24/7 Security System: 60 FPS Color EMCCD Camera with Integral Human Recognition

T. L. Vogelsong[a], T. E. Boult[b,c], D. W. Gardner[a], R. Woodworth[b], R. C. Johnson[c], and B. Heflin[c]
[a]Salvador Imaging, Inc., 5061 N. 30th Street, Suite 103, Colorado Springs, CO 80919
[b]Securics, Inc. 1867 Austin Bluffs Parkway, Colorado Springs, CO 80918
[b]University of Colorado at Colorado Springs, 1420 Austin Bluffs Parkway, Colorado Springs, CO 80933-7150

ABSTRACT

An advanced surveillance/security system is being developed for unattended 24/7 image acquisition and automated detection, discrimination, and tracking of humans and vehicles. The low light video camera incorporates an electron multiplying CCD sensor with a programmable on-chip gain of up to 1000:1, providing effective noise levels of less than 1 electron. The EMCCD camera operates in full color mode under sunlit and moonlit conditions, and monochrome under quarter-moonlight to overcast starlight illumination. Sixty frame per second operation and progressive scanning minimizes motion artifacts. The acquired image sequences are processed with FPGA-compatible real-time algorithms, to detect/localize/track targets and reject non-targets due to clutter under a broad range of illumination conditions and viewing angles. The object detectors that are used are trained from actual image data. Detectors have been developed and demonstrated for faces, upright humans, crawling humans, large animals, cars and trucks. Detection and tracking of targets too small for template-based detection is achieved. For face and vehicle targets the results of the detection are passed to secondary processing to extract recognition templates, which are then compared with a database for identification. When combined with a PTZ, the resulting system provides a reliable wide-area 24/7 surveillance system that avoids the high life-cycle cost of infrared cameras and image intensifiers.

Keywords: Camera, CCD, low light, EMCCD, night vision, surveillance, security, detection, identification, face recognition

1. INTRODUCTION

The need for advanced surveillance and security imaging systems is growing rapidly due to the greater awareness of security threats following the 9/11 terrorist attack and recent asymmetric wartime actions. Surveillance of border activities, battlefield monitoring, and high risk sites such as nuclear plants, oil fields, ports, aircraft carriers, etc. is becoming more critical. Unmanned vehicle deployment is driving size and power constraints as well as the need to operate under a broad range of difficult environmental conditions. An ideal surveillance imaging system would have the following characteristics:

- Multiple cameras to cover the field of view (scalable systems)
- Capture and process data at video rates
- Run continuously over a broad range of lighting conditions from full sunlight to overcast starlight
- Be able to capture multi-spectral and/or polarization data to enhance discrimination
- Be auto-adjusting to automatically optimize the image quality for the given illumination and scene conditions
- Process the video data in real time to detect targets and track their movement (intelligent data compression)
- Transmit target data back to a central monitoring station
- Be small, rugged and work over a wide range of environmental conditions

In order to create a system which addresses the desires indicated above, the authors are combining low light electron multiplying charge coupled device (EMCCD) cameras developed by Salvador with algorithms for image enhancement as well as detection, discrimination and tracking of objects of interest developed by the University of Colorado at Colorado Springs (UCCS) and Securics with FPGA-based real time implementation of those algorithms by Securics.

In this paper we will describe the elements of the security system being developed. The first section will describe low light EMCCD technology and provide the specifications for Salvador's SI-VGA60-EM CCD camera. The next section

Preprint of a paper to appear in SPIE Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VI, April 2007. Do no copy

compares EMCCD imaging capabilities with those obtained from intensified cameras. The next section will discuss the advantages of EMCCD cameras in defense and security applications, and consider potential new applications enabled by this technology. The following section will describe advanced detection and tracking algorithms which have been used with visible, NIR, and thermal cameras under difficult real world imaging conditions. Identification and discrimination will also be covered including face recognition at a distance. These algorithms where selected for their ability to be implemented in FPGAs to enable on-the-fly detection and tracking. Next, we describe the camera architecture with on-board FPGAs, and describe how this integrated capability can be packaged into a UAV compatible format. Finally we discuss how these security camera systems could be linked in to a central station where the information gathered may be combined with the input from other sensors such as NIR, lidar, thermal IR, etc.

2. EMCCD TECHNOLOGY AND CAMERAS

The recent development of EMCCD sensors has enabled the performance advantages of charge-coupled devices to be extended to low light visible and near IR applications, heretofore addressed by intensified CCDs. These benefits include high resolution, video rate imaging with wide dynamic range in a cost-effective robust solid-state device. An EMCCD achieves this extended capability by providing a programmable gain mechanism on-chip, prior to the introduction of amplifier noise.

