Wavelet Based Video Compression for A Low-Rate Data Link

Michael W. Landry
L3 Communications, Conic Division
9020 Balboa Ave
San Diego CA 92123

Jeffrey C. Lee
L3 Communications, Conic Division
9020 Balboa Ave
San Diego CA 92123

ABSTRACT

Conventional video image transmission requires large data bandwidths, whether the signals are transmitted in analog or digital form. Many applications are limited to bandwidths that will not support image transmission unless high compression ratios are employed. Video compression techniques based on wavelet functions provide high data compression while preserving image quality, but are computationally demanding. Recently available high-integration devices that utilize wavelet basis functions for video compression have made possible low-cost high-performance video compression systems. Image data is more useful when provided with other data such as position information, telemetry, or other user data. This paper describes the technology, features, and limitations of a versatile low data rate system incorporating compressed video data.

Key Words: Video Compression, Wavelet, Data Link

INTRODUCTION

A communication system transmitting video signals can do so by directly modulating a carrier with the analog signal or by digitizing the video signal prior to modulation. Approximately 5 MHz of baseband video bandwidth will require as much, or even more, RF bandwidth after modulating, depending on the modulation type. Amplitude modulation results in a frequency translation to the carrier frequency, while frequency modulation results in an increase in the RF bandwidth needed. The bandwidth requirement for digitized video, 100 MHz or more, is prohibitive without compression. With compression, however, the bandwidth can be reduced below that required for transmission of the analog signal. An application which has available only a limited amount of RF spectrum, a few hundred kHz for example, will require high compression of the video and also reduction of the frame update rate in order to maintain usable image quality.

Reduction in data bandwidth by compression is achieved primarily from eliminating redundancy and inefficiency in the signal. Bandwidth can also be lowered by reducing

picture quality and frame rate. The image information, contained within a standard composite video signal, has redundancy that is removed by the compression algorithm. The composite video signal has formatting overhead such as synchronization and color reference, which can be transmitted more efficiently in digital form. Originally designed over 40 years ago, these signal characteristics were chosen to simplify the consumer receiver design at a time when each additional active circuit device was costly and analog signal processing was used exclusively.

Combining analog video with other data, such as position or telemetry, had been previously achieved by using sub-carrier modulation of the video, which adds extra complexity and system cost. A digital video signal is optimum for this type of system because it can be readily combined with other digital signals and the transceiver modulation is entirely digital.

THE DCNS SYSTEM

The Digital Communication Network System (DCNS) is a multi-terminal system that provides a mechanism to transmit remote data and GPS position between nodes. The DCNS operates in the UHF band between 350 and 450 MHz using filtered frequency shift keying (FSK) modulation. Time division duplexing (TDD) is used so transmit and receive links operate on same frequency channel.

TDD allows remote units to communicate directly with each other. When transmit and receive links are operated on different channels using frequency division duplexing (FDD), remote units can only communicate with a base station. Frequency planning is somewhat more complicated with FDD because two channels are needed, generally with a substantial frequency difference. A separate channel for communication to remote units would be under-utilized in many applications because most of the information flow is from the remote units.

The basic unit is comprised of a network controller board (NCB), a transceiver, a power amplifier, and optional GPS receiver. The controller has asynchronous serial ports for communication with an internal GPS receiver and with other external devices. Additionally, two general purpose slots are provided for options, such as a video interface.

Network communication is controlled by a master unit that is in communication with one or more remote units and relay units. A poll/response mode is available or a time division multiple access (TDMA) structure can be employed. A PC based graphical user interface (GUI) is used for unit setup and operation of the network from the master station site. The GUI communicates with the DCNS unit through the host serial port at baud rates up

to 115k. NCB firmware can be reprogrammed by downloading new code through the host port.

VIDEO CAPABILITY

The video capability is achieved by the inclusion of a video compressor board (VCB) in a remote unit and a video expansion board (VEB) in another unit, either a master or another remote. An NTSC or PAL color composite video signal is connected to the remote unit though an SMA connector. A D1 serial BT656.3 digital video or S-video Y/C signal can alternatively be provided as the input. The transmitted video signal is available at the other end of the link in either analog or digital form. The resulting Video Communication Network System (VCNS) retains all the basic DCNS capabilities.

The VCNS uses a wavelet based video compression engine to maintain good image quality even at high compression ratios. Compressed data frames occupy from 10 kilobits to 500 kilobits, depending on the user selected compression ratio. Naturally, image quality is better at larger frame sizes but take longer to transmit. The frame update rate is a function of the available network data capacity and the image compression ratio.

