KIM T. CONSTANTIKES

POINT DETECTION USING A CONING SCAN IMAGER

A very compact and inertially pointed imaging device can be constructed by combining the functions of a telescope and gyroscope into a single assembly. However, the image resulting from this device is not easily processed owing to scan-induced geometric distortions. We have devised a method for adaptively processing the imager outputs to facilitate detection of bright points in a cluttered background.

INTRODUCTION

A very compact and inertially pointed imaging device can be constructed by combining the functions of a telescope and gyroscope into a single assembly (Fig. 1). The principal axis of a spinning telescope mounted on a pair of gimbals will point in a fixed direction in inertial space in the absence of external torques. Optical scanning of the imaging device's object space is accomplished simply by creating an imbalance in the telescope so that the optical axis is rotated from the principal inertial axis of the telescope. Thus, the optical axis will move with a coning motion around the principal axis, and the image formed at the focal plane of the telescope will be scanned circularly in the object space. A linear photodetector array is fixed to the inner gimbal plane, which is collocated with the focal plane. When the entire gimbal apparatus is rotated, the photodetector array will scan the object space in a somewhat complicated way.

Scan pattern complexity is the price one pays for compact, mechanically simple, and stable imaging. The image that results is continuously distorted by the variation of the photodetector array's velocity, as seen in the inertial object coordinates. We will refer to this distorted image as a pseudoimage. Our challenge has been to devise a signal processing method that can effectively detect point objects in various cluttered backgrounds and to do so with minimal computational resources.

We have devised such an approach. The scanning geometry of the imager is exploited by applying a sequence of one-dimensional (1D) signal processing algorithms to reduce image data, and hence data rate, to a single-bit pseudoimage consisting of point detections with an adequately high probability of detection. A scan-adaptive pattern recognizer then processes these data to reduce false alarm rates to acceptable levels. A desirable side effect of the approach is that 1D matched filters can be tailored to particular photodetector positions, and thus variation of optical characteristics across the focal plane can be accommodated.

We will first discuss the image formation process of the coning imager. Then we will look at alternative methods for applying two-dimensional (2D) signal processing algorithms to pseudoimages. Finally, we will present a somewhat more detailed description of our approach and an experiment that established the basic plausibility of the method.

IMAGE FORMATION

Scanning

The scanning optics are designed with a linear photodetector array mounted on the focal plane. The ray propagated from the center of a detector, through the telescope, and to a scene object is described by a look vector $\mathbf{r}$ in the coordinates fixed in the focal plane as follows:

$$\mathbf{r}^{(f)} = \left[1, -\frac{y(n)}{f}, -\frac{z(n)}{f}\right] \equiv \left[1, -\beta_y, -\beta_z\right], \qquad (1)$$

where $y(n)$ and $z(n)$ are the focal plane coordinates of detector n, f is the (back) focal length of the telescope, and (f) is the focal length coordinates. We assume that y and z are small compared with f, and thus the vector components are approximately the direction cosines of the unit look vector.

To determine how the scene is scanned by the photodetectors, one must identify the components of $\mathbf{r}^{(f)}$ in the inertial coordinates of the scene (Fig. 2). Since the telescope's axis is at a small angle with respect to its principal inertial axis, the gimbals execute small sinusoidal motions of magnitude of the coning angle (Fig. 1). In this case, the gimbal angles as a function of time and constant telescope spin, and detector rotation rates can be determined from small angle approximations that are then used to transform the look vector back to the inertial coordinates,[1] resulting in

$$\beta_y^{(i)} = \cos\alpha_2 \sin\alpha_1 \cos\theta_1 + \sin\alpha_2 \cos\theta_1 - \beta_y^{(f)} \cos\alpha_1 \cos\theta_1 + \beta_z^{(f)} (\cos\alpha_2 \sin\theta_1 - \sin\alpha_1 \sin\alpha_2 \cos\theta_1) \qquad (2a)$$

and

$$\beta_z^{(i)} = \cos\alpha_2 \sin\alpha_1 \sin\theta_1 - \sin\alpha_2 \cos\theta_1 - \beta_y^{(f)} \cos\alpha_1 \sin\theta_1 + \beta_z^{(f)} (\cos\alpha_2 \cos\theta_1 + \sin\alpha_1 \sin\alpha_2 \sin\theta_1), \qquad (2b)$$

where

$\alpha_1 = \theta_5 \sin\theta_4$, where $\theta_4 = \omega_4 t$ (ω_4 = the telescope's angular rate, t = time) and θ_5 = the coning angle,

$\alpha_2 = \theta_5 \cos\theta_4$

$\theta_1 = \omega_1 t$, where ω_1 = the photodetector array's angular rate, and

(i) = the inertial coordinates.