2.1 CCD Architectures

Before describing the gain mechanism in detail, we will first review the basic operation of a standard CCD. While all silicon CCD sensors sense light by converting photons to electrons, there are several different architectures for reading out the charge signal from each pixel. There are full frame CCDs where image capture and readout occurs in the same location. Readout is achieved by shifting the charge row by row into a horizontal readout register. Once a row's worth of charge is in the readout register, it is shifted pixel by pixel to the corner where it is converted from charge to voltage by the readout amplifier. The amplifier adds electronic noise to the signal related to the noise bandwidth of the amplifier (faster readout rate => higher readout noise), and the capacitance of the readout node. One issue with the full frame architecture is that if the light source is not turned off during readout, or a shutter used to block light during readout, the incoming light will continue to be collected as the previous image is being slowly shifted to the edge, causing smear. A second CCD architecture called a frame transfer CCD adds a frame storage area. After an integration interval, the charge from the image area is quickly transferred row by row into a full size image storage area, which has an optical light shield to block incoming light. The transfer time is typically 1 millisecond, so for short exposures or high overload light spots, image smear can still occur. Unlike the full frame CCD, integration of the next frame can occur while reading out the previous frame. A third architecture, called interline transfer reduces the smear further by including a storage site at each pixel. After an integration interval, the charge for all pixels is transferred into its local storage area, covered by an opaque layer. The next frame can start integrating while the previous frame is read out from the storage area. However, while the previous frame is being read out, a small fraction of the incoming light can leak into the readout register, causing smear. This is apparent for very short exposure times, and long readout times, or with scenes with very large intra-scene dynamic range. To further reduce image smear, a fourth architecture, called frame interline transfer (FIT) CCD's, was developed and has become standard for the broadcast video industry. In an FIT device, as shown in figure 1, there is both an interline storage register and a frame storage area. After the integration time, the charge from each pixel is transferred in parallel to the local interline storage area. Then this charge is quickly transferred (in ~ 1 millisecond) to the frame storage area. Integration of the next frame can start immediately after the interline transfer, in parallel with readout of the previous frame. The EMCCD camera described in this paper uses the FIT architecture.

2.2 CCD Noise Sources

There are several sources of noise in a charge-coupled device. This includes electronic noise from the charge to voltage conversion that takes place in the output amplifier. A second source is the dark current noise, or fluctuations in the amount of thermally generated charge in each pixel. A third source is photon shot noise or the amount of variation in the incoming photons themselves. This photon shot noise provides an ultimate limit to the signal-to-noise of a perfect image sensor. Another noise source is kTC, or reset noise due to the resetting of the output node to a fluctuating level. This source can be addressed through the use of a correlated double sampling readout method. The total noise contribution for a state of the art conventional CCD running at video speeds is a few 10s of electrons.

