

Application of integrated sensing and processing decision trees for target detection and localization on digital mirror array imagery

Carey E. Priebe, David J. Marchette, Youngser Park, and Robert R. Muise

We demonstrate the applicability of integrated sensing and processing decision trees (ISPDTs) methodology to a set of digital mirror array (DMA) hyperspectral imagery. In particular, we demonstrate that ISPDTs can be used to detect and localize targets by using just a few DMA Hadamard frames, so that an entire hyperspectral data cube need not be collected to successfully perform the given task. This suggests that such an integrated sensing–processing suite may be appropriate for extremely time-sensitive pattern-recognition applications. © 2006 Optical Society of America

OCIS codes: 100.2960, 100.5010.

1. Introduction

Modern sensors are designed to collect vast amounts of information, often more than can easily be transmitted or processed in time-critical applications. These sensors are often adaptive in the sense that they have adjustable parameters that determine the information that will be collected (for example, the spectral bands in a hyperspectral sensor). Several factors mandate that the sensor be tuned to produce only the information necessary for the task at hand: processing constraints, bandwidth constraints, and the curse of dimensionality. Ideally, the sensor system should adapt its collection based on previous observations. These observations can provide information about the current environment and the objects within the scene. By tailoring the collection to better disambiguate the specific hypothesized objects within the observed environment, the system can improve its performance on the observed scene. This allows for approximately optimal solutions on the

observed data without the need for globally optimal solutions.

One method considered for this problem is the use of integrated sensing and processing decision trees (ISPDTs).¹ As an illustration of the ISPDT approach, let us assume that a hyperspectral sensor can collect any k of a total of N possible spectral bands. The sensor collects an initial set of k bands, which have been chosen so that a representative set of data cluster into distinct groups in these bands. The idea is that these groups correspond to different environmental conditions, different scenes, or different noise distributions, depending on the physics of the specific problem. The k -dimensional observation is assigned to the nearest cluster. This cluster now determines the next k -dimensional observation to collect: either the best k spectra for detecting the target under the conditions defining the cluster, or the best spectra to allow for further refinement of the estimate of the environment.

This process is also known as iterative denoising,² because each successive clustering removes from consideration either noise observations (in the training) or invalid models (in detection or classification). As the sensor iterates the collect–cluster–collect procedure, it refines its model and ultimately selects the appropriate bands for the classification task. Note that these bands depend on the path through the tree, as does the classifier. Different leaves of the tree will not only (most likely) select different bands but will also use different training data in the design of the classifier.

Figure 1 illustrates the ISPDT. At the top left, two

C. E. Priebe (cep@jhu.edu) and Y. Park are with Johns Hopkins University, Baltimore, Maryland 21218-2682. C. E. Priebe is with the Department of Applied Mathematics and Statistics, Y. Park is with the Center for Imaging Science. D. J. Marchette is with the Naval Surface Warfare Center B10, 17320 Dahlgren Road, Dahlgren, Virginia 22448-5100. R. R. Muise is with Lockheed Martin Corporation, 5600 Sand Lake Road, Orlando, Florida 32819.

Received 10 August 2005; revised 5 January 2006; accepted 13 January 2006; posted 13 January 2006 (Doc. ID 64084).

0003-6935/06/133022-09\$15.00/0

© 2006 Optical Society of America

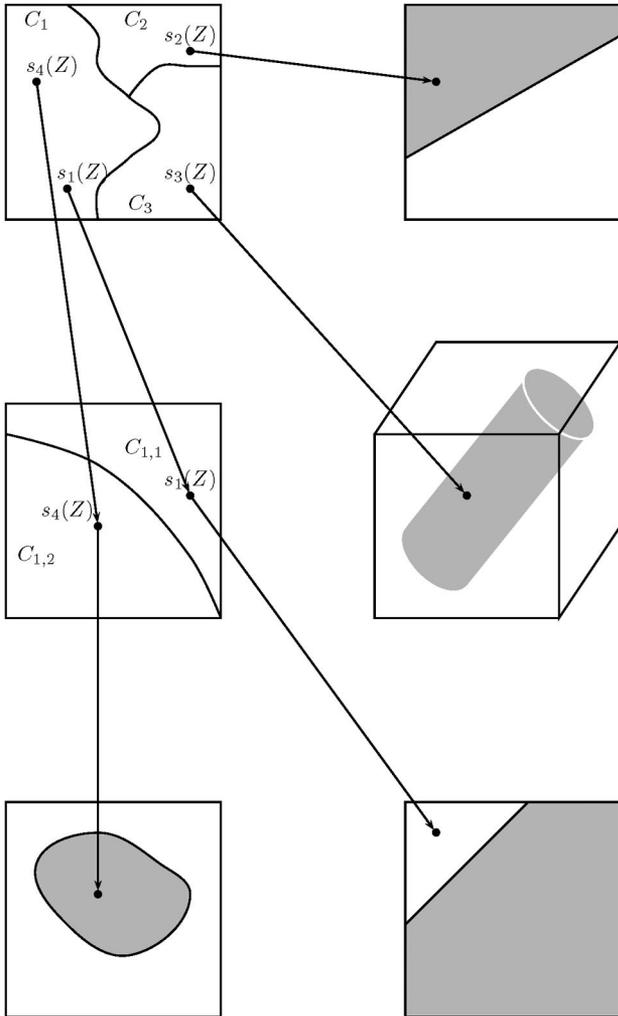


Fig. 1. Illustration of a simple ISPDPT consisting of two clustering nodes and four leaf nodes. The nodes are two dimensional except for the three-dimensional leaf node in the middle right image.