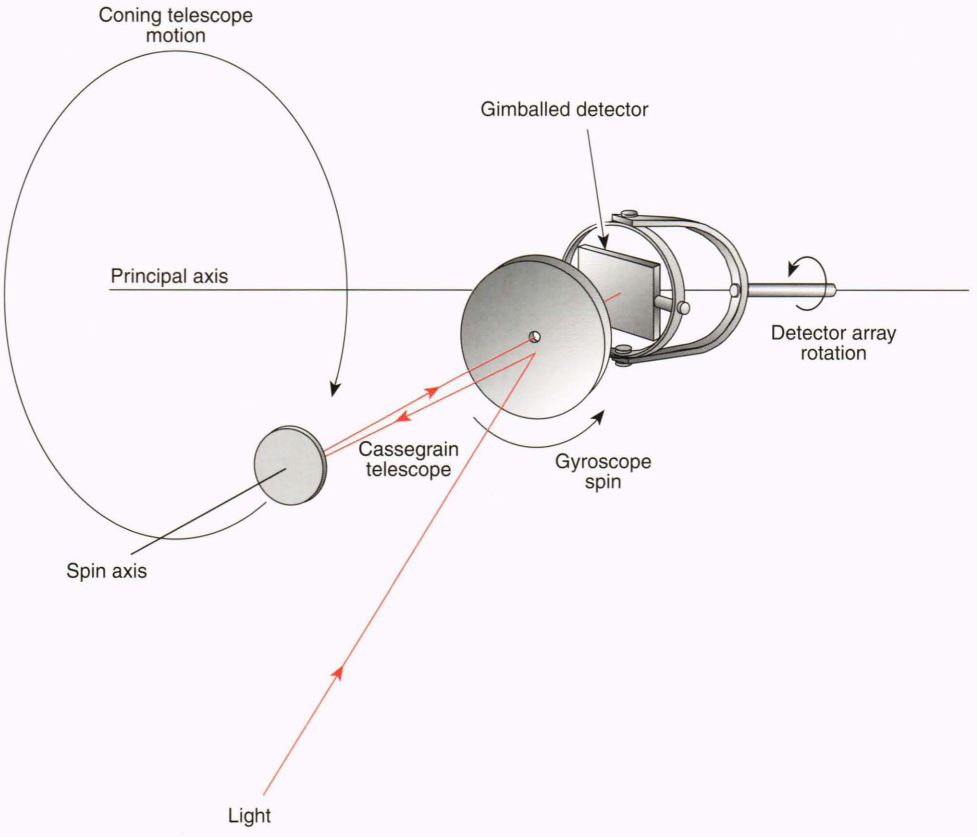


Figure 1. The coning scanner. The telescope of the imaging device is also a gyroscope mounted in two gimbals. By creating a slight imbalance in the telescope, the optical axis is rotated away from the inertial axis, which results in a coning motion of the telescope. Hence, the detector array mounted at the focal plane is scanned over the object space or scene.

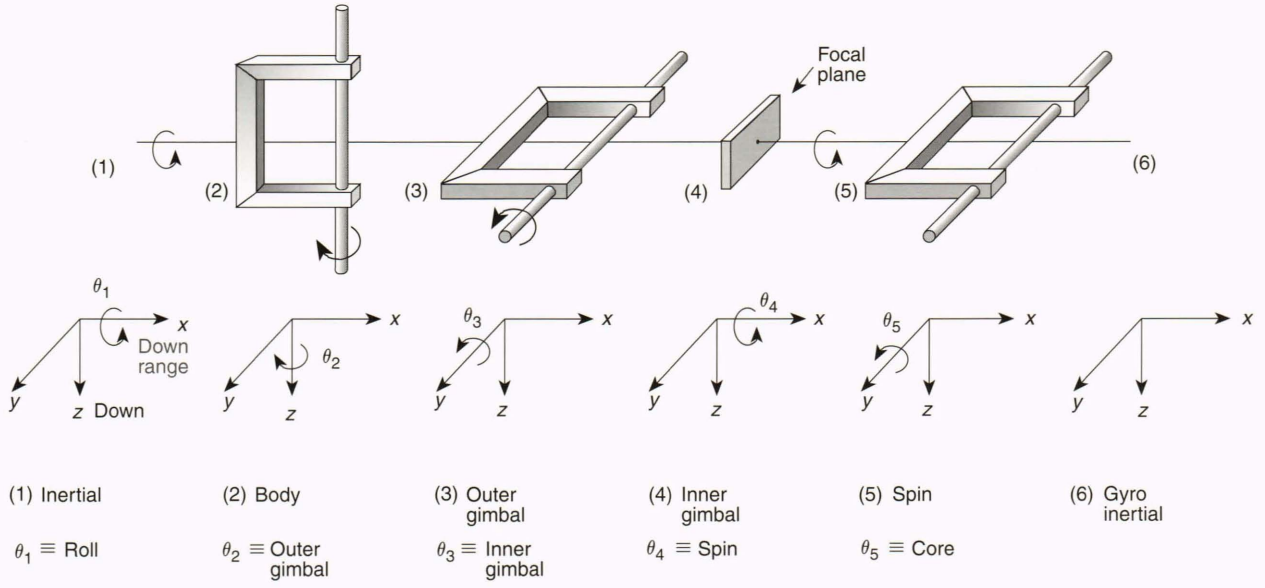


Figure 2. Imager coordinate transformations. To understand how the photodetector array scans the scene, the direction cosines of the look vector **r** in the focal plane coordinates must be transformed to inertial coordinates. If we constrain the angular velocity of the telescope's principal axis and the imager's roll axis to be along the x_1 direction, the gimbal angles θ_2 and θ_3 are small sinusoidal perturbations for small coning angles θ_5. The focal plane coordinates (4) are reflected about the *xy* and *xz* planes to represent the inverted image and transformed back to inertial coordinates (1) to derive the scanning relationship (Eq. 2).