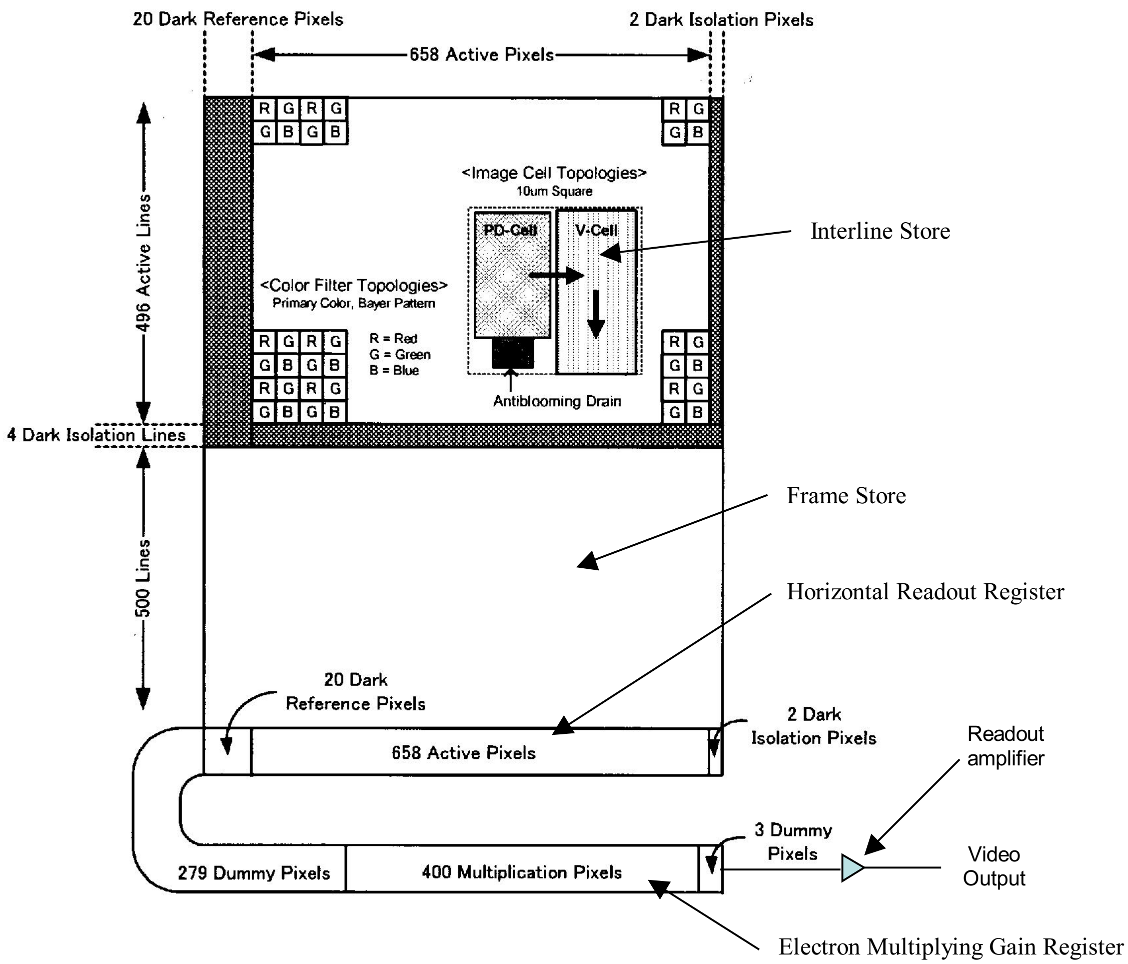

Fig. 1. Frame Interline Transfer (FIT) EMCCD architecture

2.3 Noise reduction in an EMCCD

A new advancement in CCD technology was recently developed[1,2,3,4] called electron multiplying charge-coupled devices or EMCCDs. In this type of sensor, the horizontal readout register is extended, as shown in figure 1[5], by a number (n=400) of additional transfer stages (the "gain stages"), prior to reaching the output node where the charge signal is transformed into a voltage signal by the output amplifier. These additional stages do more than just transfer charge. At each stage, one of the overlying electrodes is biased to a high voltage (typically 20-50 V, vs. 5-15 V for a standard CCD) as shown in figure 2. This creates a strong localized electric field that accelerates the electrons to the point where impact ionization may occur and the single electron into that stage may provide 2 electrons out of that stage. This gain method is statistical in nature, and the probability of ionization at any stage is low, resulting in a gain of only slightly more than 1 per stage, typically 1.01-1.02 at maximum gain setting. However, when many of these gain stages are used (typically 400-600), the resulting gain can be over 1000x ($1.015^{500} = 1710$). The amount of gain can be adjusted from 1x to the maximum by adjusting the bias on the high voltage gate as shown in figure 3. Note that this gain occurs in the charge domain, prior to reaching the output node. Thus, like an ICCD, the signal can be boosted prior to the primary source of electronic noise – the output amplifier. Unlike an ICCD, the dark current generated charge in an EMCCD is also amplified by the same gain factor. However, dark current, and the shot noise variation on that dark current can be reduced by cooling the sensor to the point where it does not contribute to the overall noise. Under these conditions, an EMCCD camera can provide effective noise levels of less than 1 electron even at high readout speeds, thus boosting sensitivity far beyond that of a standard CCD. Note that EMCCD technology can be added to any of the 4 CCD architectures mentioned in the previous section.

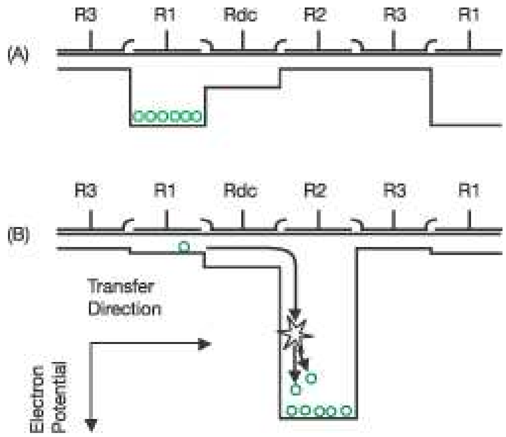

Fig. 2. EMCCD clocking waveforms illustrating electron multiplication by impact ionization.

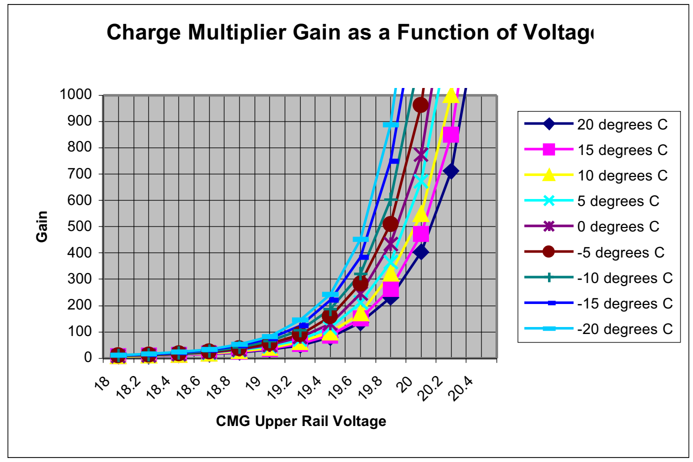

Fig. 3. Gain vs. clock voltage and temperature for Salvador's EMCCD Camera

Salvador has developed a family of cameras based on EMCCD sensors, tailored to defense and security applications. The camera for this project uses either an RGB color (Bayer pattern) or monochrome sensor with 658 x 496 pixels. The sensor is an FIT architecture which minimizes smear. The camera reads out the image progressively at a 60 Hz rate. The integration time is controllable with the electronic exposure as short as 100 usec, and as long as several seconds, in order to integrate up the scene under extremely low light levels. A picture of this camera is shown in figure 4. Salvador is also developing versions of this camera for specific customers that meet stringent temperature, shock and vibe, physical size, power, and other environmental and operational constraints. A summary of the specifications is provided in table 1.

Fig. 4. Salvador's SI-VGA60-EM camera with a color EMCCD sensor

Table. 1. Camera Performance Specifications

Parameter	***Specification***
Resolution	658 x 496 pixels
Pixel size	10 µm x 10 µm
Maximum frame rate	30 or 60 fps (programmable)
Sensor architecture	Frame Interline Transfer
Full well capacity	28 ke$^-$
Amplifier noise	22 e$^-$ @ 60 fps and EM gain = 1
Amplifier noise @ gain =100	1 e$^-$ equivalent
Electronic Shutter	500 usec - 5 sec
Blooming	Up to 500x overload
Pixel rate	25 MHz @60 fps
Data format	Camera Link or Gigabit Ethernet
Dark signal (@ 0 °C)	1 pA/cm^2
Intra-scene Dynamic range	63 dB @ 30 fps
Lens Mount	C-mount or F-mount
Size	3.15" x 3.15" x 4.51" = 80 mm x 80 mm x 114.6 mm
Weight	1.5 lbs not including lens
Operating temp	0 to 40 °C
Power supply	+24-32 VDC, 110/220 VAC with AC/DC converter

3. COMPARISON WITH INTENSIFIED CCD CAMERAS

With an effective noise level comparable to an ICCD and a quantum efficiency comparable to or superior to that of an ICCD, the EMCCD provides sensitivity equal to or better than an ICCD. The resolution of the EMCCD is superior. The EMCCD has an MTF characteristic similar to a conventional CCD, providing a sin(x)/x response comparable to that of an ideal square pixel detector. The ICCD suffers resolution degradation in the photocathode layer, the micro-channel plates, and in the fiber coupling mechanism. Therefore, even if the ICCD and EMCCD quote the same number of pixels, the resolving power of the EMCCD will be superior.