bands are collected, illustrated here as a two-dimensional plane. The training data are clustered within this space, resulting in a partitioning of the space into three sets: $\{C_1, C_2, C_3\}$. Any observation falling in C_2 results in the collection of two new bands and a linear classifier, as illustrated in the upper right-hand corner. An observation falling in C_3 results in the collection of three new bands and a quadratic classifier, as indicated in the middle right box. An observation falling in C_1 requires further processing. First, two new bands are collected, and the training has indicated that these should be partitioned into $\{C_{1,1}, C_{1,2}\}$, as indicated in the middle left figure. Then a further two bands are collected, resulting in either a linear classifier or a more complex classifier, as indicated in the bottom two figures. With each collection, a decision is made either to classify the observation or to partition the space and collect further information, as indicated by the partition in which the current observation falls. Thus observations $s_1(Z)$ and $s_4(Z)$ require three collections for a total of six bands, $s_2(Z)$ requires two collections for a

total of four bands, and $s_3(Z)$ requires two collections for a total of five bands.

In this paper we demonstrate the applicability of the ISPDPT iterative denoising methodology to a set of digital mirror array (DMA) imagery.³

2. Sensor and Data

A. Sensor

The sensor under consideration is a prototype system developed by PlainSight Systems, which incorporates a DMA device to realize a Hadamard multiplexed imaging system.

The known signal-to-noise-ratio (SNR) advantage in Hadamard spectroscopy⁴ extended to imaging systems^{5,6} allows for the collection of a hyperspectral cube of data with more efficient light collection than that of standard push-broom hyperspectral imagers.

The PlainSight sensor is a spatial light modulator-based multiplexing hyperspectral imaging camera, operable in the near-infrared spectral range of $\sim 900\text{--}1700$ nm. The system uses a DMA commercially available from Texas Instruments for projector display applications. The DMA contains 848 columns and 600 rows of mirrors and measures $10.2\text{ mm} \times 13.6\text{ mm}$. When the scene is illuminated on the DMA device, a standard raster scan could be implemented by turning the first column of mirrors ON, sending this column to a diffraction grating, which causes a spectral representation of the first spatial column of the scene to be illuminated on the detector array.

If one opens multiple slits in the DMA, the detector array will be presented with the superposition of many columns of spectra. This system would have the advantage of optimal SNR when the pattern of the open slits forms a Hadamard pattern.⁴ Each individual frame at the detector array has less physical meaning than in the push-broom method, but when all the patterns of the Hadamard sequence have been recorded, the full hyperspectral data cube is recoverable.

The PlainSight sensor implements the process wherein the detector array is a standard Indigo Phoenix large-area InGaAs camera operating in near-infrared wavelengths. During standard operation of the system, the sensor collects 512 frames. Each frame is 522×256 pixels and represents spectra versus spatial row. The 512 frames that are collected are based upon 256 Walsh (0's and 1's) patterns. Since the theory of optimal SNR is based upon Hadamard (1's and -1's) patterns, one needs to collect two Walsh patterns to generate a single Hadamard pattern. Thus the 512 collected frames represent the Walsh patterns required to form a full set of 256 Hadamard patterns. Since each column in the DMA array will hit the diffraction grating at a different location, the spectra will hit the detector array at a different location per column. We describe this as a skewness in the spectra, which results in the 522 pixels in the spectral dimension needed to represent the 266 actual spectral bins. Of course, this spatial-



Fig. 2. Hadamard frame 110 (full 256×256 image) for I_2 .

spectral mixing and skewness is invertible once all 256 Hadamard patterns have been collected. The resultant hyperspectral scene has dimensions of 256×256 with 266 spectral bands from 900 to 1700 nm.

The sensor has the capability of utilizing adaptive Hadamard frame-collection schemes as follows. The sensor is dynamically programmable in that it can be tasked to render any Walsh pattern at any scale and at any location on the DMA. For example, standard operation of the sensor renders 512 Walsh patterns from the lowest spatial frequency to the highest, at a resolution of 256 mirrors. In adaptive operation of the sensor, any one of these patterns can be rendered on the DMA; an image frame can be captured; some data processing can occur, leading to a decision regarding the next frame; and then any of the other patterns can be rendered. This sensor capability allows one to task the DMA to render patterns that have been algorithmically calculated to be optimal given the data that were previously observed. (This adaptive

sensor operation implies that the actual hyperspectral cube may no longer be recoverable from the frames collected, as one requires that the full 256 Hadamard patterns be collected to invert the spatial-spectral mixing caused by the multiplexing. If the desired processing does not require that the inverse operation be available, then it is more efficient not to collect all the data. We collect only what the algorithm needs to perform its function.) Thus the dynamic programmability of the sensor allows one the flexibility to sense only that data required for the image-processing task at hand.

Our task for integrating sensing and processing is to develop a scheme to collect these Hadamard frames one at a time and to make inferences about the current hyperspectral scene—and about the next Hadamard frame to be collected—as the frames are being sensed. Application may involve target detection, localization, and classification as well as object tracking.

