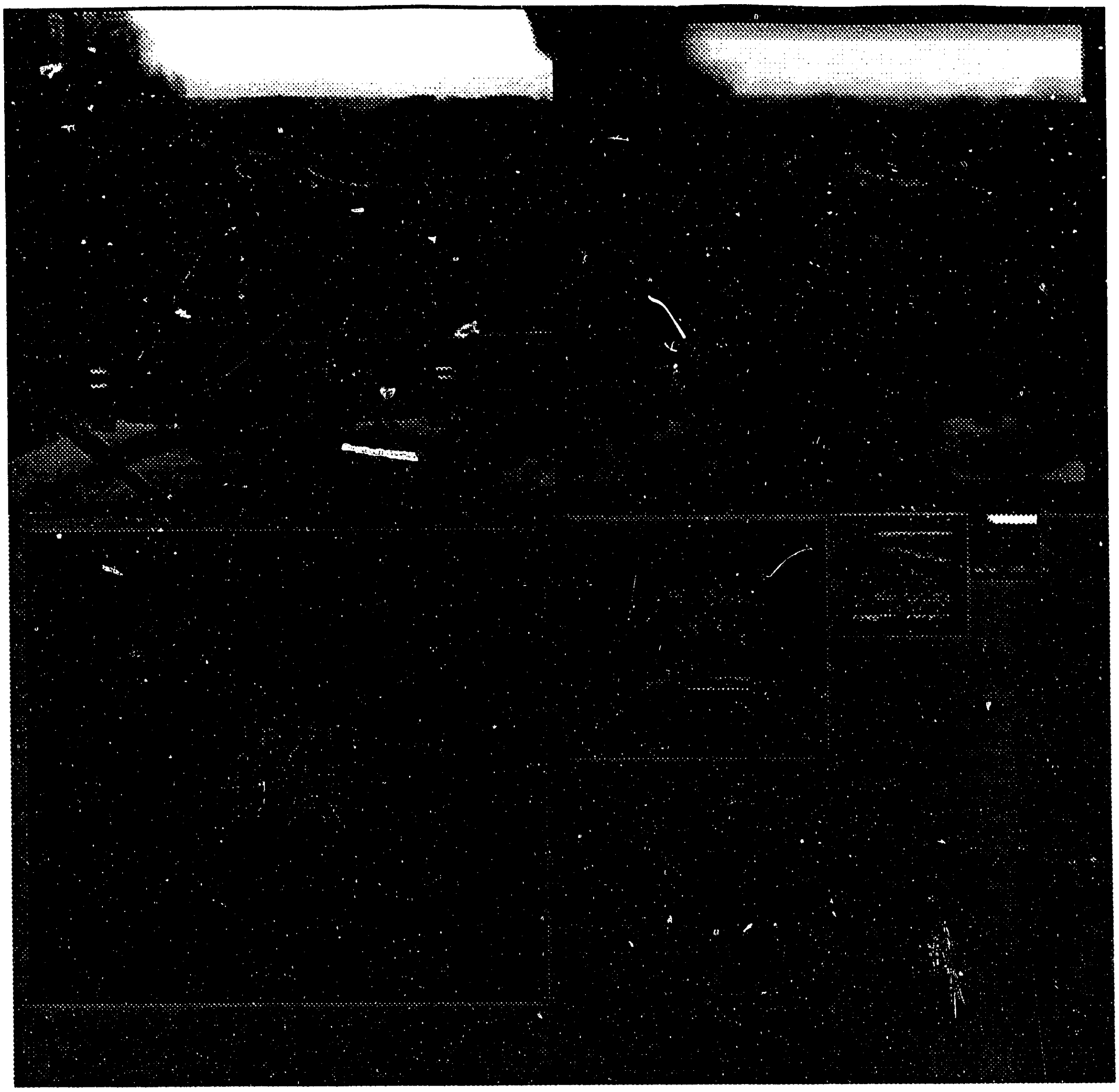

Fig. 1. An original image (upper left), a compressed image (upper right), and a quantized and foveated Laplacian pyramid.

Finally, the quantized and foveated pyramid is zero-run-length encoded and then compressed with a Huffman entropy coder [5]. This maps the variable-sized portions of the subbands into a bit stream of variable-length code words. The Huffman code word assignment is determined by an a priori analysis of representative driving imagery.