We can apply these relationships to visualize how the imager samples object space and time. In Fig. 3, the locus of scene angles sampled over one roll cycle is shown as a ribbon in angle space and time. Although the expressions given in Eq. 2 are suitable for a global analysis of the scan pattern, investigating the local properties of the sample lattice is also useful. The velocity of the array in angle space can be calculated as[2]

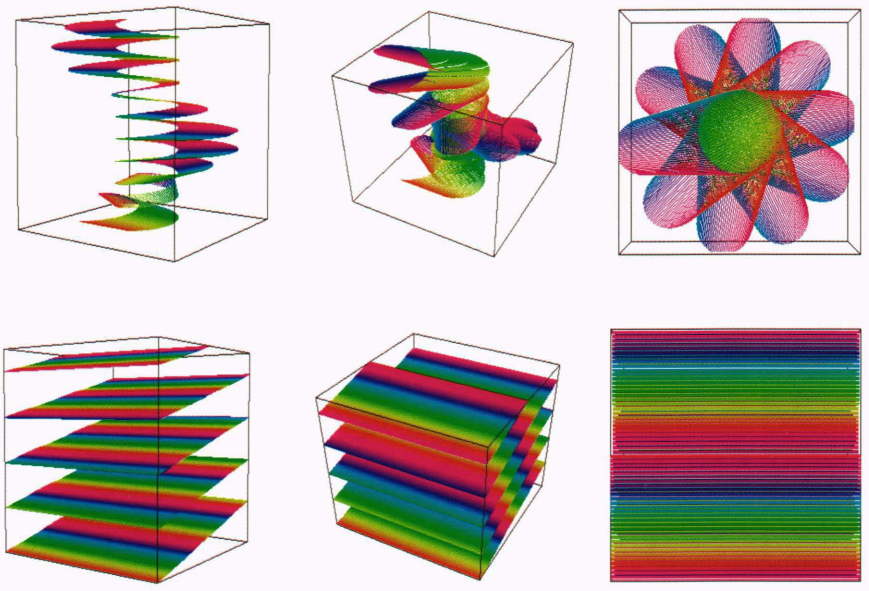


Figure 3. Sampling locus in angle space and time for one roll cycle. (Top) Three views of the locus of samples in angular coordinates versus time illustrating the inefficiency of the scanning mechanism. In this example, the spin rate is 10 times the photodetector's rotation rate. Note the 10 lobes in the view on the right. (Bottom) Three views of the locus of samples from a line scanner with similar photodetector array structure and sampling rates. Colors are used to encode the time at which the array is sampled. Red to red represents one complete telescope spin.

$$\frac{d\mathbf{r}^{(i)}}{dt} = \frac{d\mathbf{r}^{(f)}}{dt} + \omega \times \mathbf{r}^{(f)} = \omega \times \mathbf{r}^{(f)} , \qquad (3)$$

where ω is the angular velocity of the focal plane coordinate system in the inertial coordinate system. Using this expression and the gimbal rates derived in Eq. 2, we get

$$\frac{d\beta_y}{dt} = -\omega_1 \beta_z + \theta_5 \omega_4 \cos\left[(\omega_1 + \omega_4)t\right] \qquad (4a)$$

and

$$\frac{d\beta_z}{dt} = \omega_1 \beta_y + \theta_5 \omega_4 \sin\left[(\omega_1 + \omega_4)t\right] . \qquad (4b)$$

Figure 4 shows the velocity of a linear array in the inertial coordinates. Note that the magnitude squared of scan velocity is a function of scan angle as follows:

$$\left|\dot{\beta}\right|^2 = \frac{\beta_y^2 \omega_1^2 + \beta_z^2 \omega_1^2 + 1}{\theta_5 \omega_4} + 2\omega_1 \left\{\beta_y \sin\left[(\omega_1 + \omega_4)t\right] - \beta_z \cos\left[(\omega_1 + \omega_4)t\right]\right\} ; \qquad (5)$$

hence, a variation of scan velocity occurs at the detector's rotation plus spin frequency, which is worst at the end of the linear array. The effect of this variation is a scan-angle-dependent warping of the time scale that causes scale distortions in the single photodetector outputs. The magnitude of the effect is about ±9% for a nominal selection of parameters. One-dimensional matched filters must be designed to accommodate this variation with accompanying performance loss, which is probably minor.

Matched Filtering

One common and robust approach to detecting a signal in noise is to apply a matched filter to the data and then threshold the output of the filter. If the filter output exceeds the threshold, we accept the hypothesis that the signal is present in the noise. Matching the filter to the signal is accomplished by finding the filter that maximizes output signal-to-noise ratio, that is, the filter output with signal present divided by output power when the signal is not present. Under relatively general conditions, it can be shown that the matched filter has the time-reversed impulse response of the signal we are looking for. We implement a matched filter by correlating the noisy signal with the true signal. If the noise itself is correlated, the optimal filter is preceded by a filter that decorrelates the noise. Under somewhat more restrictive conditions, we can demonstrate that the matched filter and threshold detection can also implement the optimal minimum error or constant false alarm rate detector. The work of Turin[3] is an excellent tutorial on the subject.

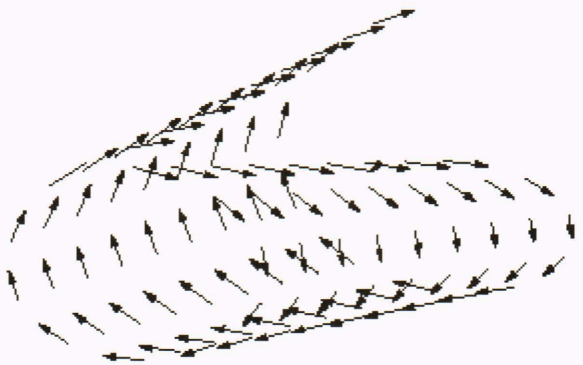


Figure 4. Photodetector array velocity in inertial coordinates for one spin cycle. Each arrow represents a detector, and the length and direction show the motion of the detector over about 25 sample times. The array moves parallel to itself twice during one telescope spin, and the scan velocity rotates around the array's long axis.