Unit position data may require only a small part of the available data bandwidth. For a basic VCNS unit with a link data rate of 76.8 KBPS, an allocation of 40 KBPS might be made available for video, resulting in frame rates from a few frames per second to a few seconds per frame. A feature of the wavelet compression is preservation of a viewable image even at this low data rate.

Data integrity is extremely important when transmitting compressed video data. A single bit error, if occurring in the wrong place, can corrupt the entire frame and prevent expansion of the image. The VCNS uses a cyclic redundancy code (CRC) appended to each message packet for error detection. If an error is detected, the sending unit can retransmit the packet to complete transmission of an error free video frame.

COMPRESSOR/EXPANDER HARDWARE

A high integration wavelet compression engine, the ADV611, is used in the VCB and the VEB. The same device operates as either a compressor or expander. In the compression mode, computations are needed to set compression parameters on a field by field basis. This is done using an ADSP 2181 16-bit fixed-point digital signal processor (DSP). In the expansion mode, only configuration registers need programming, no computations are required.

WAVELET VIDEO COMPRESSION

The image compression algorithm is based on a wavelet transform, a mathematical function developed and applied during the last 20 years. Two-dimensional spatial video information is transformed to the frequency domain using wavelets as the basis function, similar to the use of sine waves as the basis function in a Fourier transform. The frequency domain information is computed in subbands that are then adaptively quantized according to the way human visual perception occurs. Visual sensitivity decreases as the spatial frequency of the image increases; thus, fewer quantization levels are used to encode the higher frequency subbands. The lower frequency image components are quantized with more levels and therefore are allocated more of the bandwidth. This is the first stage of compression.

The quantized image contains sequences of consecutive zero and nonzero data. These sequences are run length coded, reducing the data needed for transmission. A single value represents the length of the data sequence in place of the longer sequence of constant data. Further compression occurs from a Huffman coder, which codes sequences that are more probable with a shorter code and less probable sequences with a longer code.

The compression is done on each independent field so no history of previous frames is needed. If there are dropouts in the link, the image is restored and displayed at the receiving end of the link upon completing the transmission of the next video field.

Wavelet compression has distinct advantages over other compression methods, such as JPEG. The wavelet transformed subbands contain information across the entire image field. JPEG image compression is done on 8x8 pixel regions, which creates undesirable block artifacts at high compression ratios. Referring to the image examples comparing JPEG and Wavelet compression, note the visible blocks of the JPEG image. The wavelet compressed image has smother transitions between intensity levels. This difference is also clearly visible in a color image.

Discrete Cosine Transform (DCT) compression used by JPEG and MPEG is significantly more computationally demanding to compress than decompress. These approaches are used for broadcast systems where there are numerous receivers for one transmitter, so receiver cost is most important. In a point-to-point link, the transmitters may out number the receivers. Hardware complexity for wavelet compression is comparable to decompression, lowering total system cost for these applications.

Another property of wavelet compression is the parameter control available. It is possible to control which frequency subbands are allocated more data; thus optimizing the compression based on the particular image.

Wavelet compression is employed in the VCNS because of the unique advantages for a low-rate data link. The functionality developed for the VCNS is equally applicable to a high bandwidth full-motion system.

VIDEO COMPRESSOR BOARD

The VCB accepts a composite video signal in either NTSC or PAL format. A multiplexer selects one of three composite video inputs or a single S-video Y-C input. A video decoder performs synchronization and 8-bit digitizing of the luminance and chrominance components. Y-Cr-Y-Cb samples are combined with embedded synchronization data according to BT656 format. Saturation, brightness, contrast and other picture settings are selectable through the GUI to compensate for line signal characteristics. Once set, these values normally do not require changes for a given source. The video decoder is programmed via an I2C (inter-integrated circuit) bus serial clock and data protocol from the DSP. Status values from the video decoder are also read over the I^2C bus.

A data selector selects the parallel BT656 data path from either the video decoder or the D1 serial interface circuitry. The selected data stream is fed to the ADV611, which performs the compression one field at a time. The ADV611 has access to a dedicated 4 Mbit DRAM used for manipulation of the video data and data buffering. Each video field is analyzed and image statistics are generated then read by the DSP to compute a set of 42 bin width values that are written back to the ADV611 for use in the compression algorithm. The bin widths set the amount of quantization, and thus compression, performed on the image. Each bin corresponds to a frequency subband processed by the algorithm. A frequency band is preserved or attenuated based on the respective bin value.