The decompression side of the system decodes the bit stream, adds the prediction errors (if appropriate), and paints the reconstructed pyramid into a buffer. The pyramid is then "collapsed" by a series of expand and add operations [3]

IMPLEMENTATION

The implementation of the ORNL system consists of two VersaModule European (VME) racks, one for compression and one for decompression. The compression rack is mounted on board a High-Mobility Multipurpose Wheeled Vehicle (HMMWV) outfitted for teleoperation. The decompression rack is mounted in an environmental enclosure adjacent to a VME-based Sparc station. An operator interface runs on the Sparc, providing vehicle control functions. The driving station and teleoperated HMMWV were developed by the Information Processing Branch of the Army Research Laboratory (ARL), Adelphi, Maryland [6].

Figure 2 shows the basic hardware components in the overall system. Six VME single-board computers are arranged in a pipelined architecture, three in the compression rack, and three in the decompression rack. Each rack also contains Datacube image processing hardware. The first processor on board the vehicle (CPU 1) is responsible for controlling a Datacube image-processing board that provides real-time generation of the quantized and foveated Laplacian pyramids. CPU 2 performs the zero-run-length and Huffman coding. The third processor interfaces to a packet radio. The processors in the decompression rack perform symmetrical functions. A more detailed description of the system implementation can be found in reference [2].

The packet radios used are the ARLAN 620 Ethernet radios from Telesystems, Inc., Toronto, Canada. These radios use a spread-spectrum type of modulation and operate in the 902- to 928-MHz band. The units provide a wireless Ethernet bridge between the VME systems. The low-bandwidth communication channel is emulated by using a real-time clock to maintain an average data rate of 64 kbps. The channel's data rate is also monitored at the receiving unit for verification.

When in driving mode with a 64-kbps channel, this implementation provides a nominal performance of ~125:1 compression per frame, 4 frames per second, and a processing latency of ~1 s (roughly 0.4 s for compression processing, 0.3 s for low-bandwidth transmission, and 0.3 s for decompression processing). This performance has demonstrated capabilities of supporting remote operation of the HMMWV at speeds up to 15 mph (24.1 kph) on a moguled dirt track.

The ORNL compression system also provides progressive transmission of still images and subimages in support of an Automatic Target Acquisition (ATA) mission package on board the HMMWV. The ATA system was developed also by the Information Processing Branch of the ARL [6]. At the start of targeting operations, the ATA system passes to the compression rack seven images that cover the area to be monitored. These are transmitted at full quality and stored in video buffers in the decompression rack. When the ATA system acquires targets, it assigns a target number and passes the image, along with the target's pixel coordinates and size, to the compression system. The compression system then transmits only the small rectangular portions of the images containing the tracked targets, which are typically ~40 x 60 pixels. At the operator station, the decompression system pastes targets at the appropriate locations within the appropriate

frame buffers. The ATA system keeps the compression system informed as to the pixel location of a given target. As the target moves, the decompression rack pastes in the background at the old location and pastes the updated target image at the new location. The contents of the tracking buffers are displayed on the operator control station at the operator's request. A socket-based custom protocol is used to communicate among the ORNL compression system, the operator control station, and the ATA mission package. The system has been demonstrated to be capable of supporting remote ATA at 64 kbps by using this approach.

