Use of Hyperspectral Data with Intensity Images for Automatic Building Modeling*

A. Huertas[1], R. Nevatia[1] and D. Landgrebe[2]

[1]Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, California 90089-0273
huertas|nevatia@iris.usc.edu

[2]Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana 47907-1285
landgreb@ecn.purdue.edu

Abstract

Geospatial databases are needed for many tasks in civilian and military applications. Automated building detection and description systems attempt to construct 3-D models using primarily PAN (panchromatic) images. These systems can make use of cues derived from other sensor modalities to make the task easier and more robust. The recent development of hyperspectral sensors such as HYDICE (HYperspectral Digital Imagery Collection Experiment) can provide reasonably accurate thematic maps. Such data, however, tends to be of lower resolution, have geometric distortions and camera models are needed to map points between the different sensors. We use the thematic map to provide cues for presence of buildings in the PAN images for accurate delineation. It is shown that such cues can not only greatly improve the efficiency of the automatic building detection system but also improve the quality of the results. Quantitative evaluations are given.

Key Words: Information Integration, Sensor Fusion, HYDICE, Hyperspectral data, 3-D Building Modeling, Thematic Map.

1 Introduction and Overview

Three-D models of man-made structures in urban and sub-urban environments are needed for a variety of tasks. The principal sensor products used for this task have been panchromatic (PAN) images acquired from an aircraft [Noronha & Nevatia, 1997, Collins et al., 1998, Grün, et al. 1997, Grün & Nevatia, 1998, Paparoditis et al., 1998]. PAN images have many advantages: they are relatively easy to acquire at high resolution (say of the order of 0.5 meters/pixel) and humans find it is easy to visualize them and to extract the needed information from them. However, their use for automatic extraction has proven to be quite difficult. One of the principal causes of this difficulty is the high density of features present in the images. PAN image pixels encode reflected light intensity that gives little information to the nature of the material reflecting it. While it is possible to apply analyses that help recover structure from image elements, the problem of segmenting aerial scenes accurately remains a challenge.

* This research was supported in part by the U.S. Army Research Office under grant No. DAAH04-96-1-0444.

In recent years, advances in the solid state electronics have made possible the construction of hyperspectral sensors with an orders of magnitude increases in the number of bands possible, while at the same time providing improved signal-to-noise ratios. One such sensor, called HYDICE collects data of 210 bands over the range 0.4-2.5 μm with a field of view 320 pixels wide at an IFOV (pixel size) of 1 to 4 m depending on the aircraft altitude and ground speed. Given the spectral detail in such data it becomes practical and effective to construct a *thematic map* of an area that shows the layout of the various types of land cover and distribution of various materials in the scene.

In this paper, we focus on the task of building detection and reconstruction with the assistance of corrected and geo-referenced thematic maps derived from HYDICE data. The complementary qualities of conventional images and HYDICE image data provide an opportunity for exploiting them in different ways to make the task of automatic feature modeling easier.

Combining the two data sources at the pixel level is difficult as there is not a one-to-one correspondences between the pixels in the two sources, in general; hyperspectral data poses major challenges in terms of geometric corrections and terrain normalization. Instead, we propose to extract information from each which is then combined and perhaps used to guide extraction of additional information. In particular, we feel that the HYDICE data is suited for *detecting* possible building locations as buildings may be characterized by their roof materials. However, hyperspectral analysis results in a label for each pixel, but does not, by itself, combine pixels into objects such as buildings. HYDICE image data tends to be of lower resolution than conventional PAN images. Object boundaries are not likely to be precise and it may be difficult to distinguish a building from other nearby objects such as roads. PAN images, with much higher resolution can provide precise delineation as well as distinguish a building from other high objects much more reliably.

In the next sections, we describe how thematic maps are derived and how useful *cues* can be extracted from the HYDICE data. Use of these cues in the building extraction process is then described. Results comparing the effects of these cues are presented in section 4. Other approaches to use of HYDICE data may be found in [Ford et el., 1998, Bea & Healey, 1998, Healey, 1999, Madhok & Landgrebe, 1999].