Transfer Function

A net transfer function for the imaging apparatus can be conveniently derived. This function is represented as a convolution kernel and yields, when applied to the radiance at the imager dome, the output of a detector located at some position on the telescope's focal plane. This convolution kernel is in turn the convolution of the optical point spread function[4] (PSF) g and the detector response function h, which is a circular aperture function multiplied by diode responsivity.[5] Let $\{t_n\}$ be the times at which the charge accumulated on the photodetector apparatus during an interval Δt is sampled. The charge Q accumulated will then be

$$Q(\boldsymbol{\beta}_n; t_n) = \left[g(\boldsymbol{\beta}; \boldsymbol{\beta}_n) * h(\boldsymbol{\beta}) * \delta(\boldsymbol{\beta} - \boldsymbol{\beta}_n - \dot{\boldsymbol{\beta}}_n t) * R_{\Delta t}(-t)\right] * L(\boldsymbol{\beta}), \quad (6)$$

where

$$\boldsymbol{\beta} = \left[\beta_y, \beta_z\right],$$

$$\boldsymbol{\beta}_n = \boldsymbol{\beta}(t_n),$$

$$\dot{\boldsymbol{\beta}} = \frac{d\boldsymbol{\beta}}{dt},$$

$R_{\Delta t}(t) = 1$ for the interval between 0 and Δt (and 0 otherwise),

L = the scene radiance in direction $\boldsymbol{\beta}$, and

* denotes a convolution.

The scanning geometry is such that a detector is located at a fixed radius with respect to the optical axis. The detector will have an associated rotation with respect to coordinates fixed in the telescope, and this angle will change at the telescope's spin rate. Since the PSF of the telescope will likely not be uniform over the focal plane, we can write the PSF in product form as

$$h(\boldsymbol{\beta}) = \mathrm{P}\left(\sqrt{\beta_y^2 + \beta_z^2}\right)\Phi\left[\tan^{-1}\left(\frac{\beta_z}{\beta_y}\right)\right], \quad (7)$$

where P is the PSF component dependent on radial position, and Φ is the part dependent on angular position.

In an ideally symmetric optic (i.e., perfectly aligned and with perfectly symmetrical optical elements), the angular component will be a constant. This fact is worth noting, since in this case, although the PSF may vary over the focal plane, any detector at a fixed focal plane location will have a constant PSF associated with it.

IMAGE PROCESSING FOR ACQUISITION

The most obvious way to apply a 2D matched filter to imager outputs is to reconstruct the 2D image from the pseudoimage. An example of this is shown in Fig. 5, where Eq. 2 was used to sample and reconstruct an image, assuming a perfectly stable gyroscope. Aside from the issue of temporal separation of samples (see Fig. 3) and gyroscope drift, we can see that the original image is undersampled on the periphery of the reconstructed image and grossly oversampled in the central section. Before a filter could be applied, the reconstructed image would have to be interpolated at the edges, and some mechanism for combining oversampled data would need to be employed. Since application of the filter would require the entire reconstructed image, the latency of detection would be one detector rotation. This delay and the computational burden associated with reconstruction and interpolation are prohibitive obstacles. The direct approach in this case is thus unworkable.

An alternative would be to apply a matched filter in the photodetector-sample coordinates (i.e., to the pseudoimage in Fig. 5). Unfortunately, the impulse response of the system varies in the pseudoimage coordinates. Most of the variability is due to a changing sample lattice, whereas some is the result of a blurring component caused by changing motion (see Eq. 6). Figure 6 shows the correspondence between samples in the pseudoimage and locations in the scene. More to the point, the effect is shown in Fig. 7, where an impulse, convolved with a rectangular detector aperture function and Gaussian PSF to approximate Eq. 6, was sampled for different scan angles. A scan angle is the angle between the velocity of a detector (Eq. 4) and its position vector (Eq. 2). It is evident that the scan angle changes through 360° in about one detector array rotation. When we look at the images in Fig. 7, we can see that the impulse response changes considerably as a function of scan angle. Any local technique would need to take this relationship into account by using temporally variant matched filtering synchronized to the scan geometry. Again, this approach would be prohibitively complex for our purposes. A single-instruction multiple data stream (SIMD) architecture is being used for our front-end signal processing. Although temporally variant matched filtering could be implemented, given adequate resources, the SIMD architecture does not lend itself to such a technique.

An alternative signal processing method allows us to use a 1D matched filter and detection preprocessor followed by a scan-angle-dependent, 2D false alarm rejection process in the pseudoimage coordinate system. As we noted, any single detector in the array has a constant impulse response associated with it. Thus, we can implement slightly different matched filters for each detector, which is particularly easy in an SIMD architecture with multiple processors per detector, and then process the detected data to reject false alarms using the 2D context of the detection. This approach is attractive for several reasons, for it maps well to the SIMD computer, it can accommodate variation of the PSF over the focal plane, and it applies 2D signal processing in a scan-dependent way at a low-data-rate point in the signal processor.