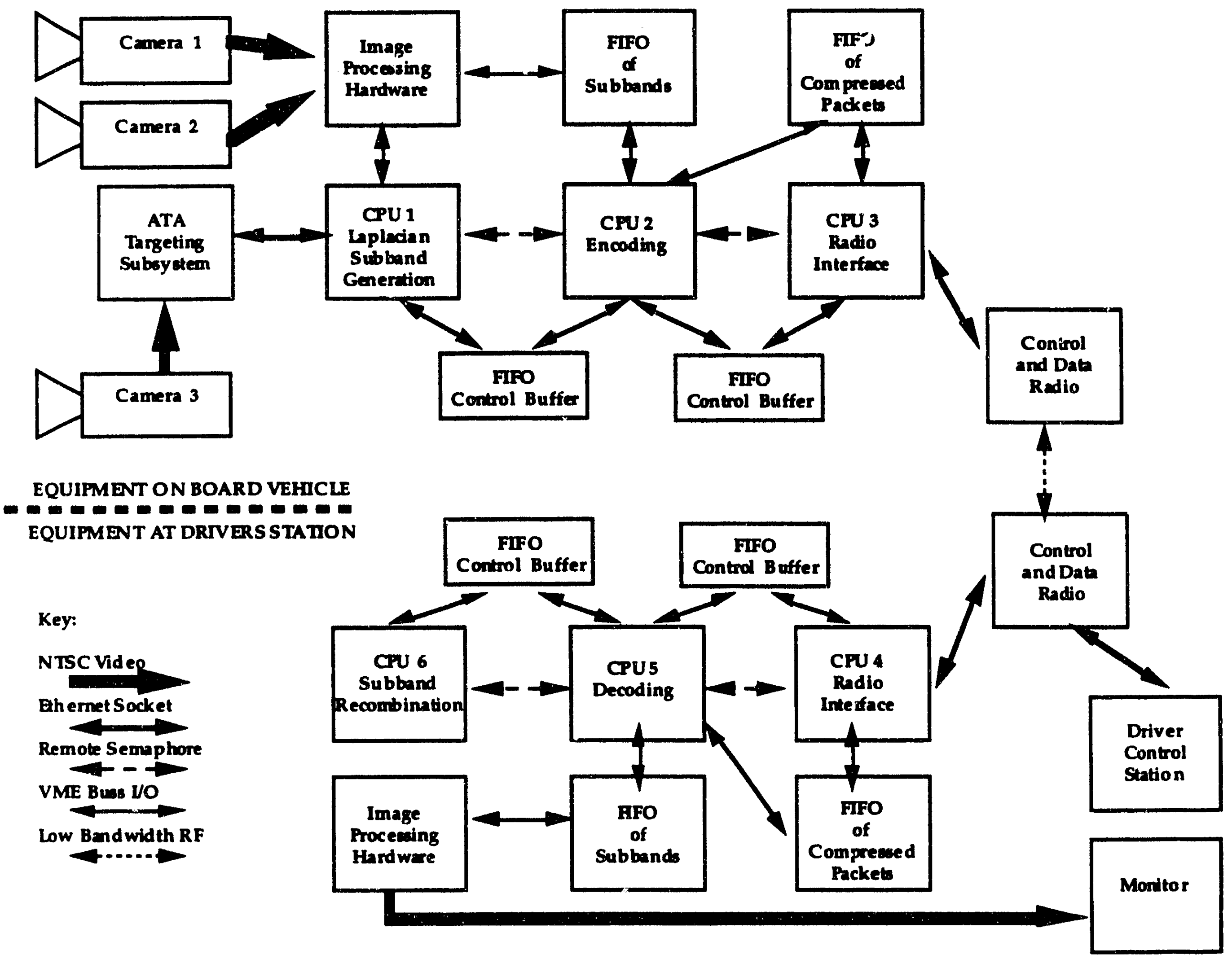

Fig. 2. Block diagram of Oak Ridge National Laboratory video compression system. (NTSC = National Television Standard Code, and FIFO = First in, first out)

DRIVING WITH THE ORNL SYSTEM

In April and May of 1992, the ORNL video compression system was tested and demonstrated at the Office of the Secretary of Defense's Robotics Testbed Demo I at the U. S. Army's Churchville test facility near Churchville, Maryland. During this and subsequent testing, many practical lessons were learned about driving with the ORNL video compression system.

The first task when initially setting up the system was to determine the best trade-off of quality vs frame rate for remote driving. Five parameter settings were determined that provided a range of quality vs frame rate operating points at 64 kbps. The parameters include the size of each subband's foveal area and the bin width of the quantizers for each subband. Codebooks were determined for each subband/quantizer combination to optimize the compression for each parameter setting. This was done using a video tape from the driving camera of the HMMWV while the test course was driven at 10 mph (16.1 kph). This optimization is exhibited by a maximized frame rate for a given parameter setting because the system is designed to transmit images as fast as possible for a given parameter setting and bandwidth. Frame rates ranged from 1 to 6 frames per second. It was quickly determined that the image quality should be somewhat compromised so that the frame rate could approach the upper part of this range. The previously discussed setting that provides a nominal frame rate of 4 Hz with a per frame compression of ~125:1 was determined to be the most suitable choice for driving with the ORNL system. The corresponding latency was determined to be approximately 1 s.

Early attempts at teleoperation made it obvious that speed control would be difficult given the 1 s latency. For example, as the operator released the brake, the vehicle began to move, but this motion was not immediately evident in the low-bandwidth video due to the latency. As the operator compensated, the vehicle soon went faster than desired. This led to a series of rapid starts and abrupt stops. Preoccupation with speed control detracted from the operator's attention to steering, and driving performance was unacceptable. This was addressed by putting both the HMMWV's transfer case and transmission in their lowest gears. This provided a more sensitive, limited range speed control and allowed the operator to concentrate on the overall driving task. In general, latency should be minimized and stable. This permits the operator to anticipate driving conditions by visualizing the vehicle slightly down range in the scene. Future systems would benefit from a cruise-control type of speed input as an operator aid.