2 Thematic Maps from HYDICE Data

The intent for multispectral and hyperspectral image data analysis is to rapidly and inexpensively associate a ground cover label to each pixel in the image. Given the multivariate nature of such data, the process of data analysis is one of dividing up the N-dimensional feature space into M exhaustive but non-overlapping regions where M is the number of classes of materials existing in the scene. The process involves defining the M classes of interest in a quantitative fashion, such that each pixel in the scene, which exists as a discrete location in the N-dimensional space, can be uniquely associated with one of the M classes. Frequently, this is done by using a small number of samples in the scene, called design samples or training samples, to define an N-dimensional probability density function for each of the M classes. Then an unknown pixel can be evaluated in terms of the likelihood of each possible class to determine the most likely class membership.

The onset of high dimensional hyperspectral data, on the one hand, greatly increases the potential of such a process. However, it has also introduced significant new challenges to the analysis process to achieve this potential, because such high dimensional feature spaces are much more complex. Not only can a 210-dimensional probability distribution not be visualized, but even the ordinary rules of geometry of 2- or 3-dimensional space do not apply in such high dimensional spaces [Lee & Landgrebe, 1993; Jimenez & Landgrebe, 1998]. Much progress has been made in recent years in understanding such high dimensional spaces and in devising effective analysis procedures for them [Landgrebe, 1999]. The following example from Fort Hood, Texas, will serve to illustrate some of the tools available for this process.

In this case, bands in the regions where the atmosphere is opaque were not considered and 171 bands in the 0.4 to 2.45 μm region of the visible and infrared spectrum were used. This data set contains 1208 scan lines with 307 pixels in each scan line. It totals approximately 130 Megabytes. The primary intent of the analysis of this data set was to identify rooftops and other impervious materials in the scene. With data this voluminous and complex, one might expect a rather complex analysis process, however, it has been possible to find quite simple and inexpensive means to do so. The steps used and the time needed on an inexpensive personal computer for this analysis are listed in the following table and are briefly described below.

Table 1: Thematic Classification Time

Operation	CPU time	Analyst time
Display Image	15 sec.	
Define Classes		30 min.
Feature Extraction	11 sec.	
Reformat	117 sec.	
Classification	86 sec.	
Total	229 sec.	30 min.

Define Classes

A software application program called MultiSpec, available to anyone at no cost from http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/, was used. The first step is to present to the analyst a view of the data set in image form so that training samples, examples of each class desired in the final thematic map, can be marked. A simulated color infrared photograph form is convenient for this purpose; to do so, bands 60, 27, and 17 are used in MultiSpec for the red, green, and blue colors, respectively. The image is shown in Figure 1. (Color versions of the figures in this paper are available at http://iris.usc.edu/home/iris/huertas/www/hydice.)

Feature Extraction

After designating the training areas, a feature extraction algorithm is applied to determine a feature subspace that is optimal for discriminating between the specific classes defined. The algorithm used is called Discriminate Analysis Feature Extraction (DAFE). The result is a linear combination of the original 171 bands to form 171 new bands that automatically occur in descending order of their value for producing an effective discrimination. From the MultiSpec output, it is seen that the first 15 of these new features should be adequate for successfully discriminating between the classes.

Reformatting

The new features defined above are used to create a 15 band data set consisting of the first 15 of the new features, thus reducing the dimensionality of the data set from 171 to 15.

Classification

Having defined the classes and the features, next a classification is carried out. The algorithm in MultiSpec used was the standard Gaussian maximum likelihood algorithm in which the mean vector and covariance matrix for each class are estimated from the training samples. These estimates then allow calculating the likelihood of each

Figure 1. Simulated infrared image using HYDICE bands 17, 27 and 60.