1D MATCHED FILTER/2D RECOGNITION PROCESSOR

A block diagram of our approach is shown in Fig. 8. Each photodiode output is processed independently by a matched filter and a threshold detector. Thresholds are set adaptively on the basis of signal statistics to ensure

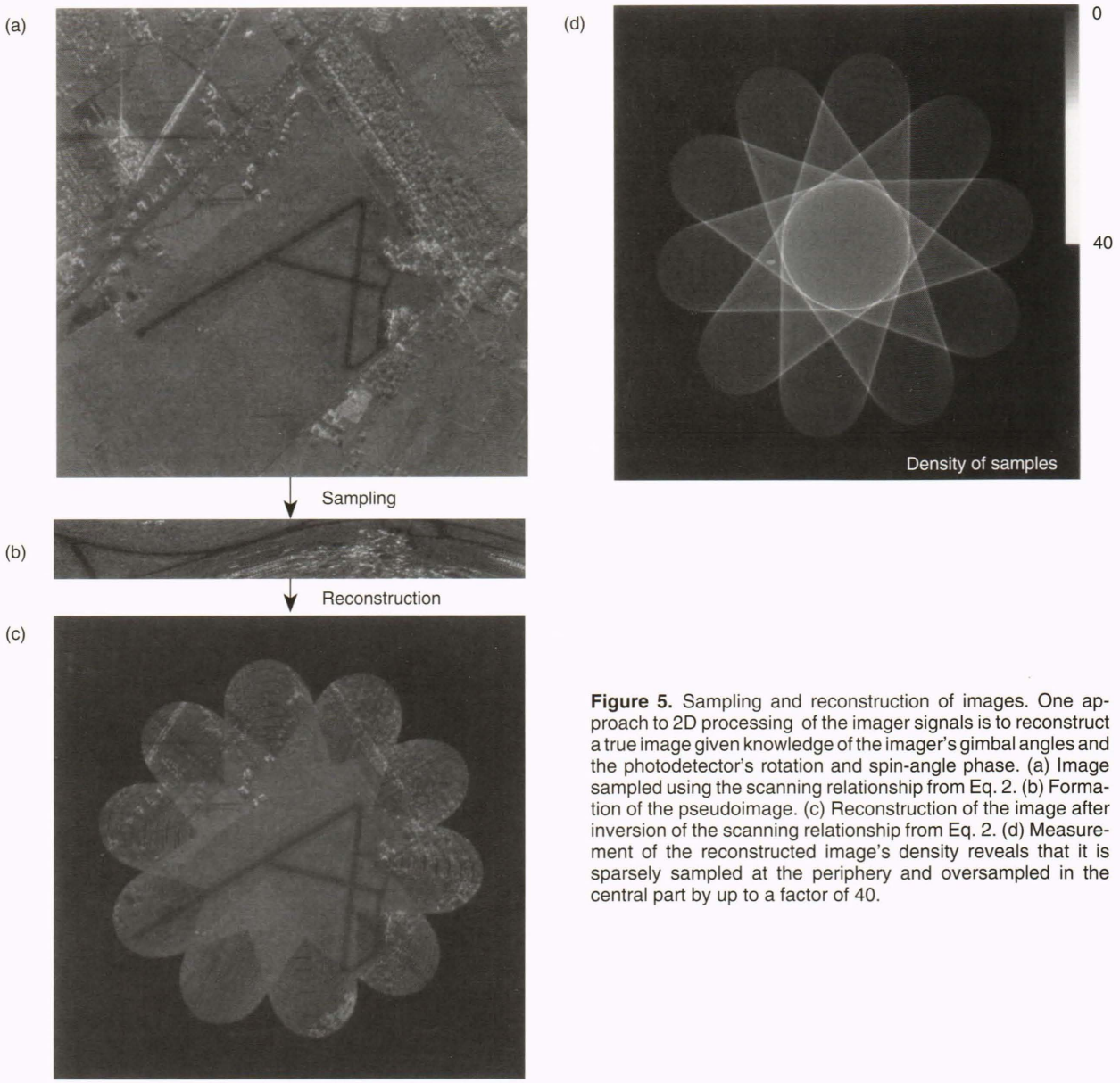


Figure 5. Sampling and reconstruction of images. One approach to 2D processing of the imager signals is to reconstruct a true image given knowledge of the imager's gimbal angles and the photodetector's rotation and spin-angle phase. (a) Image sampled using the scanning relationship from Eq. 2. (b) Formation of the pseudoimage. (c) Reconstruction of the image after inversion of the scanning relationship from Eq. 2. (d) Measurement of the reconstructed image's density reveals that it is sparsely sampled at the periphery and oversampled in the central part by up to a factor of 40.

adequate detection probability. The stream of single-bit output is then captured in a shift register, the neighborhood of a bit is extracted, and several bits of scan-angle data are appended to this data word. Then a pattern classifier is applied to determine which detections could have come from unresolved targets sampled with that scan angle. Finally, feasible detections are passed to additional algorithms that may apply further algorithms incorporating supplementary information relevant to the detection task. The later algorithms operate in a low-data-rate environment and can be considerably more complex than the those that preceded them.

The design of the matched filter and threshold is relatively straightforward, given the net impulse response of the imager and an assumption about the background clutter distribution function. The design and implementation of the 2D false alarm rejection network are somewhat more novel. Let us suppose that false alarm rejection on a neighborhood of three detectors and three temporal samples has been implemented and that the network is the same for all neighborhoods. If 4 bits of resolution in scan-angle data are adequate, the false alarm rejection becomes a mapping from a 13-bit vector to a single bit: 9 neighborhood bits and 4 scan-angle bits determine whether the detection was valid or a false alarm. This mapping can be realized as a memory containing 2^{13} bits, which is a relatively small memory.