One additional factor affecting the sensitivity comparison is the noise factor. This factor is due do the statistical nature of the multiplication mechanism and results in a multiplication of the noise of a signal entering the gain register (e.g. photon-generated and dark current generated shot noise). This noise factor has been estimated through analysis[6] at $\sqrt{2}$ for large values of gain and measured as 1.2-1.3x, and thus is better than the 1.6-2.0x excess noise factor of an ICCD[7].

An additional advantage of EMCCDs is apparent under extremely low light conditions. ICCD cameras suffer from scintillation noise and ion backstreaming, which at low light level produce a noise characteristic which is non-Gaussian and shows correlation from frame to frame. The EMCCD characteristic noise at low light level is Gaussian as it is due to shot noise. Therefore, spatial and temporal filtering techniques can increase the signal-to-noise with the EMCCDs and the impact can be quantitatively predicted. In fact, a sign-to-noise improvement of an order of magnitude has been shown with even moderate spatial and temporal filtering. The same methods are much less effective with an ICCD due to the correlated nature of their noise characteristic.

Being solid state, the EMCCD sensor is inherently as compact and rugged as a standard CCD, and is eminently suitable for field deployable systems. In fact, Salvador Imaging is developing cameras for military surveillance operating from full daylight to overcast starlight in compact, rugged housings for deployment in manned and autonomous vehicles. Salvador has previously developed cameras in a "block camera" format for deployment on UAV's as shown in their SI-1M30-FT camera shown in figure 5.

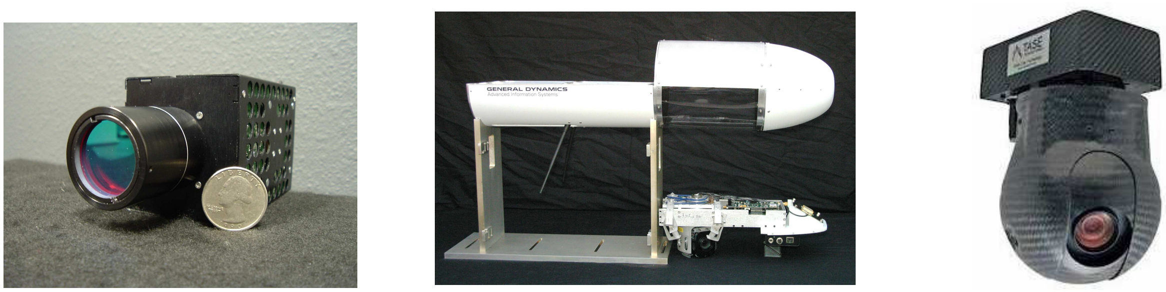

Fig. 5. Salvador's SI-1M30-FT block camera with polarization sensor mounted in compact Cloud Cap turret for General Dynamics UAV

The ICCD does have one advantage over the EMCCD for applications where time-gating is critical. The MCP of an ICCD can be turned on and off in nanoseconds, where the minimum integration time of an FIT EMCCD presently is in the range of a few microseconds. We are working on methods of reducing this minimum time further for applications where very short time gating is required.

3.1 Comparison with InGaAs NIR detector arrays

One additional comparison has been made between silicon EMCCD's and NIR detectors based on InGaAs detectors. For actively illuminated scenes, illuminated in the 1.2-1.7 um range, and InGaAs array will of course perform better. However, in passive operation, preferred for clandestine scenarios, the contrast in many scenes, combined with the lower electronic noise level achieved in silicon EMCCDs provides superior detectability. We have experienced this most dramatically in operation at dusk and during low light face recognition applications. Due to the investment made in silicon technology, the cost/resolution ratio favors silicon EMCCDs over more costly high resolution InGaAs arrays. Example low light images taken of the same scene with a 1k x 1k EMCCD and a 320 x 240 InGaAs array are shown in figure 6 (courtesy of SAIC). *The* ideal solution for applications which are not space or cost constrained is to provide both cameras and use multi-spectral fusion processing.

Fig. 6. Experimental comparison by SAIC for Sceptre Finder program – Images taken with 75 mm lens at 400 yards under ¼ moonlight. Image on left is from Salvador's SI-1M30-EM camera (1k x 1k, 30 fps EMCCD). Image on right from Goodrich (Sensors Unlimited) 320 x 240 InGaAs camera.

4. DEFENSE AND SECURITY APPLICATIONS ENABLED BY THE EMCCD CAMERA

There are a number of attributes of the EMCCD camera which change the way that surveillance and security systems are envisioned. This section discusses attributes of the EMCCD cameras which enable new capabilities at the systems level.