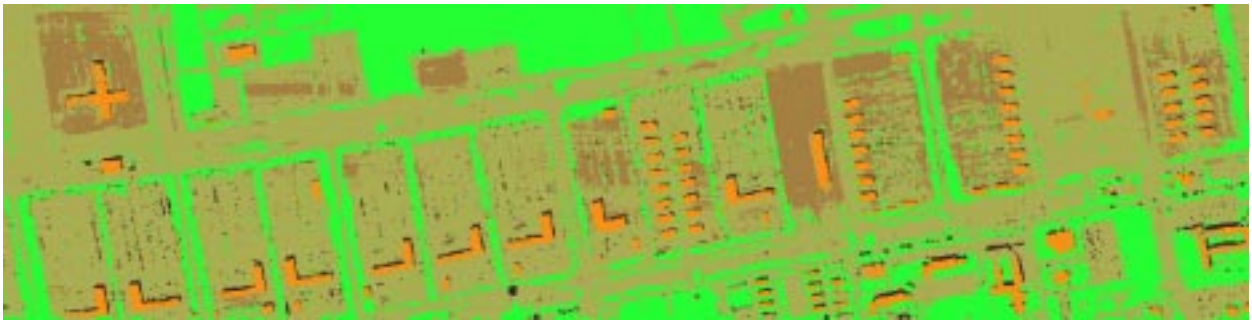

Figure 2. Thematic Map of the classes roof (black), road, lot, field (grays) and shadow (white)

class for a given pixel. The label of the most likely class is assigned to the pixel.

Hyperspectral data provides the capability to discriminate between nearly any set of classes. Research has shown that, of all the variables to the data analysis process, the most important one is the size and quality of the classifier training set. There are a number of additional steps that could be taken to further polish the result, but the current result appears to be satisfactory for the current use.

3 Integration of HYDICE and PAN Information

In order to integrate cues extracted from HYDICE data into the building detection and description system we require that the thematic map be rectified and registered to the PAN imagery as described next.

Geometric Rectification

Geometric rectification is needed to correct for the oscillations and "waviness" introduced by the nature of the HYDICE pushbroom sensor. Rectification is performed on the thematic map rather than on the hydice data directly. The method utilizes ground control points and control linear features typically found in urban scenes together with the pushbroom sensor model and a gauss-markov platform model to yield coordinate relationships between ground and image spaces. See [Lee, et al. 1999] for details. The accuracies achieved are in the 0.5 to 1 pixel range. Figure 3 shows a geometrically rectified thematic map of a portion of the Ft. Hood site. Note the straight roads. The waviness of the image boundaries gives an idea of the extent of rectification required.

Registration with PAN Images

The corrected thematic map has the geometric characteristics of an orthographic projection. The estimation of the sensor parameters, or "camera" model, associated with this overhead (nadir) viewpoint is straightforward. The camera model allows us to derive the appropriate 3D- to-2D and 2D-to-image transforms needed to register the available PAN images to the thematic map. We use these transforms to project EO 2-D and 3-D features onto the thematic map to assist and support the building detection system at various stages of processing. We describe in more detail, and illustrate these processes, with an example, below in section 4.

Cue Extraction

Figure 4 shows some of the barrack buildings in Fort Hood, Texas. The corresponding thematic map is shown in Figure 5. We first extract the roof pixels from the thematic map. These are shown in Figure 6. Many pixels in small regions are misclassified or correspond to objects made of similar materials as the roofs. The building cues extracted from this image are the connected components of certain minimum size. These components are shown in Figure 7; Except for one region, these components correspond to building roofs.