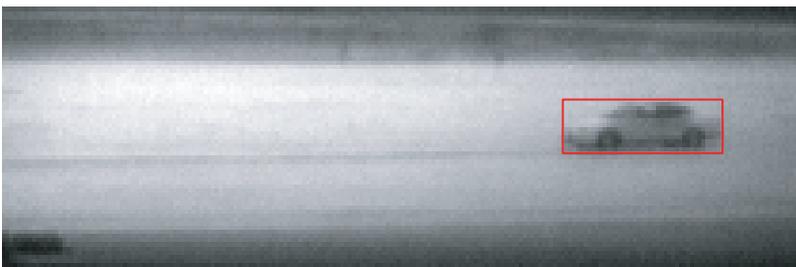


Fig. 3. (Color online) Hadamard frame 110 (64×210 swath) for I_2 with the target box used for training.

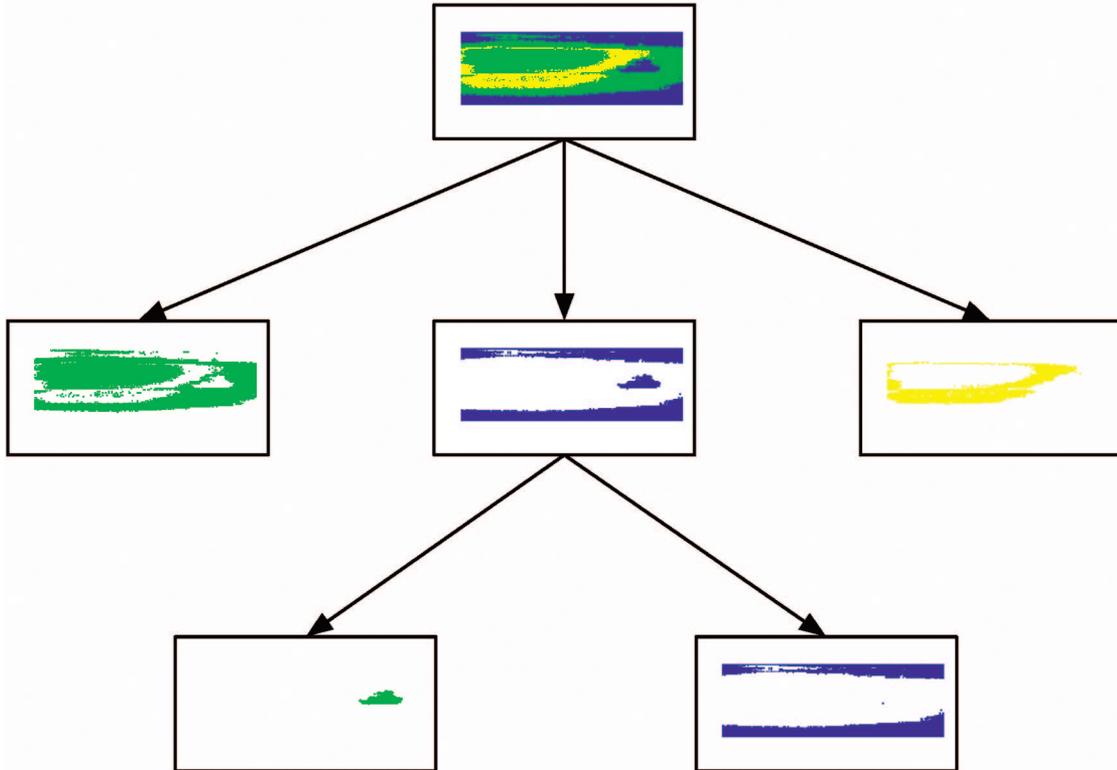


Fig. 4. (Color online) ISPDT constructed on I_2 (training). The leftmost node at level three of the tree is the target leaf.

B. Data

Data were collected at Lockheed Martin in Orlando, Florida, in April 2005. From a collection tower, a full set of 256 Hadamard frames was collected for a background scene (I_0) as well as for a target (a vehicle) moving from right to left at four time steps (I_t , $t = 1, 2, 3, 4$). Figure 2 depicts one Hadamard frame at time $t = 2$. This collection of five hyperspectral scenes is used to demonstrate the detection and localization capabilities of an integrated sensing–processing suite consisting of a dynamically programmable DMA sensor and ISPDT processing.

3. Methodology

We apply the ISPDT iterative denoising methodology to the hyperspectral vehicle imagery, I_t , $t = 0, 1, 2, 3, 4$, described above. We use the hyperspectral data cube collected at time $t = 2$, I_2 as our training data; Fig. 2 depicts Hadamard frame 110 at time $t = 2$. We focus on 64×210 image swaths so as to obviate issues of sensor skewness and edge effects; Fig. 3 depicts the swath for Hadamard frame 110 of the training data I_2 , with the region delineated by the small box representing the training information regarding the target of interest. (No additional information regarding vehicle–nonvehicle pixels within the target box is utilized.) Training proceeds as follows (see Fig. 4).

A. Step 1

No single Hadamard frame provides adequate performance in segregating target from nontarget pixels,

and so Hadamard frame 110 is selected for use at the root by virtue of its performance in providing the best clustering clarity—a new Hadamard frame at each node is determined based on the best separation between the target box and the nontarget pixels by using the adjusted Rand index criterion.^{7,8} Model-based clustering⁹ is employed in which a Bayesian information criterion (BIC) is used to determine the complexity and type of a Gaussian mixture fit to the data. The clusters are then defined in terms of the posterior likelihoods of the individual components. The Hadamard frame 110 pixels (gray-scale intensity) cluster into three clusters; the spatial locations of the pixels for these three clusters are represented by blue (middle), green (left), and yellow (right) in the root node of the tree depicted in Fig. 4 (color online).