The DSP computes the bin values using a combination of open and closed loop processes, attempting to match the compressed image field size to the target field size. The target field size is calculated to achieve the user selected video frame rate at the specified data rate allocated to the video data. Before compressing the image, bin values are calculated using the target field size and an estimate of compressed field size. After compression, an error signal is generated by comparing actual compressed image size with the target. This error drives the next bin calculation to either increase or decrease compression. From field to field, the exact compressed field size will vary by 5-10%. The change will be more pronounced if there is a significant scene change such as if the source is switched.

The ADSP2181 is a high integration single chip DSP. The instruction cycle rate is 27 MHz, twice the externally supplied 13.5 MHz. It contains 16k words of program memory and 16k words of data memory on chip. All instructions execute in a single cycle giving the processor 27 million instructions per second (MIPS) capability at this clock rate. Although the instruction set is optimized for digital signal processing, a rich set of bit

manipulation and input/output (IO) instructions, along with multiple sources of interrupts, allow the processor to function as an embedded controller as well as DSP.

The 24-bit wide program memory is loaded from a slower byte-wide non-volatile flash EPROM device, which also stores user settings. The program code can be downloaded into the flash memory over the host interface serial port allowing field upgrades to be performed easily. New code can be distributed on disk or by e-mail to the user's site. The GUI is used to perform the download of the hex formatted file.

When data is available from the compression engine, an interrupt to the DSP is generated. The DSP reads the data blocks and inspects the data for framing information. The data is intrinsically 32-bit wide, but is read and processed 16 bits at a time. When the beginning of a field is detected, the data is written to a frame storage first-in-first-out (FIFO) buffer. A complex programmable logic device (CPLD) contains sequencing logic to control the FIFO writing operation. Upon detecting the end of a field, the DSP signals the controller that a complete field is available for transmission over the link. By counting words written to the FIFO, total field size is tabulated for use in the compression control loop.

The controller reads data from the frame storage FIFO as 8-bit values and forms packets with header information identifying the source and destination of the video data. An expander located in another unit within the network recognizes the source identifier and receives the compressed image data.

DSP code, written in assembly language, controls the operation of the VCB. Communication with the NCB microcontroller is through a pair of 8-bit wide registers addressable from both processors. Configuration and code download transfers use this mechanism. The NCB has direct access to the compressed video data from the frame storage FIFO.

VIDEO EXPANSION BOARD

The ADV611 operating in the expansion mode does not require any support except initial programming of mode information. All other data needed to perform the expansion is contained in the compressed image data. The VCNS controller programs the needed registers in the ADV611. A CPLD contains logic needed for interfacing between the VEB functions and the controller bus.

Two FIFO frame storage memories are used in reconstructing the compressed image. One memory is available to receive a new compressed image from the controller. The other FIFO supplies data to the ADV611 at a 60/50 Hz rate so that a flicker free viewable image can always be displayed, even for still frames.

A standard NTSC/PAL video frame at 30/25 Hz consists of two interlaced fields updated at a 60/50 Hz rate. At the low data rate and bandwidth utilized by the VCNS, this frame update rate is not possible with an acceptable image quality. A few frames per second is the highest practical rate. The VCNS compresses one field image but not the other and duplicates the omitted field in the expansion process. This cuts in half the data transmission requirement with barely perceptible image degradation. The expanded image may remain static on the viewing monitor for several updates of the 60/50 Hz field rate. If two interlaced fields were transmitted with any movement in the intervening 16/20 mS between fields, the moving object would be displayed with noticeable 30/25 Hz dithering. A less costly solution to motion compensation is to process only one field.

After the ADV611 expands the compressed image, 8-bit BT656 data is supplied to a video encoder that reconstructs composite and S-video Y-C signals. The same video adjustments available to the compressor are also available to the expander to compensate for monitor characteristics. The BT656 data is also supplied to a D1 digital interface.

VIDEO PROCESSING FEATURES

A cropping feature is available to allow the user to select a region of the image to have enhanced quality leaving lower quality outside the region. The degree of attenuation of contrast outside the quality box is programmable from 6 dB to 30 dB, which ranges visually from a modest reduction to almost no contrast. By using this mode, the image quality inside the box can be increased for a given data rate, or the frame rate can increase. This and all programmable values may be programmed locally at a unit or over the link from the master unit site.