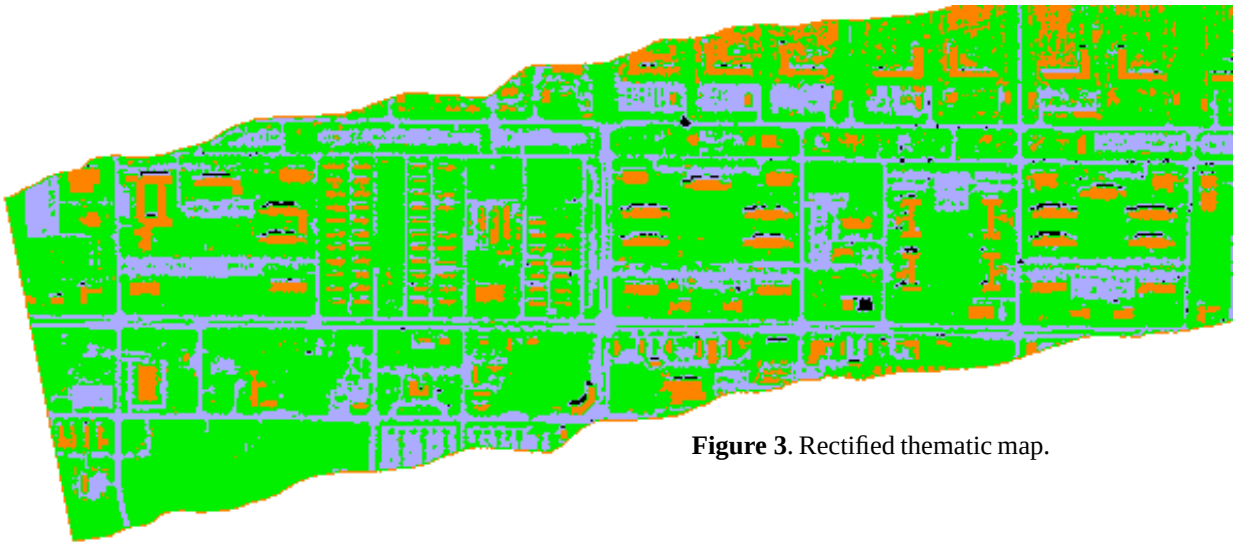

Figure 3. Rectified thematic map.

Figure 4. Barrack buildings at Fort Hood

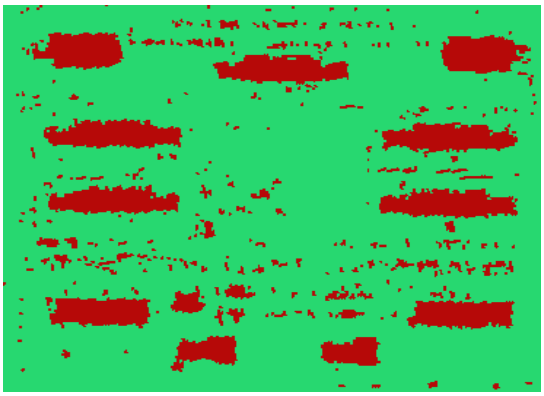

Figure 6. Roof Class from Thematic Map

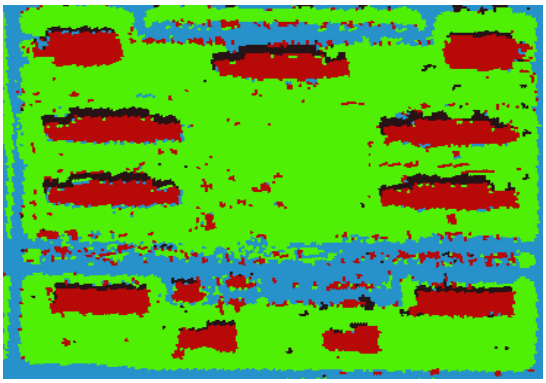

Figure 5. Thematic map.

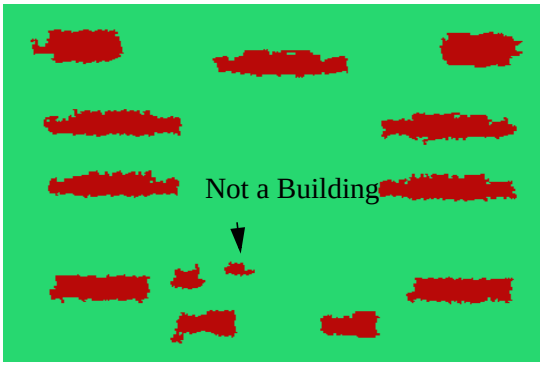

Figure 7. Building Cues

4 Multi-View System

We next describe the use the HYDICE cues in the multi-view building detection system described in [Noronha & Nevatia, 1997]. This system has three major phases: hypothesis formation, selection and validation. This system assumes that the roofs of buildings are rectilinear though the roofs need not be horizontal (some forms of gables are allowed). Hypotheses are formed by collecting a group of lines that form a parallelogram in an image. Multiple images and matches between lines are used in the hypotheses formation stage. As line evidence can be quite fragmented, liberal parameters are used to form hypotheses. Properties of resulting hypotheses are used to select among the competing hypotheses. The selected

hypotheses are then subjected to a verification process where further 3-D evidence, such as presence of walls and predicted shadows are examined.

The cues extracted from the HYDICE data can help improve the performance of the building description system at each of the three stages described above. We show some details and results of these processes.