4.1 Persistent Surveillance

Because of the MCP/phosphor architecture of an ICCD, the device will burn out if exposed to bright scenes. Therefore, it is typical to turn off the ICCD camera during daylight operation and add a visible light camera in parallel, with the additional complication of splitting the optics path and aligning the multiple sensors. An EMCCD can be operated in direct sunlight as well as in night by reducing the gain down to 1x, and operating the device as a standard CCD. For operation in even brighter conditions, the FIT architecture enables electronic exposure control, with the ability to reduce the integration time from a full frame time (~16 msec), down to a few microseconds. The combination of programmable EM gain, electronic exposure, and controllable iris extends the camera's inter-scene dynamic range to over 100,000,000:1! The ICCD also suffers from flare under bright spot illumination due to the repulsion of a concentration of electrons from the photocathode. A bright spot on the EMCCD will only expand due to the intensity profile of the light itself, as the on-chip anti-bloom capability prevents charge spillover into adjacent pixels. Therefore the use of an EMCCD camera enables *true single camera 24/7 operation*.

4.2 Color, hyperspectral and polarization imaging

There are several other less obvious advantages of EMCCDs which can open up additional applications. Since the EMCCD is composed of a standard array of pixels, different filters may be applied on a pixel-by-pixel basis, similar to the Bayer color pattern used on many consumer level CCD cameras. Thus *color low light imaging is now possible*. An example color image acquired with the SI-VGA60-EM color camera is shown in figure 7.

Fig. 7. Color image acquired with Salvador's SI-VGA60-EM camera

It must be understood however, that the use of spectral filters reduces the number of photons impinging on each pixel, and thus sensitivity is reduced. However, studies[8] have shown that under certain scenarios, the added information provided by color, outweighs the loss of signal to noise from spectral filtering for human detection of objects of interest. Most common R, G, and B dye filters are transparent in the near infrared, so an optimal solution may be to use an IR blocking filter for day-to-dusk operation, and allow the IR during night operation to maximize the sensitivity.

Other spectral filters could be applied for multi-spectral acquisitions. In addition, Salvador has worked with General Dynamics to demonstrate cameras with polarization filters applied on a pixel-by-pixel basis. The use of multi-spectral and polarization information can be used as further discrimination factors enhancing target detectability and recognition even in camouflaged situations. This can be particularly advantageous for drug interdiction for camouflaged vessels, soil which has recently been disturbed, decoys, or man-made objects intended to blend into the background such as those hiding IEDs. Otherwise unobservable variations in surface topology can be made to stand out using polarization imaging. Recent studies have been made of the advantage of using spectral information and polarization information in enhancing the capabilities of face recognition. For example figure 8[9] shows how information from different spectral bands and different integration times can be used to accentuate different facial features used for information extraction and identification. Figure 9[10] (courtesy of General Dynamics) shows how in polarization images the Stokes vector information is used to create an estimate of the surface normalization at each point on the face. From the front on view of the face, a reconstruction of a rotated view has been made. This additional polarization information aids the recognition process.

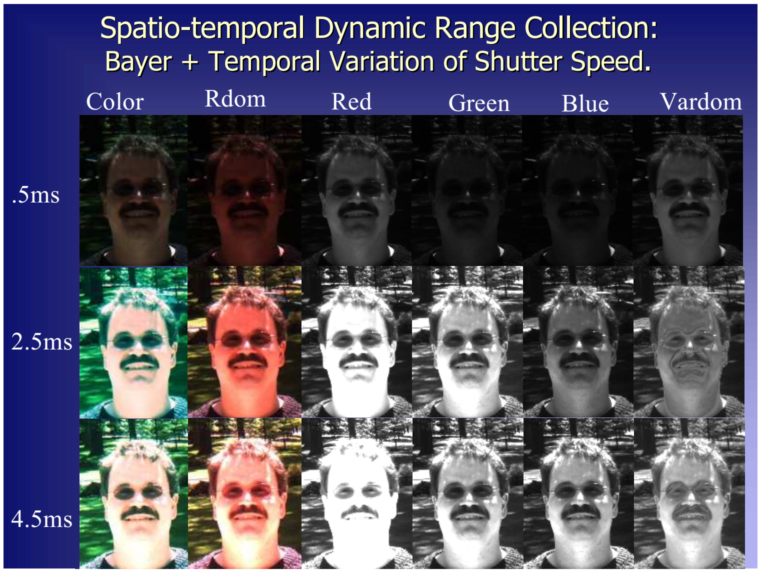

Fig. 8. Multi-spectral and multi-integration time information used to accentuate different facial features used in face recognition

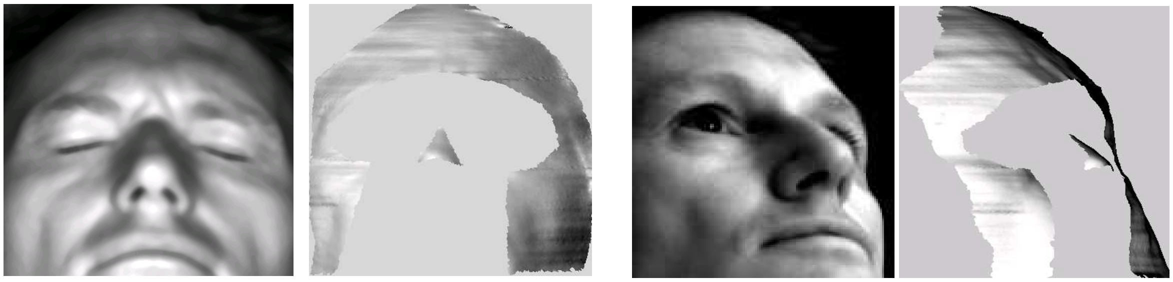

Fig. 9. Enhancement of face recognition based on the use of polarization images. The image pair on the left shows the front on view acquired by the camera and the orientation information extracted. The right image pair is an image created using the surface normal information extracted from the front on view.