NETWORK CONTROLLER BOARD

The Controller uses an 80C320 embedded 8-bit high performance processor. In addition to the two serial ports in the processor, a serial communications controller (SCC) provides two additional channels. One serial port communicates with the host computer; another interfaces to the transceiver. A third port communicates with the internal GPS receiver and a fourth is spare. A general purpose extension of the processor bus is routed to the motherboard connector to communicate with option boards such as the VCB and VEB. Each board is assigned a 3-bit address to locate it within the address space of the processor. A VCB and VEB can both be installed in a single chassis if required by an application.

TRANSCEIVER AND POWER AMPLIFIER

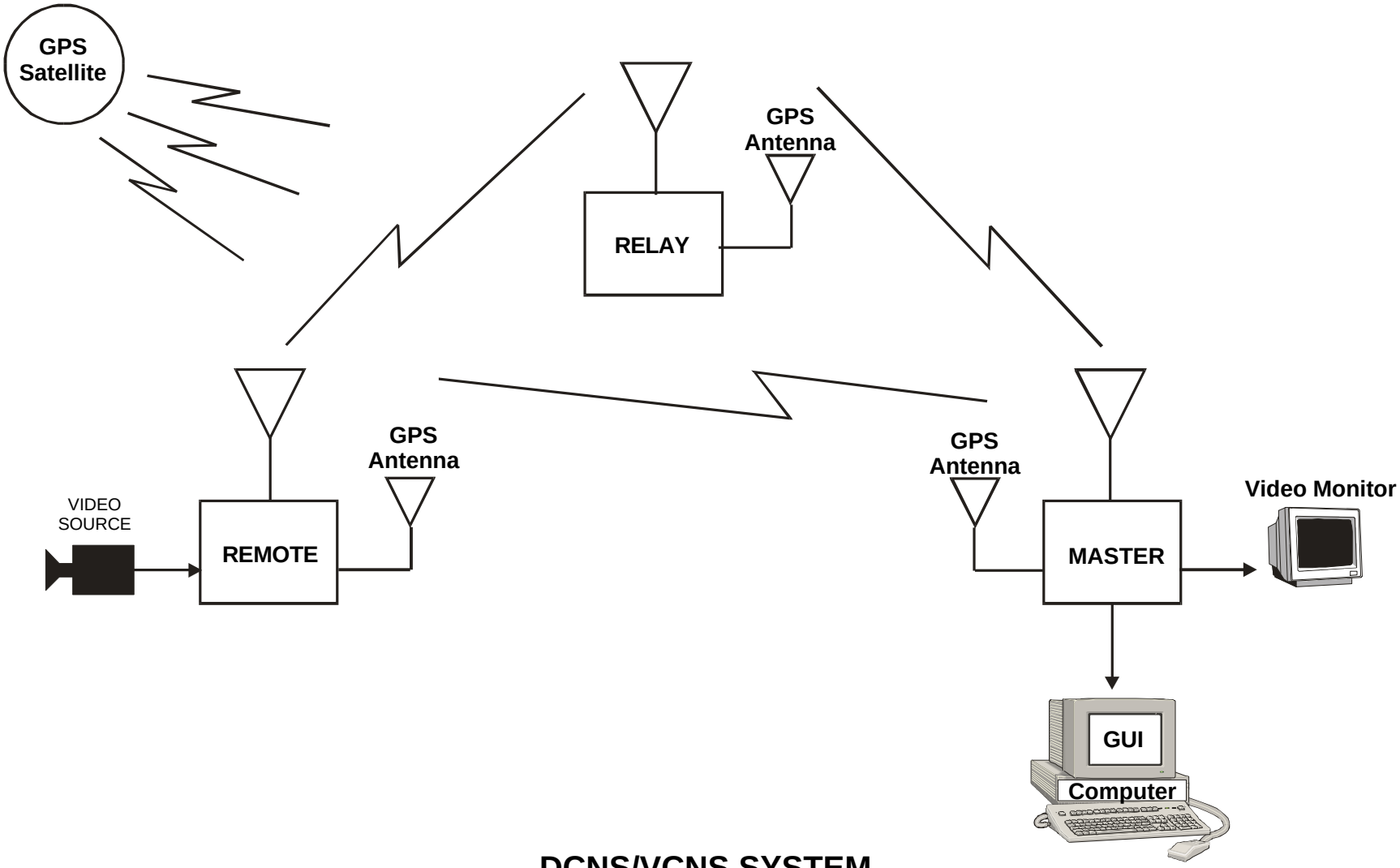
A 100 mW transceiver performs FSK modulation and demodulation. With a 76.8 KBPS data rate, the 99% occupied bandwidth is 250 kHz. The RF carrier is user selectable over a 6 MHz band factory set in the range from 350 MHz to 450 MHz.