Hypothesis Formation

Cues can be used to significantly reduce the number of hypotheses that are formed by only considering line segments that are within or *near* the cue regions. The 3-D location of a line segment in the 2-D PAN images is not known. To determine whether a line segment is near a HYDICE cue region we project the line onto the cue image at a range of heights, and determine if the projected line intersects a cue region. Figure 8 shows the line segments detected in the image of Figure 4 (using a Canny edge detector); Figure 9 shows the lines that lie near the HYDICE cues. As can be seen, the number of lines is reduced drastically (84%) by filtering without loosing any of the lines needed for forming building hypotheses.

This not only results in a significant reduction in computational complexity but many false hypotheses are eliminated allowing us to be more liberal in the hypotheses formation and thus including hypotheses that may have been missed otherwise.

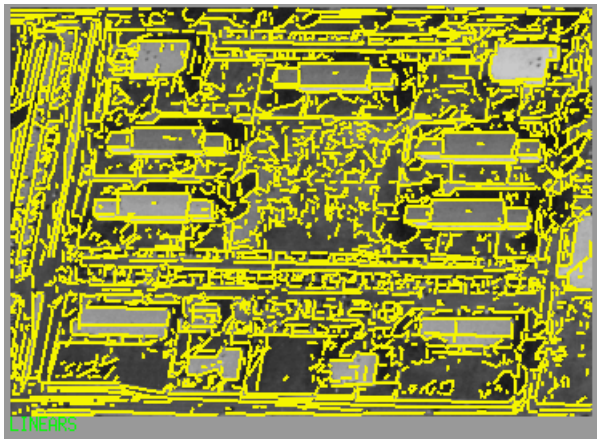

Figure 8. Line segments from PAN image.

Hypothesis Selection

The building detection system applies a series of filters to the hypotheses formed. The remaining hypotheses are then evaluated in the basis of the geometric evidence (underlying line segments that support the hypothesized roof boundaries), in an attempt to select a set of "strong" hypotheses. With HYDICE cues available we skip the initial filtering stages and introduce cue evidence into the roof support analysis. The evidence consists of support of a roof hypotheses in terms of the overlap between the

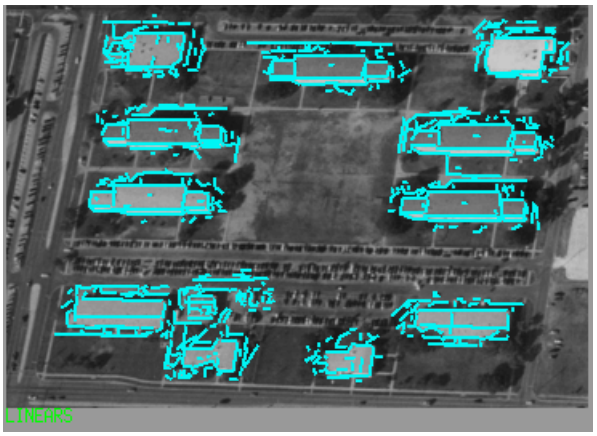

Figure 9. Lines near HYDICE cues.

roof hypotheses and the HYDICE cue regions. The hypotheses are constructed from matching features in multiple (two in this example) images and are represented by 3-D rectilinear components in 3-D world coordinates. We can therefore project them directly onto the HYDICE cues image to compute roof overlap (See Figure 10). The system requires that the overlap be at least 50% of the projected roof area.

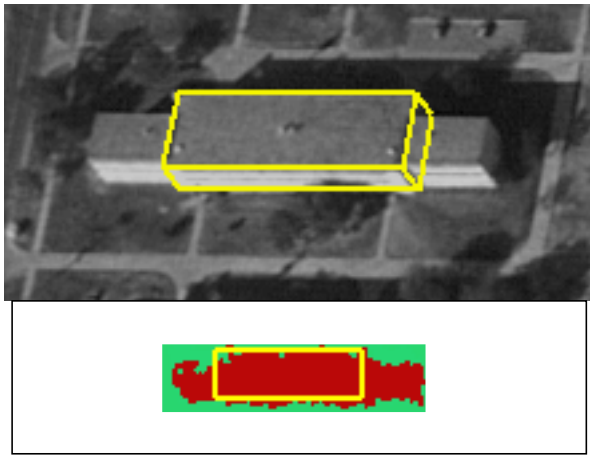

Figure 10. A 3-D hypotheses projected on PAN image (top) and on cue image.