5. DETECTION, TRACKING, AND IDENTIFICATION ALGORITHMS

Intelligent networked video surveillance is a well-established commercial technology using automated identification algorithms. Dr. Boult and his students have been leading advanced video detection/tracking work for more than a decade[11] and working on advancing identification methods for almost as long. One of the goals of this work has been to develop methods that are tailored for embedded processing. There are a plethora of approaches for detection and tracking and for identification. The issues are not defining new ones, but in choosing ones that are well suited to use with the EMCCD camera and also reasonably well matched to potential implementation inside the camera, providing real-time recognition, tracking all at low power.

Background subtraction is a very common approach. But with low contrast targets and the speckle noise that is common with any low light imager, traditional algorithms, e.g. mixture of Gaussians, will either have a high variance, hence become insensitive to low-contrast targets, or will have a high false detection rate. The quasi-connected components (QCC) algorithms[12], which addresses these problems, is summarized in figure 10. The QCC algorithm, unlike a mixture of Gaussians, is also extremely well suited to an embedded implementation since it uses only fixed point math and is very parallel. The team at UCCS is currently mapping this algorithm to a Virtex4-FX12, with the low level pixel processing in the FPGA and the higher levels of target tracking in the hard-core PPC. Analysis shows that it will run comfortably at 60 fps on the full mega-pixel image, leaving sufficient resources for compression and networking support as well.

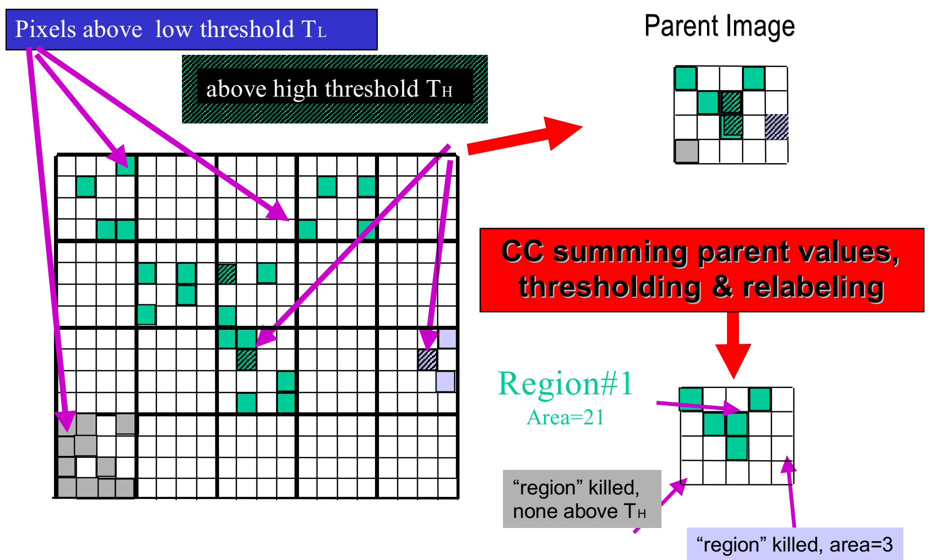

Fig. 10. Quasi-Connected Components: Thresholding with hysteresis combined with multi-resolution processing significantly improves segmentation.

QCC combines multi-resolution processing with a threshold-with-hysteresis approach. A resolution-reduction phase allows the algorithm to fill gaps. As it is reducing the resolution, it accumulates the evidence from the magnified images while tracking the number of pixels above the low and high threshold separately. It then applies a standard connected components algorithm at the reduced resolutions and again tracks pixels above low and high threshold separately. The algorithm then prunes away regions that either have no pixels above the high-threshold or that are simply too small. The low-threshold provides for a very sensitive detection even with low contrast, while requiring some of the region pixels to be above the high-threshold in order to maintain robust detection. AGC artifacts such as scene lighting or atmospheric impacts which can produce small amplitude changes.

The use of dual-thresholds and hysteresis lets the approach have low threshold that can detect parts of the target with very low contrast, but they must connect to components above the high-threshold to keep large regions of "noise" from being considered targets. Figure 11 shows an example of the QCC algorithm being applied to imagery from the EMCCD camera. Imagery was taken with a 400mm f2.8 lens at between 200m and 250m an hour after sunset with a moonless sky. The box in the images shows the detected target. Note the very low contrast of much of the target.

Fig. 11. Target detection and tracking using the QCC method on an image from a video sequence taken one hour after sunset on a moonless night.

Beyond the detection and tracking algorithms, it is also extremely useful for the system to be able to identify particular targets. Again, many object identification algorithms exist and the question is which ones are well suited to mapping into FPGAs. Our approach is to use a general object identification scheme based on Ada-boosted Cascaded Haar wavelets, followed by an enhanced PCA-based recognition engine.