The standard VCNS uses a 5-watt power amplifier (PA). For applications requiring more power, the PA output can be used to drive a higher power external PA. The transmit/receive switch control signal is available to control an external duplexer.

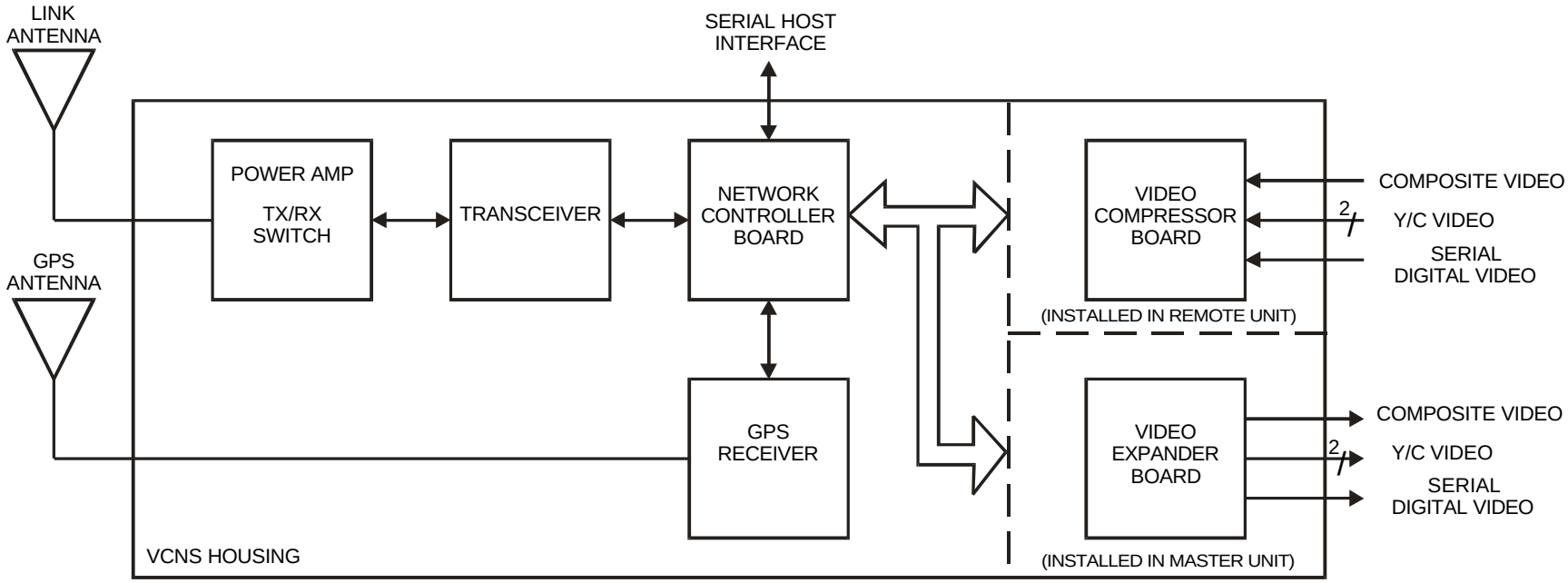
Other transceiver/power amplifiers can be used to operate in a different frequency band.

HOUSING

The VCNS circuit boards are contained in a sealed aluminum housing design to meet NEMA 4 requirements. Fins support convection cooling of the unit.