Another problem that had to be overcome was the impact of competing radio-frequency (RF) traffic. The emphasis of the overall program effort was on the teleoperation of vehicles and demonstration of multiple integrated mission packages, not on the demonstration of radio technology. Unfortunately, this resulted in multiple radios operating in the 902 to 928-MHz band at the demonstration site. Although the spread spectrum radios provide some channel separation via their spreading codes, the resulting loss in system performance was unacceptable. Namely, packets containing subbands became garbled if another radio keyed up during the transmission, causing the built in features of the ARLAN radios to reject the subband packet, which initiated a

retransmission. This caused a drop in net available bandwidth below 64 kbps, as measured at the decompression rack's radio interface CPU (CPU 4 in Fig. 2). The impact on the compression system's performance was a reduced frame rate, and increased latency.

To reduce the effects of RF collisions, the ARLAN radios were operated on a higher bandwidth channel. Also, the Ethernet driver of the compression rack's radio interface CPU (CPU 3 in Fig. 2) was modified to use a smaller maximum packet size. However, the system still limited the video bandwidth to 64 kbps. These modifications lowered the amount of time the radio was transmitting for each physical packet. This, in turn, lowered the probability that a packet would be corrupted by a competing radio transmission. The impact of the competing RF traffic on the system's performance was significantly reduced, but there were still rare instances when the throughput would suffer. With a formal communications plan, this might have been avoided, but it was an issue with the prototype fleet. Hence, it became important that the operator be able to tell when the image was not being updated at some minimum rate (e.g., 2 Hz). To alert the operator, a small bar graph would appear if a half second elapsed without a new image. The bar graph grew in proportion to the age of the current image. When the bar graph appears, the operator should consider taking an appropriate action, such as slowing down.

In June of 1993, additional testing of the system was performed at the Aberdeen Proving Ground's Unmanned Vehicle Test Site. The test site is a 13-acre grass field with sections of gravel road. At this particular site, the gray level of the gravel road and that of the adjacent grass was essentially the same in the black-and-white driving image. Even with full-bandwidth black-and-white video, it was difficult for the operator to discern road from grass. This had not been a issue at the Churchville site, where a dirt road resulted in a different gray level than the surrounding grass. Experimenting with a number of color and infrared optical filters resulted in only limited success. This underscores the fact that good source video is very important for image compression systems. Adding color capability to the compression system will improve image understandability but does increase bandwidth requirements. Another possibility is to use an infrared camera. This will not increase bandwidth requirements, but may improve image contrast for such operationally significant concerns, such as being able to discern road from grass.

CONCLUSIONS

The ORNL video compression system supports teleoperation of an HMMWV at speeds up to 15 mph (24.1 kph) on a moguled dirt track. Experience in driving the system identified a quality vs frame rate operating point that favors frame rate to be suitable for driving video over a 64-kbps channel. Latency of the system was determined to be ~1 s, which affects driving performance. The system also supports remote automatic target acquisition at 64 kbps. Future work will address improvements to the compression algorithm, which will improve latency and quality for a given bandwidth. The use of color and/or infrared will be investigated to improve image contrast when needed.

ACKNOWLEDGMENTS

The authors express their gratitude to C. M. Shoemaker and T. W. Haduch for their programmatic guidance throughout this project and to the members of the Advanced Sensor Systems Branch of the Army Research Laboratory for their cooperation and excellent development efforts. We thank Tim Gee of The Georgia Institute of Technology for his significant contributions.

REFERENCES

1. F. W. DEPIERO, T. E. NOELL, and T. F. GEE, "Remote Driving with Reduced Bandwidth Communication," *Proc. Space Operations, Application and Research Symposium*, Houston, Tex. (1992).

2. F. W. DEPIERO, T. E. NOELL, and T. F. GEE, "A Video Transmission System for Low-Bandwidth Remote Driving," *Proc. American Nuclear Society Fifth Topical Meeting on Robotics & Remote Systems*, Knoxville, Tenn. (1993).

3. P. J. BURT and E. H. ADELSON, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. Commun.* **31,** 532–40 (1983).

4. P. J. BURT, "Multiresolution Techniques for Image Representation, Analysis and 'Smart' Transmission," *Proc. SPIE Conf. 1199: Visual Communications and Image Processing IV*, Philadelphia, Penn. (1989).

5. A. K. JAIN, "Image Data Compression: A Review," *Proc. of the IEEE*, **69,** 349–84 (1981).

6. S. BALAKIRSKY, P. DAVID, F. FISHER, P. EMMERMAN, and P. GAYLORD, "Semi-Autonomous Mobile Target Engagement Systems," *Proc. AUVS-93 Unmanned Systems: Expanding Government & Commercial Applications*, Washington, D.C. (1993).