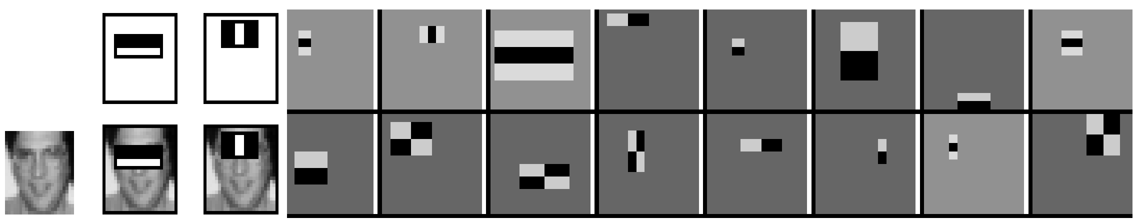

Fig. 12. The left images show 4 "Haar" wavelets and how they interact for face detection. The filter weight is the sum of the per pixel product of the value under the mask values and either –1 (black) or 1 (white). The overall response is the weighted sum of the cascaded thresholding of the filter response. In this example, the weight of each filter is shown as the brightness of the background of the Haar wavelets on the right.

Ada-boosting with Haar-like features has become a significant basic recognition algorithm for vision, in part because the computations are fast and in part because with Ada-Boosting, building a good classifier from a collection of weak classifiers is easy and effective. The face-detection system developed by Viola and Jones[12] was an early system showing its power. The basic concept of a Cascaded-Haar-wavelet (CHW) system used simple features, built from sums over rectangular regions.

The CHW process begins by computing an "integral image", II(x,y) which is the sum of the values above and to the right of x,y. After an initial computation of the integral transform, $II(x,y)$, the sum of pixels within any rectangle in image I which has vertex points A,B,C, and D is $Area = II(B) - II(C) - II(A) + II(D)$. Using this approach, an important aspect of the Haar wavelets is that they are scaled rectangular filters applied to the image window and hence can be computed with only a few operations. Viola-Jones defines a simple weak classifier based on a Haar wavelet features. This weak classifier is 1 if $pf(x) < p\theta$ and 0 otherwise, where h is the classifier, p is the parity, θ is the threshold, and f is one of the integral transform-type features discussed above. The classifier reports a value of 1 when it believes that a target is found, and it reports a value of 0 when a target has not been found. Training a classifier, which uses a given feature f, consists of discovering the threshold and parity which maximize its classification performance within the training set. Such a simple classifier, by itself, is not particularly useful, but using a collection of many such weak classifiers can make a strong classifier. Importantly, the classifier need not evaluate all of the features. If it can determine the object cannot match after only a few features are analyzed it can stop processing and move on. With this processing, though the worst case may be 1 second for object identification, in our FPGA implementation, the average expected detection time is only 6ms per frame.

Rather than using the traditional approach of "scaling" of features, we found that scaling the image on the fly, in the FPGA, was more efficient as we have more compute power than random IO access.[13] The analysis shows that with no constraints on the face-size, using the standard approach for a mega-pixel imager, requiring 4Mb of off-chip memory for the integral images and cascades, and assuming a 32-bit interface to memory, will need 18MB/second memory bandwidth. If, however, we do the real-time scaling and can limit the scales, e.g. using expected target sizes, we can reduce this to only requiring 72KB of memory for the integral image and 4KB of memory for the object descriptors. The initial implementation was designed for Elphel 333 camera, with a Spartan 3 FPGA. Figure 13 shows examples of the face detection algorithm applied to EMCCD imagery. The images were taken 1 hour after sunset at 100m and 150m, with a street light being the dominant illumination being 50m away for the left image and 75m away for the second. In both cases the camera used a 400mm F2.8 lens and was on a roof top 25m higher than the subject.

Fig. 13. Face detection examples. The left two images were taken at 100m with a street light about 50 m to the left and slightly behind the subject. The right image is at 150m and the street light is in front of the subject on the left. Note the significant difference in illumination that must be handled.

The face identification takes more than just the face detector. After detection the process must localize the eye regions, normalize the resulting image and then apply the face recognition process. For the eye localization we reuse the CHW engine, this time with an eye region localizer rather than a face detection. The image is then normalized and passed to the processor for recognition. However, before any face-recognition algorithm can be applied to his class of imagery a new approach to normalization is needed. The low-lighting and noise of intensified imagery, coupled with the potential for strong directional lighting (after amplification) makes this a challenge. Securics has developed a new patent pending lighting normalization algorithm just to address these issues. It is also extremely well suited to an FPGA implementation as it is based on a dual-LUT mapping of the data. Figure 12 shows an example of this applied to multiple low-light examples as well as comparing it to the standard normalization from the CSU face toolkit[14]

CSU Standard Normalization

Fig. 14. Examples of low-light imagery normalized with the top row showing the standard equalization approach and the bottom row showing Securics' new dual-LUT approach.

The recognition algorithm being used is the Biotope algorithm developed by Dr. Boult[15]. The core algorithm is based on a PCA algorithm, but with a significantly modified distance metric, to make it far more robust than prior PCA-based face algorithms. The performance on the standard FERET test sets shows this algorithm performing at or above all previously published results on that dataset. Mapping it to embedded systems hardware, including the formal fixed-point analysis is underway.

6. CAMERA ARCHITECTURE WITH FPGAS

The SI-VGA60-EM camera has been designed for 24/7 automated image capture and embedded target detection discrimination and tracking. The camera incorporates high performance FPGAs for developing the precise timing used to drive the EMCCD sensor and maximize performance across a broad range of illumination conditions as well as to process the images. This will be achieved by embedding the algorithms described in the previous section in on-board

FPGAs to perform on-the-fly processing for both image enhancement and detection. A block diagram of the camera is shown in figure 15.