Finding the best mapping is somewhat more complicated, as no reason exists to assume that vectors from valid detections are linearly separable[6] from false alarms. Further, as the size of the feature vector increases to a practical memory limit of about 2^{26} bits, it is not clear

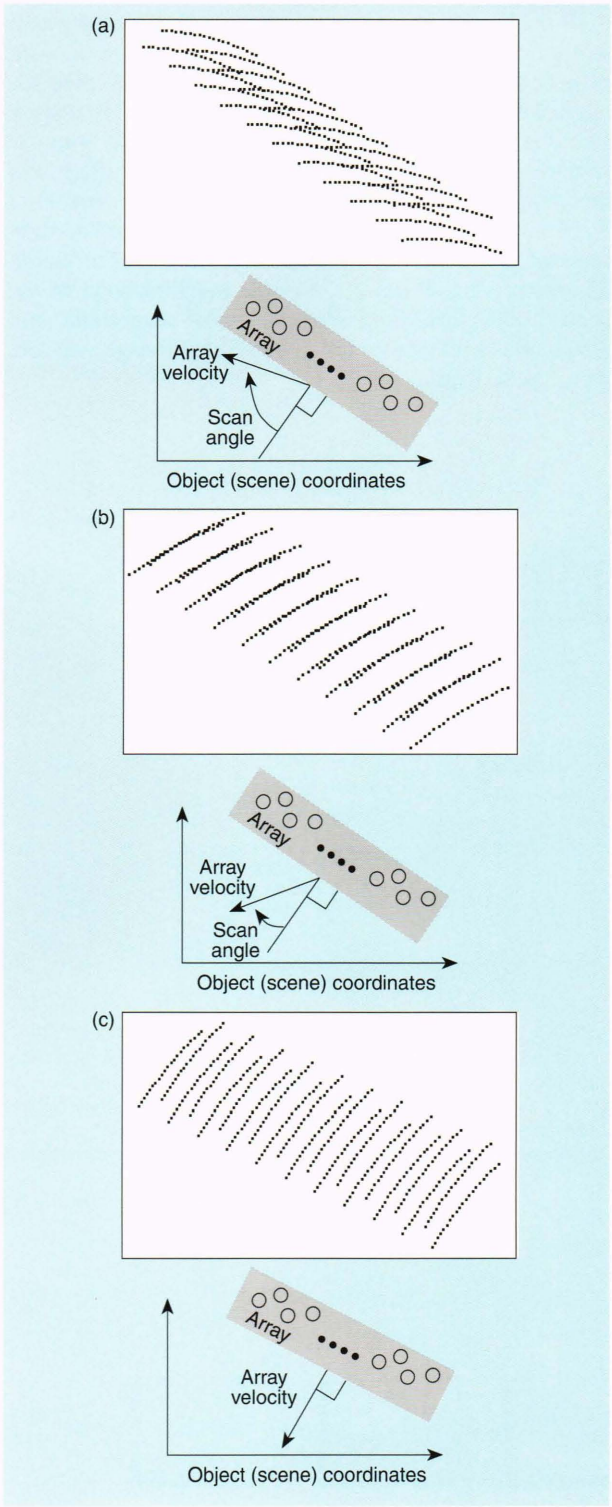


Figure 6. Sample points in scene coordinates. The scan-angle-dependent distortion of the sampling lattice is also evident in these portrayals of sample points in the scene coordinates. These are calculated using a staggered linear array geometry in which alternate detector diodes are shifted by 1/4 detector pitch spacing on either side of the array's long axis. Dots are placed for each time sample. In (a) the array moves along its axis, in (b) at about 45°, and in (c) about 90° from the long axis.

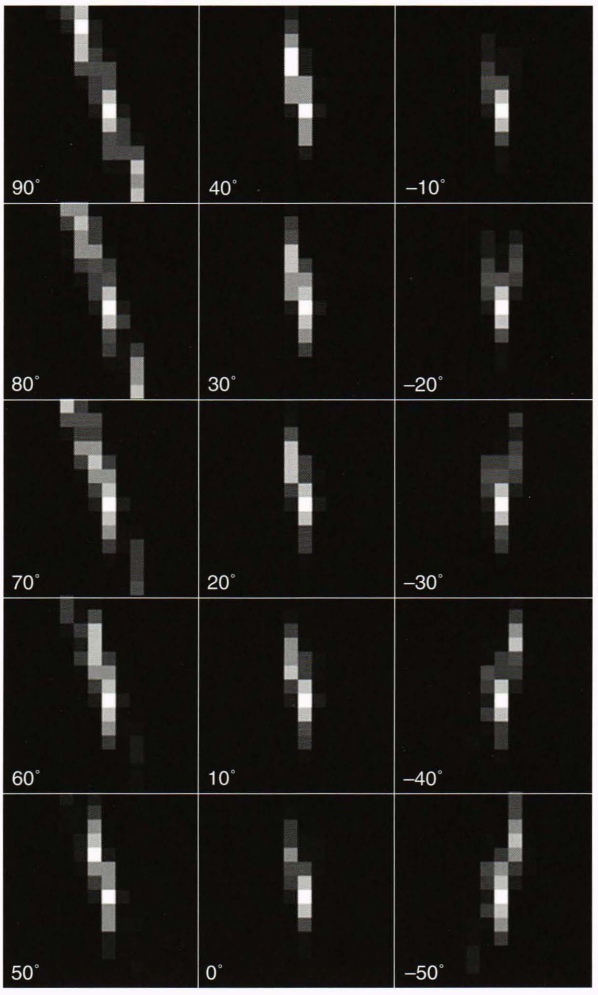


Figure 7. Impulse response as a function of scan angle. These images show the pseudoimage output of the scanner when it views an impulse in the scene. Equation 6 was approximated by a Gaussian PSF and rectangular aperture function computed at a resolution several times the array's resolution. The resultant image was sampled in such a way that the central pixel in each image corresponds to the instant the associated detector diode passed over the impulse for different values of the scan angle.

how one goes about exhaustively categorizing the input to output mapping. Thus, we need a supervised learning method that accommodates nonsimply connected and nonconvex decision regions and that can be used to generalize from a smaller training set to a relatively large working set. Although several different approaches can be taken (e.g., decision trees of linear classifiers, nonlinear mappings followed by linear classifiers, etc.), we chose to use a feed-forward neural network trained using a backpropagation algorithm.[7]

A PLAUSIBILITY TEST

False alarms from the 1D matched filter detector in a cluttered scene are in part due to the presence of edges

and lines in the clutter. These features appear to be impulses in a single photodiode's output. Thus, to demonstrate the feasibility of our signal processor, we applied the architecture to the detection of points, lines, and edges. If they proved to be distinguishable from each other, we had a reasonable hope that the approach would be usable in a more general environment.