Hypotheses Validation

Just as poor hypotheses can be discarded because they lack HYDICE support, the ones that have a large support see their confidence increase during the verification stage. In this stage, the selected hypotheses are analyzed to verify the presence of shadow evidence and wall evidence. Details of the shadow and wall analysis are given in [Lin et al., 1994]. When no evidence of walls or shadows is found, we require that the HYDICE evidence (overlap) be higher, currently 70%, in order to validate a hypotheses. The 3-D Models constructed with HYDICE support from the validated hypotheses are shown in Figure 11. For comparison, the model shown in

Figure 12 was derived without HYDICE support. Note that false detections are eliminated with HYDICE cueing. Also, the object cue on the lower middle in (Figure 7) is not found to be a building, even with HYDICE support, as the lack of geometric evidence prevented a hypothesis to be formed there. Also, the building components on the top left and on the lower left are not found without HYDICE support but found with it.

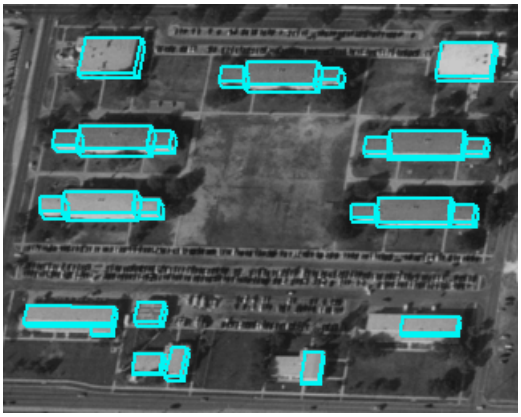

Figure 11. 3-D HYDICE assisted model

Figure 12. Model from PAN without HYDICE cueing

Once a 3-D model of the buildings is obtained, it is possible to reclassify the roof pixels in the thematic map more accurately by improving the delineation of the roof pixels boundaries and marking missclassified pixels (see Figure 13.)

An evaluation of the quality of results is given next.

5 System Evaluation

Table 2 gives a comparison of the number of features and final result component counts with and without use of HYDICE cues for the Fort Hood example. The two figures given for the line segments and linear structures correspond to the two images that were used, one of which was shown earlier in Figure 4.

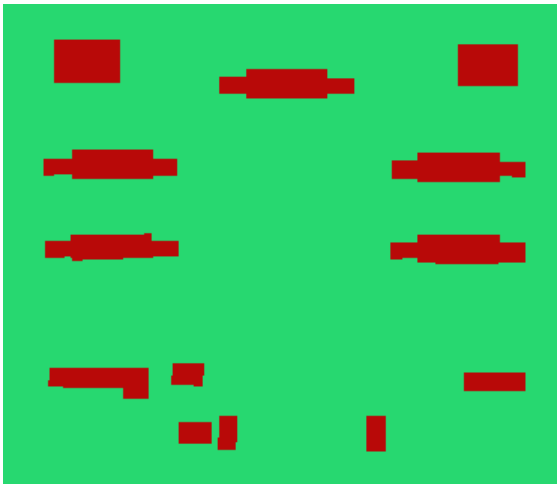

Figure 13. Refined roof class

Table 2: Execution Statistics

Feature	PAN Only	With HYDICE
Line Segments	15938/6976	
Linear Structures	6363/2693	796/652
Hypotheses	3793	636
Selected hypotheses	273	172
Verified hypotheses	115	127
Final hypotheses	20 (2 false)	24 (0 false)

To characterize the increase in performance of the system when HYDICE cues are available we use two basic metrics (see [Nevatia, 1999] for details), detection rate and false alarm rate, as follows:

$$\textit{Detection Rate} = \frac{TP}{(TP + FN)}$$

$$\textit{False Alarm Rate} = \frac{FP}{(TP + FP)}$$

TP, **FP** and **FN** stand for true positives, false positives and false negatives. Note that with these definitions, the detection rate is computed as a fraction of the reference features whereas the false alarm rate is computed as a fraction of the detected features.