B. Step 2

In building an ISPDT, after clustering at a node, each cluster is processed in a branch of the tree. This subsequent processing proceeds in an analogous fashion to that of the root: if the pattern-recognition task at hand can be adequately addressed (for the data falling to that branch), then tree growth (along that branch) is halted. It is necessary, of course, to perform the search for the best Hadamard frame once again in each branch, conditionally upon the results of all previous clusterings.

In this application the leftmost (green) and rightmost (yellow) clusters at the root of the ISPDT depicted in Fig. 4 yield essentially pure nontarget branches, and no further tree growth is necessary in

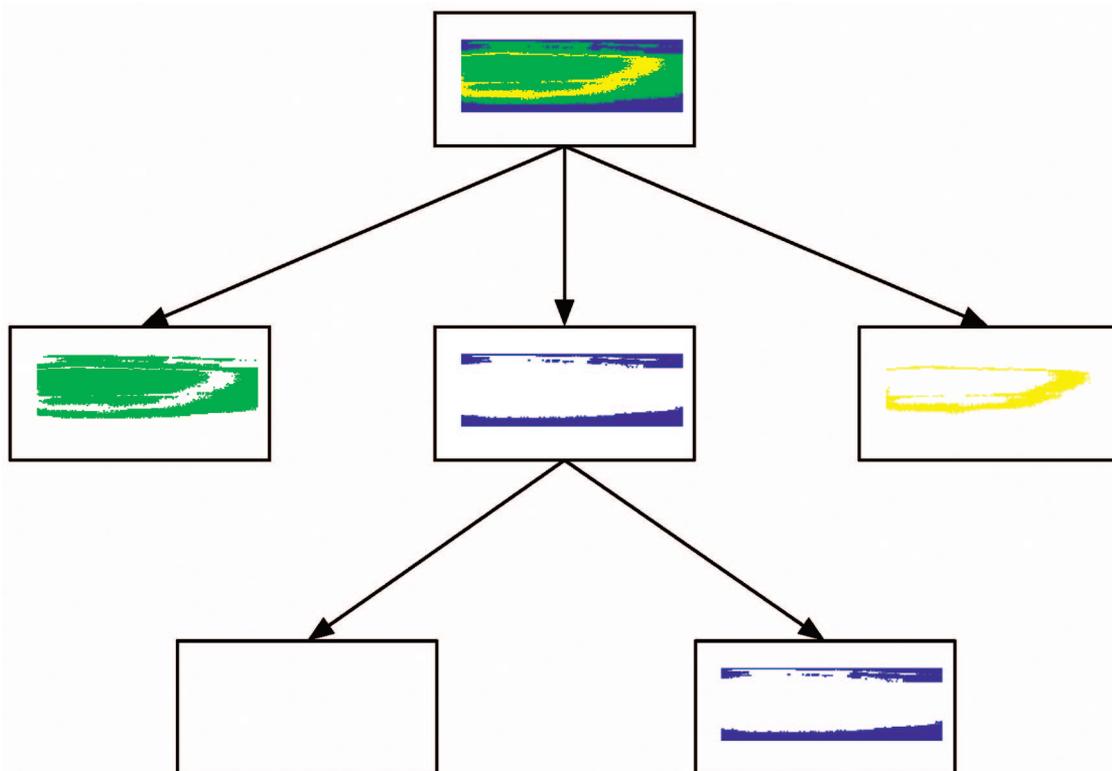


Fig. 5. (Color online) ISPDT detection-localization results for I_0 (testing). No vehicle is present at time $t = 0$, and no pixels falling into the target leaf implies no detection, as desired.