We generated a training and testing data set by simulating imager response in a noiseless case, as diagrammed in Fig. 9. Images of an impulse, line, and edge at a spatial resolution 5 times the imager detector array pitch were convolved with a simplified form of Eq. 6. Then these images were sampled using Eq. 2 for 18 uniformly spaced scan angles and for 25 different phase relationships (i.e., subpitch positions of the point with respect to the detector array). These data were then filtered with the convolution of a whitening filter chosen as a 3-element Laplacian approximator and the impulse response of the imager. The output of the filter was thresholded, and output values exceeding the threshold were set to a binary 1. These data were ordered so that, for each phase and scan angle, a feature vector of length 32 was produced: a 5×5 sample neighborhood of the central pixel for 25 elements, 4 bits of scan angle, and 3 bits indicating which class (point, line, or edge) the data were taken from.

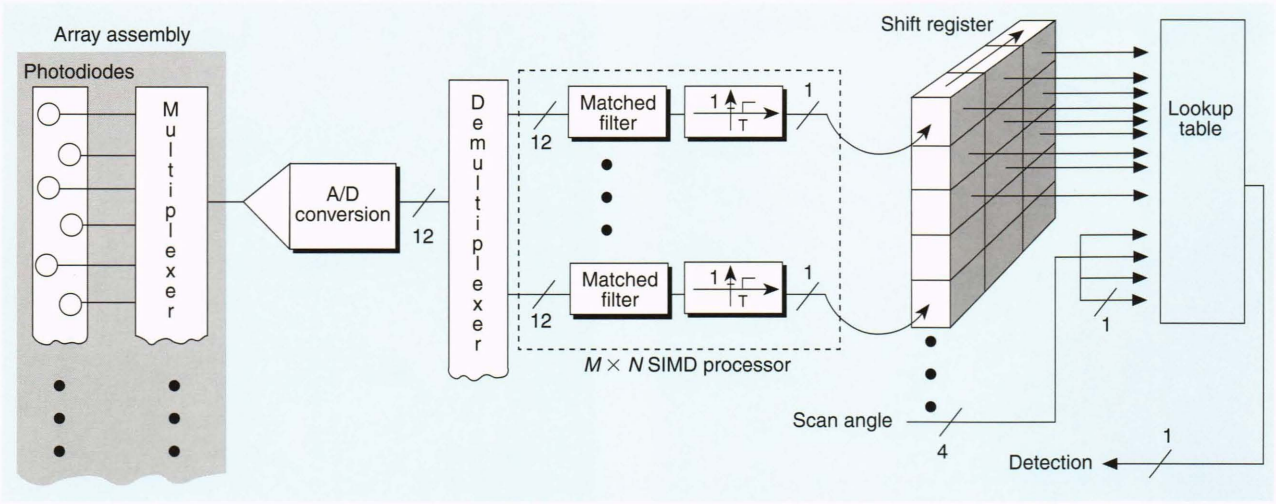


Figure 8. Acquisition processor block diagram. The analog photodiode outputs are demultiplexed and converted to digits at the sampling rate. The digital data are input to a single-instruction multiple data stream (SIMD) processor consisting of bit serial computational elements arranged in an $M \times N$ lattice, where M is the number of photodiodes and N is the number of taps in the matched filter to each photodiode, which may vary for each diode. On the basis of signal statistics, the filter output is thresholded using some detection criterion, such as constant false alarm rate. Output values that exceed the threshold are set to 1; otherwise, the output is 0. These data are delayed and multiplexed to construct L-bit neighborhood vectors, and a few bits of scan-angle data are appended. This sequence of vectors is then used to address a lookup table, which applies previously learned 2D false alarm rejection mapping, resulting in a stream of bits representing detections consistent with unresolved targets with that scan angle.