In the definitions given above, a feature could be an object, an area element or a volume element. The first level of evaluation is to measure the detection and false alarm rates at the object levels such as for buildings or wings of a complex building. We consider each rectangular part of a rectilinear building as a separate object. A building object will be considered to be detected, if *any* part of it has been detected. Consider the reference model shown in Figure 14. Table 3 shows a summary of detection and false alarm results for the Ft. Hood example in terms of object parts.

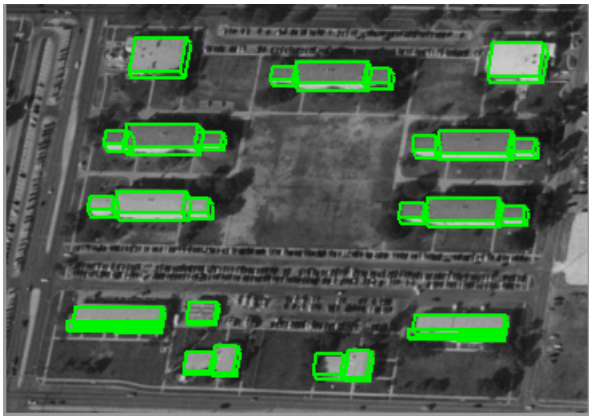

Figure 14. Reference model for evaluation.

Table 3: Component Evaluation

	PAN only	With HYDICE
Reference Model	26	
TP	20	25
FP	2	0
FN	6	1
Detection Rate	0.769	0.961
False Alarm Rate	0.09	0.00

To better reflect the quality of the detected components we also compute the accuracy the overlap between the footprints of the detected and the reference models and in the overlap between the 3-D volume occupied by them.

The area (volume) elements of the reference model that overlap with some area (volume) element of an extracted model can be considered to give the true positive (**TP**) values for the area (volume) elements of the reference model (the remaining elements of the reference models are the false negatives, **FN**). The area (volume) elements of the extracted model that do not overlap with any area (volume) element of the reference model give us the false positives (**FP**) for the area (volume) elements of the extracted model.

One way to combine the results of the above area (or volume) overlap analysis is to consider each area element as an object and count the detection and false alarm rates for all the area elements in the models. Table 4 shows these results for our Ft. Hood example. Ground detection rate is computed for the ground area elements (all elements that are not part of other objects); ground false alarm rate is not shown.

To better characterize the accuracy, we compute the detection rates for the area elements of each reference building component and the false alarm rates for each extracted building component separately. To visualize the result we compute a cumulative distribution of the detection and false alarm rates. Specifically, we can compute the percentage of building components of the reference model whose area (volume) elements detection rate (**TP**) is at a give value or *higher*. A curve plotting such a distribution is called a **CDR** curve [Nevatia, 1999]; Figure 15a shows the CDR curve for area elements of our Ft. Hood example. Similarly, we can compute the percentage of the building components of the extracted model whose false alarm rate (**FP**) is at a given value or *lower*. A curve plotting such a distribution is called a **CFR** curve; Figure 15b shows the CFR curve for the *area* elements of our Ft. Hood example. We also compute CDR and CFR curves for the *volume* elements for the reference and extracted building components. These are not shown for lack of space. A CDR curve that is consistently higher than another CDR curve indicates consistently better performance (similarly, a CFR curve that is consistently lower is consistently better).

Table 4: Ft. Hood Combined Area Evaluation

	PAN Only	with HYDICE
Detection rate	0.7116	0.8453
False Alarm rate	0.1510	0.0768
Ground Detection rate	0.9819	0.9907