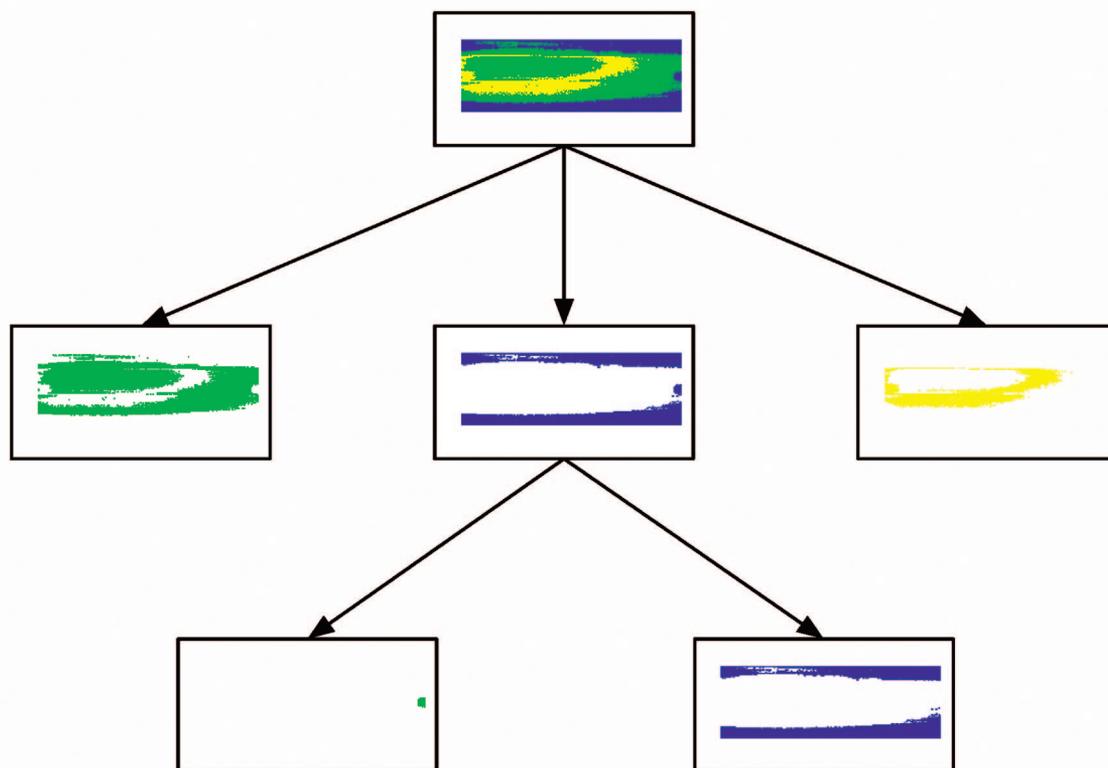


Fig. 6. (Color online). ISPDT detection-localization results for I_1 (testing). There is a vehicle present at time $t = 1$, and the existence of pixels falling into the target leaf implies detection, as desired. The spatial location of these pixels indicates that the detection is indeed on target.

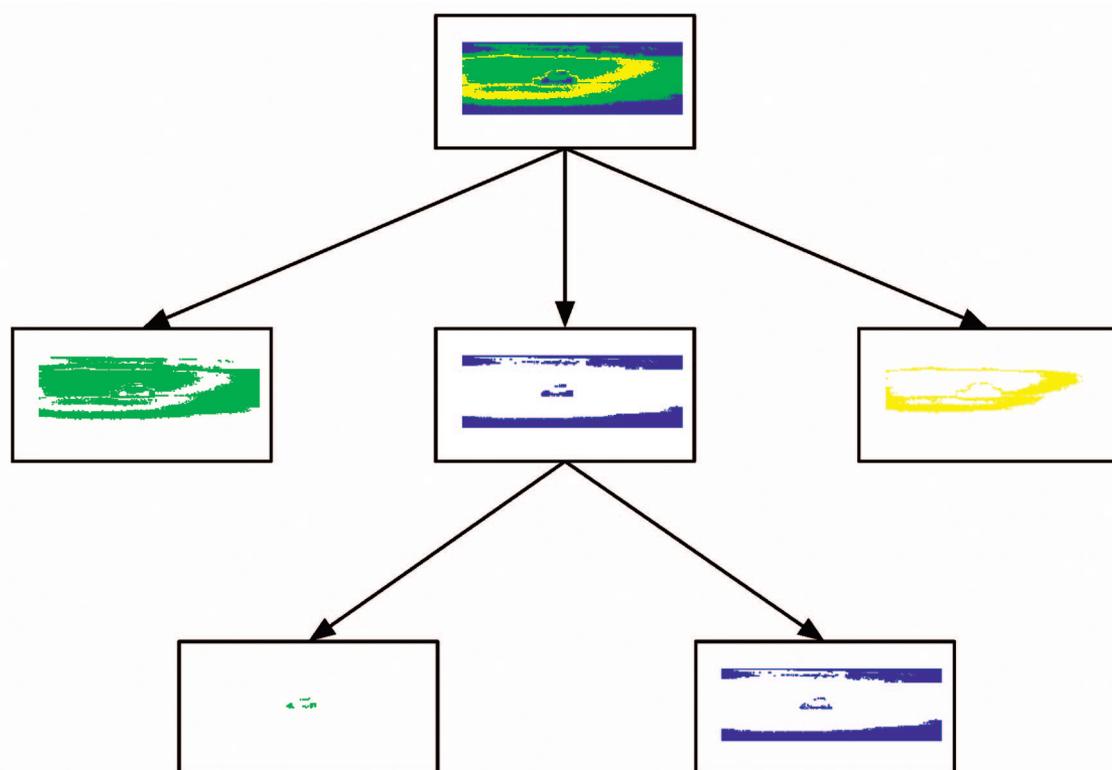


Fig. 7. (Color online) ISPDT detection-localization results for I_3 (testing). There is a vehicle present at time $t = 3$, and the existence of pixels falling into the target leaf implies detection, as desired. The spatial location of these pixels indicates that the detection is indeed on target.