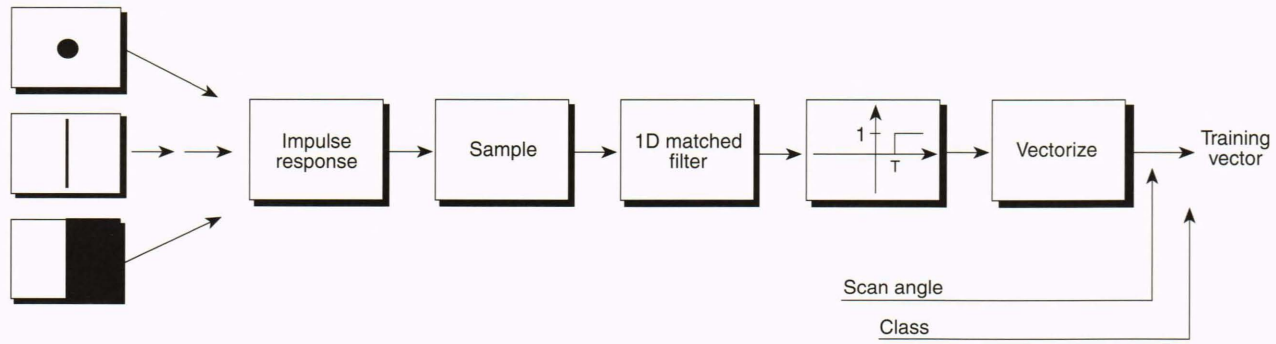


Figure 9. Training set construction. We constructed an impulse, edge, and line image at a spatial resolution about 5 times the photodiode's angular separation and then convolved them with the optical PSF and photodiode aperture function. These images were sampled by stimulating the scanning mechanism to create pseudoimage fragments 5 elements wide and 18 elements long. The signals were detected using a 1D matched filter and thresholded. The matched filter is the convolution of the PSF, the photodiode aperture function, and a Laplacian whitening filter. Finally, a 5×5 neighborhood of elements was vectorized, and 4 bits of scan-angle data and 3 bits of class (point, line, edge) data were appended. One of these vectors was generated for 25 subarray phases, 18 equally spaced scan angles, and each of the 3 classes.

The set of vectors was used to train a feed-forward neural net with 29 input nodes, 10 hidden nodes, and 3 output nodes. Although the input vectors are strictly binary, we did not use a Hamming net since we felt that a Hamming distance measure would inadequately weight the scan-angle data. Backpropagation training was able to separate the classes completely; no errors were made. Testing schemes such as "leave one out" to estimate error rates were unnecessary. A larger data set could probe error rate, but, for the purposes of our study, this result is adequate: rejection of false alarms caused by clutter is plausible using a neural net. The final step in an implementation would be to use the neural net to program the mapping memory by presenting all possible input vectors to the net and recording its output in the associated memory cell.

SUMMARY

Coning imagers are attractive owing to their compactness. The price of this compactness is that the pseudoimages they form are difficult to process optimally, and the imager's field of view is processed inefficiently, resulting in comparatively high image data rates, along with scan-angle-dependent distortions of the actual image.

Two-dimensional image processing is desirable for point detection, since extended objects can be rejected by the process. In principle, the actual image can be reconstructed from a pseudoimage, but the process is complicated and imposes a heavy computational burden. Local 2D filtering could also be used but would require scan-angle-dependent processing.

Implementing scan-angle-dependent filters would be difficult. A similar benefit may be achieved by first applying 1D matched filters and thresholds followed by a 2D pattern classifier on the single-bit data. This technique is considerably easier to implement, as the pattern classifier could be just a lookup table in a modest-sized memory. The approach has the added advantage that the radial part of the nonstationary PSF can be incorporated into the matched filters on a photodiode by photodiode basis, perhaps even as a part of the imager production process, since the radial dependence of the PSF will change from telescope to telescope.

We have demonstrated the plausibility of this approach by showing how a feed-forward backpropagation neural network can be trained to distinguish between true detections caused by point targets and false detections caused by lines and edges. The trained neural network can be used to generalize from a smaller training set into the complete working set implemented in the lookup memory.

The next task is to apply the method in a more general setting consisting of measured clutter images with radiometrically correct simulated point insertion, actual measured point and clutter data, and laboratory simulations using a prototype imager and some mechanism for point and scene synthesis. The Applied Physics Laboratory's dual mode facility would be an ideal environment for the simulations. We are reasonably confident that, in the general setting, our signal processing approach will provide performance near the optimal 2D case and with considerably less hardware and complexity.

REFERENCES

[1] Pio, R. L., *Algebra of Piograms*, M78-170, Hughes Aircraft Co. Radar Systems Group, Culver City, CA, pp. 6-52–6-57 (Oct 1978).

[2] Goldstein, H., *Classical Mechanics,* 2nd Ed., Addison-Wesley, Reading, MA, p. 174 (1980).

[3] Turin, G. L., "An Introduction to Matched Filters," *IRE Trans.* **IT-6** (Jun 1960).

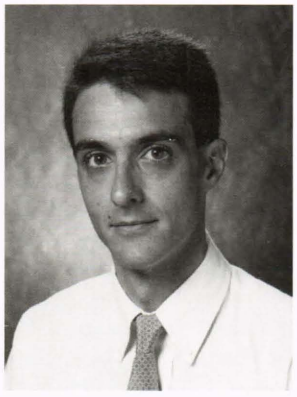
[4] Dudzik, M. C., *Electro-Optical Systems Design, Analysis, and Testing*, Vol. 4 of *The Infrared and Electro-Optical Systems Handbook*, J. S., Accetta and D. L. Shumaker (eds.), SPIE Optical Engineering Press, Bellingham, WA, p. 68 (1993).

[5] Boyd, R. W., *Radiometry and the Detection of Optical Radiation*, John Wiley & Sons, New York, p. 109 (1983).

[6] Duda, R. O., and Hart, P. E., *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, p. 138 (1973).

[7] Krogh, A., Palmer, R. G., and Hertz, J., *Introduction to the Theory of Neural Computation*, Addison-Wesley, New York (1991).

THE AUTHOR



KIM T. CONSTANTIKES is a Senior Staff engineer in APL's Fleet Systems Department. He received his B.S. degree in applied physics in 1983 from Ohio University. He then attended Carnegie Mellon University with a Benjamin Franklin Fellowship and received an M.S. degree in electrical and computer engineering in 1985. At Carnegie Mellon, he researched topics in pattern recognition and sensors in the Smart Sensors Laboratory of the Robotics Institute. Mr. Constantikes joined APL's Electro-Optic Systems Group in 1985, where he has worked on various computational and algorithmic aspects of imaging systems used for target acquisition, tracking, and recognition, as well as image-matching navigation. His current interests include image and scene modeling, image processing algorithms, image-matching navigation, target detection in clutter, computer graphics, and computational theory of vision.