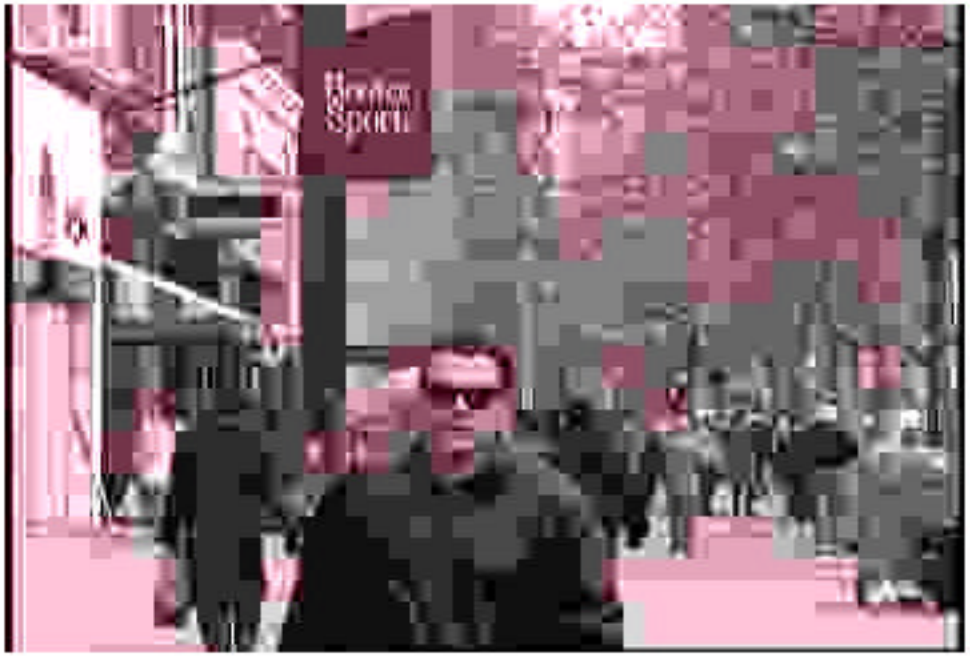


DCNS/VCNS SYSTEM

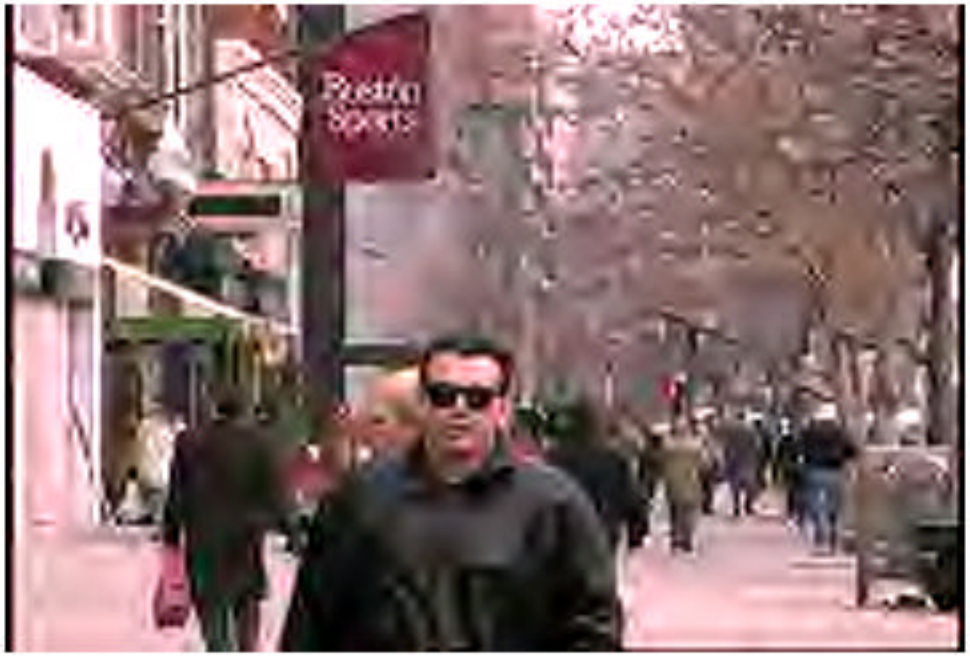

LINK
ANTENNA
GPS
ANTENNA
SERIAL HOST
INTERFACE
POWER AMP
TX/RX
SWITCH
TRANSCEIVER
NETWORK
CONTROLLER
BOARD
GPS
RECEIVER
VIDEO
COMPRESSOR
BOARD
(INSTALLED IN REMOTE UNIT)
VIDEO
EXPANDER
BOARD
(INSTALLED IN MASTER UNIT)
COMPOSITE VIDEO
2
Y/C VIDEO
SERIAL
DIGITAL VIDEO
COMPOSITE VIDEO
2
Y/C VIDEO
SERIAL
DIGITAL VIDEO
VCNS HOUSING

VCNS UNIT

JPEG COMPRESSION 70:1 3.5KB



WAVELET COMPRESSION 70:1 3.5KB



JPEG COMPRESSION 100:1 2.5KB

WAVELET COMPRESSION 100:1 2.5KB