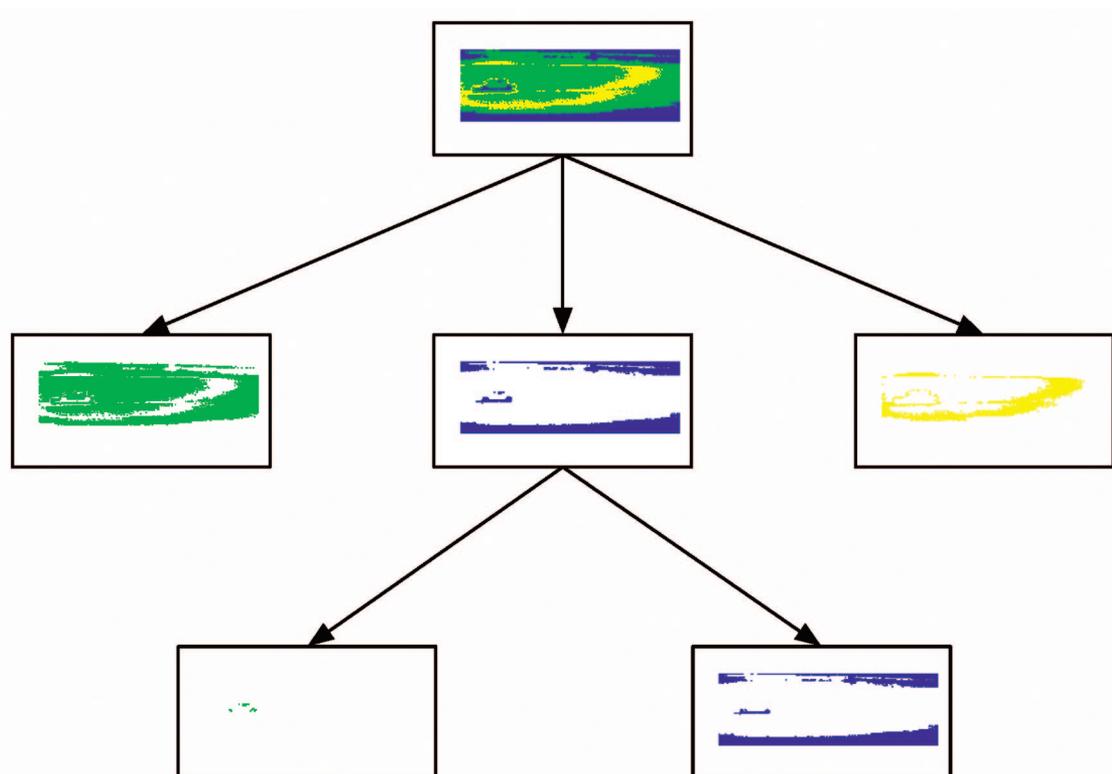


Fig. 8. (Color online) ISPDT detection-localization results for I_4 (testing). There is a vehicle present at time $t = 4$, and the existence of pixels falling into the target leaf implies detection, as desired. The spatial location of these pixels indicates that the detection is indeed on target.

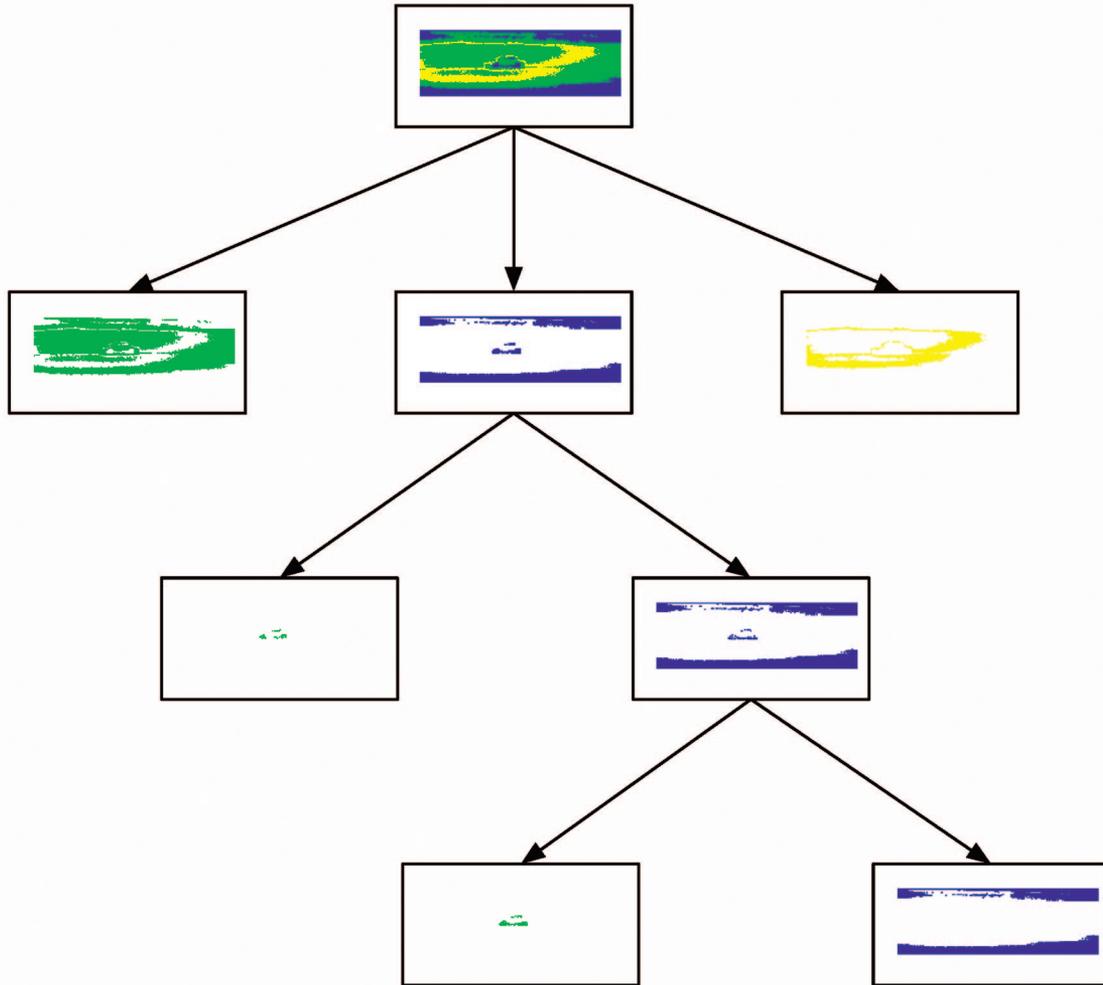


Fig. 9. (Color online) Illustration of a more elaborate ISPDT, requiring nontrivial conditioning and adaptation for I_3 .

these branches, while the middle (blue) root cluster contains many of the target pixels as well as a significant number of the nontarget pixels. This middle cluster must be processed further. In this middle cluster we find that Hadamard frame 151 provides clustering into two clusters; furthermore, one of these clusters (the leftmost node at level three of the ISPDT depicted in Fig. 4) contains a significant number of target pixels and no nontarget pixels. Ergo, this leaf is labeled as the “target leaf.”