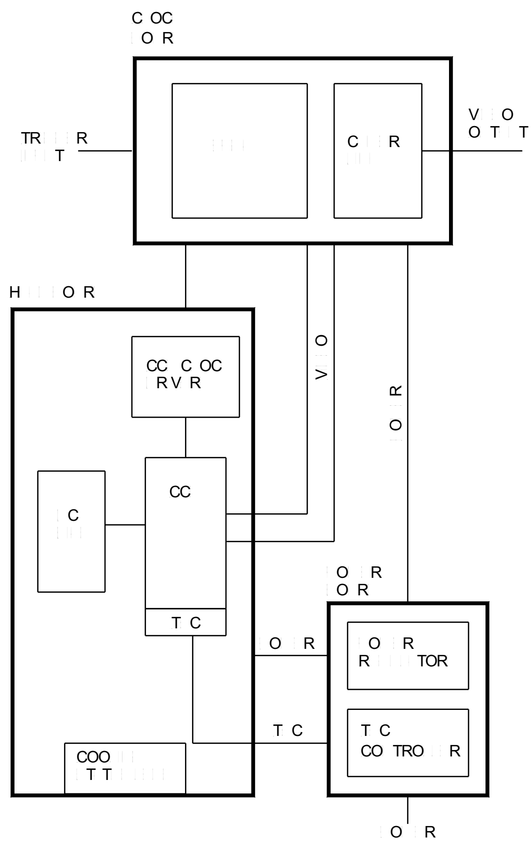

Fig. 15. SI-VGA60-EM camera architecture showing FPGAs for on-board image processing (need to add GigE board with frame buffers and Vertex FPGA)

The FPGA utilized on the clock I/O board is an Altera Stratix II. Image enhancement functions such as AGC, image filtering, and image feature pre-processing are being implemented on this FPGA. The image processing board includes a Xilinx Vertex IV FPGA along with frame buffer space to perform more complex image processing functions as well as target detection, tracking and recognition functions. The board also includes the circuitry required for presenting the data over a Gigabit Ethernet connection, using the GigE Vision standard. This enables the use of multiple cameras, each with its own IP address.

7. CONCLUSIONS – INTEGRATION OF A SECURITY IMAGING SYSTEM

A security imaging system has been presented based on the combination of advanced low light image sensors and innovative target detection, discrimination and identification algorithms. The low light camera uses the features of electron multiplying CCDs including programmable gain, anti-blooming, color, and uncorrelated noise to produce high quality images over a broad range of illumination conditions for bright sunlight to overcast starlight. The incorporation of algorithms which can detect targets of interest amongst a cluttered background can produce low false alarm rates under real world imaging conditions. Face recognition algorithms have also been developed permitting identification under non-ideal imaging conditions with non-cooperative subjects. These algorithms have been optimized for real time operation on FPGAs. This capability will allow the combination of the 24/7 image capture with on-board processing, to dramatically reduce the data transmission requirements in a system. and enable a scalable multi-camera security system.

By having cameras with automated image enhancement and on-board recognition capabilities, connected by Gigabit Ethernet to a central monitoring station enables a truly scalable multi-camera security system. This system can be enhanced with sensors providing complementary information such as thermal IR, NIR, polarization, imaging radar, lidar, etc. The cameras can be mounted on poles, combined with pan-tilt-zoom, or 360^o mirrors or mounted on UAVs and UGVs. These capabilities allow us to approach the goals of ideal security imaging systems for a broad range of defense and homeland security operational scenarios.

REFERENCES

1. P. Jerram et al, “The LLLCCD: Low Light Imaging without the need for an Intensifier”, SPIE Vol 4306, 2001.
2. Donal J. Denvir, Emil Conroy, “Electron Multiplying CCDs”, www. Andor-tech.com.
3. J. Hynecek, “Impactron – A New Solid-State Image Intensifier”, IEEE Trans. On Electron Devices, Vol. 48, No. 10, Oct. 2001.
4. C. D. Mackay et. Al. “Sub-Electron Read Noise at MHz Pixel Rates”, SPIE 4306 Conference Proceedings, 289-298, Jan. 2001.
5. Texas Instruments Data Sheet for the TC-246RGB-B0.
6. Robbins and Hadwen, “The Noise Performance of Electron Multiplying Charge Coupled Devices”, IEEE Transactions on Electron Devices, Dec 2002.
7. www.andor.com/library/glossary/index.asp?app=326
8. J.Kriesel and N. Gat, “Performance Tests of True Color Night Vision Cameras”, Military Sensing Symposia (MSS) Specialty Group on Passive Sensors, Orlando, FL, Feb. 2006.
9. T. E. Boult ref for colored faces
10. R. Tippets ref
11. T.E. Boult, R.J. Micheals, X. Gao and M. Eckmann, ``Into the woods: visual surveillance of non-cooperative and camouflaged targets in complex outdoor settings'' , the Proceedings of the IEEE, Oct 2001.
12. P.Viola and M.Jones. “Robust real time object detection,” ICCV Workshop on Statistical and Computational Theories of Vision, Vancouver, Canada, 2001.
13. R. Woodworth, “Object Detection In Realtime Video Using Verilog VLSI with Haar Cascade Classifiers”, MS Thesis, U. Colorado at Colorado Springs. Technical Report EAS_ECE_2006_10. 2006.
14. Bolme, D.S., Beveridge, J.R., Teixeira, M. and Draper, B.A. (2003). “The CSU face identification evaluation system: its purpose, features, and structure”. ICVS 2003: 304-313
15. T.E. Boult, “"Robust distance measures for face recognition supporting revocable biometric tokens", IEEE Conf. on Face and Gesture, April 2006