Notice that, as was mentioned above, no information regarding vehicle–nonvehicle pixels within the target box is utilized. Nevertheless, the target leaf (the leftmost node at level three of the ISPDT depicted in Fig. 4) successfully—and unsupervisedly, with respect to the pixels within the target box—makes this distinction. The clustering at each level of the ISPDT can be used to produce a partition of the input space. Thus the ISPDT described above and depicted in Fig. 4 can be presented as

$$g(x) = I\{x \in C_{1,2} \wedge x \in C_{2,1}\}, \quad (1)$$

where $C_{i,j}$ represents the j th partition cell at level i of the tree and g is a classifier for pixel x .

Note that while the entire hyperspectral data cube was required for the training data, Eq. (1) allows processing for detection, localization, and classification with the collection of just two Hadamard frames, 110 and 151, as opposed to all 256.

The ISPDT depicted in Fig. 4 is quite simple. In general, for more complex applications, there may be more elaborate conditionality: for instance, depending on what is sensed in the first few frames, different choices may be required for subsequent sensing. Such a tree results in multiple target leaves and requires dynamic programmability of the sensor.

4. Results

We present successful detection–localization results on the vehicle imagery test data I_t , $t = 0, 1, 3, 4$, described above, by using the ISPDT trained on I_2 and depicted in Fig. 4 and Eq. (1). There is no vehicle present at time $t = 0$, and (see Fig. 5) no pixels fall

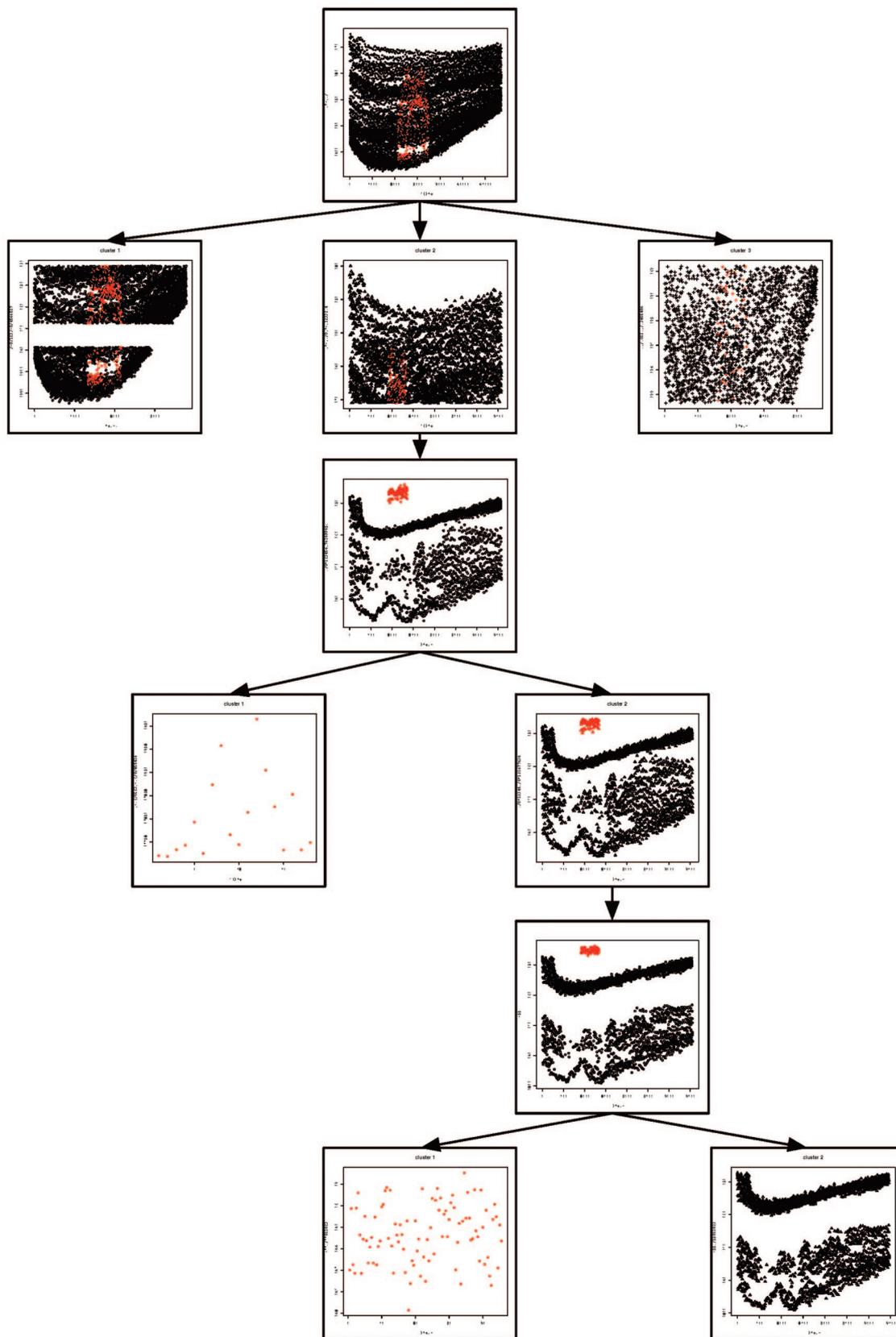


Fig. 10. (Color online) Scatterplot for the illustrative ISPDT depicted in Fig. 9. See text for description.

into the target leaf. This implies no detection, as desired. For $t = 1, 3, 4$, there is a vehicle present, and (see Figs. 6–8) in each case the existence of pixels falling into the target leaf implies detection, as desired. Furthermore, the spatial location of these pixels indicates that the detection is indeed on target.

Note: the fact that not all target pixels fall to the target leaf is inconsequential; the goal is detection and localization, and in each test case pixels on target are identified ($t = 1, 3, 4$) or no pixels are identified ($t = 0$) as desired.

Figure 9 (color online) depicts a more elaborate ISPDT, requiring nontrivial conditioning and adaptation, generated for I_3 . In this tree, as before, Hadamard frame 110 is used at the root, and in the middle cluster at level two, Hadamard frame 151 provides clustering into two clusters. While the target leaf—the leftmost node at level three—provides target detection, there is a significant number of target pixels that fall to the rightmost node at level three, a nontarget node.

When training on I_2 , tree building stops at this point. If training on I_3 (presented in Fig. 9), it is preferable to continue the tree building by using Hadamard frame 145 for additional accuracy. Thus the leftmost node at level four is an additional target leaf for this illustration, and the two target leaves together allow more target pixels to be identified (while still yielding no false detections). This tree would be realized by dynamically programming the sensor to collect Hadamard frame 145 only for cases in which frames 110 and 151 leave detection unresolved.

The scatterplot tree for the illustrative ISPDT depicted in Fig. 9 is given in Fig. 10 (color online). Note that the scatterplots have a horizontal location on the x axis versus a pixel value on the y axis. Thus separability of the red (target box) pixels and black pixels is envisioned by projecting onto the y axis. For example, initial clustering of Hadamard frame 110 (the root node) produces the three nodes depicted in row two of the figure. These nodes are still depicting frame 110, as at the root. Notice that the target pixels in the middle cluster do not separate from nontarget pixels. However, Hadamard frame 151 (the next node in the tree) for these same pixels does provide significant separability.

5. Conclusions and Discussion

We have successfully applied ISPDT to DMA hyperspectral imagery for detection and localization, demonstrating the potential of an integrated sensing–processing suite consisting of a dynamically programmable DMA sensor and ISPDT processing.

This example demonstrates detection and localization. When multiple target types are possible, a subsequent classification step—possibly involving additional Hadamard frames—is employed. If, as is likely in practice, detection is not so perfect and some few off-target pixels are identified as falling into the target leaf (but still many on-target pixels are so identified), then postprocessing under some spatial dependence scheme, such as maximum *a posteriori* spatial filtering, can be used to perform the ultimate detection and localization.

Finally, while it may seem (and is indeed the case) that processing more elaborate than our individual pixel-based approach (such as the use of an edge detector) would make the detection–localization task trivial, the extremely time-sensitive nature of the pattern-recognition applications envisioned here preclude the use of elaborate processing schemes.

This research is supported in part by the Applied and Computational Mathematics Program of the Defense Advanced Research Projects Agency.

References

1. C. E. Priebe, D. J. Marchette, and D. M. Healy, “Integrated sensing and processing decision trees,” *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 699–708 (2004).
2. C. E. Priebe, D. J. Marchette, Y. Park, E. J. Wegmen, J. L. Solka, D. A. Socolinsky, D. Karakos, K. Church, R. Guglielmi, R. Coifman, D. Lin, D. M. Healy, M. Q. Jacobs, and A. Tsao, “Iterative denoising for cross-corpus discovery,” in *COMPSTAT: Proceedings in Computational Statistics*, J. Antoch, ed. (Physica-Verlag, 2004), pp. 381–392.
3. R. A. DeVerse, R. R. Coifman, A. C. Coppi, W. G. Fateley, F. Geshwind, R. M. Hammaker, S. Valenti, F. J. Warner, and G. L. Davis, “Application of spatial light modulators for new modalities in spectrometry and imaging,” in *Spectral Imaging: Instrumentation, Applications, and Analysis II*, R. M. Levenson, G. H. Bearman, and A. Mahadevan-Jansen, eds., *Proc. SPIE* **4959**, 12–22 (2003).
4. M. Harwit and N. J. A. Sloane, *Hadamard Transform Optics* (Academic, 1979).
5. R. A. DeVerse, R. M. Hammaker, and W. G. Fateley, “Realization of the Hadamard multiplex advantage using a programmable optical mask in a dispersive flat-field near-infrared spectrometer,” *Appl. Spectrosc.* **54**, 1751–1758 (2000).
6. A. Wuttig and R. Riesenberger, “Sensitive Hadamard transform imaging spectrometer with a simple MEMS,” in *Sensors, Systems and Next-Generation Satellites VI*, H. Fujisoda, J. B. Lurie, M. L. Aten and K. Weber, eds., *Proc. SPIE* **4881**, 167–178 (2003).
7. L. Hubbert and P. Arabie, “Comparing partitions,” *J. Classification* **2**, 193–218 (1985).
8. W. M. Rand, “Objective criteria for the evaluation of clustering methods,” *J. Am. Stat. Assoc.* **66**, 846–850 (1971).
9. C. Fraley and A. E. Raftery, “Model-based clustering, discriminant analysis, and density estimation,” *J. Am. Stat. Assoc.* **97**, 611–631 (2002).