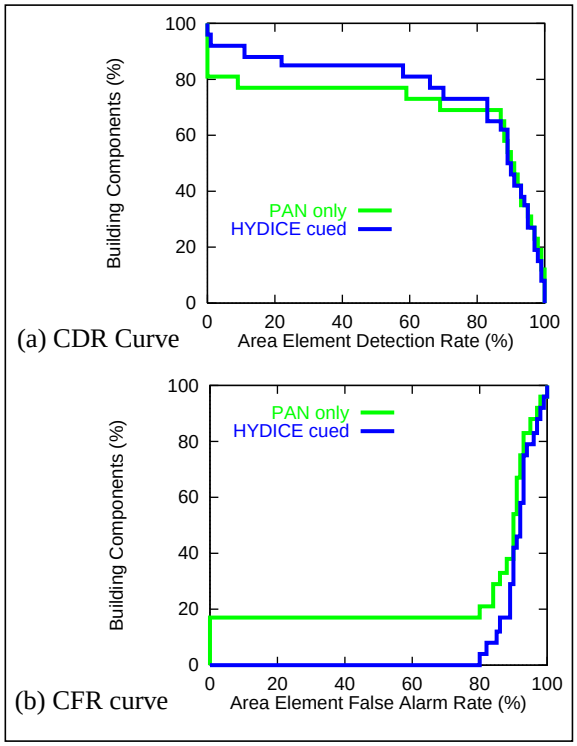

Figure 15. Evaluation curves for *area* analysis.

6 Conclusions

Many challenges remain in terms of data normalization and sub-pixel image registration for successful of data fusion of PAN and HYDICE types of imagery at the sensor level. Hyperspectral data however, provides the capability to discriminate between nearly any set of classes. By introducing an optimal feature design calculation on the 171 bands, we have shown that a good classification of materials can be achieved for production of a thematic map providing effective cues for objects of interest

We have presented a methodology for detection and reconstruction of building structures by using conventional intensity images with cues data derived from HYDICE sensors. Even though the HYDICE data is of a lower resolution and contains some missing elements and artifacts, it has been shown that it can be used to enhance the results of PAN image analysis while substantially reducing the computational complexity. This was accomplished not by combining the information at the sensor level but rather by using analysis of one to guide the analysis of the other. We believe that this paradigm will be suitable for other tasks as well as sensors of different modalities become available for more domains.

References

[Bea & Healey, 1998] T. Bea and G. Healey. "Invariant Subpixel Material Identification in Hyperspectral Imagery," Proceedings DARPA Image Understanding Workshop, Monterey, CA, November, 1998, pp. 809-814.

[Collins et al., 1998] R. Collins, C. Jaynes, Y. Cheng, X. Wang, F. Stolle, A. Hanson, and E. Riseman, "The ASCENDER System: Automatic Site Modeling from Multiple Aerial Images", Computer Vision and Image Understanding Journal, Vol 72, No. 2, November, pp 143-162.

[Ford et al., 1998] S. Ford, C. McGlone, S. Cochran, J. Shuffelt, W. Harvey and D. McKeown, "Analysis of HYDICE Data for Information Fusion in Cartographic Feature Extraction" Proceedings of the International Geoscience and Remote Sensing Symposium, IGARSS'98, Seattle, Washington, July 6-10 1998, Vol. V, pages 2702-2706.

[Grün & Nevatia, 1998] Computer Vision and Image Understanding Journal, Special Issue on Automatic Building Extraction from Aerial Images. A. Grun and R. Nevatia, editors. Academic Press, Vol. 72, No. 2, November.

[Grün et al., 1997] Proceedings of the Ascona Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images II, A. Grün, E. Baltsavias & O. Henricsson, Ed., Brinkhauser Verlag, Switzerland, May.

[Healey, 1999] G. Healey, "Reflectance Estimation and Material Classification for 3D Objects in Aerial Hyperspectral Images," To appear in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, June 1999.

[Jimenez & Landgrebe, 1998] L. Jimenez and D. Landgrebe, "Supervised Classification in High Dimensional Space: Geometrical, Statistical, and Asymptotical Properties of Multivariate Data," IEEE Transactions on System, Man, and Cybernetics, Volume 28, Part C, No. 1, pp. 39-54, February 1998.

[Landgrebe, 1999] D. Landgrebe, Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data, Chapter 1 of Information Processing for Remote Sensing, edited by C. H. Chen, World Scientific Publishing Co., Inc., River Edge, NJ

[Lee et al., 1999] C. Lee, H. Theiss, J. Bethel, and E. Mikhail, "Rigorous Mathematical Modeling of Airborne Pushbroom Imaging Systems." to appear in the Photogrammetric Engineering and Remote Sensing Journal, 1999.

[Lee & Landgrebe, 1993] C. Lee and D. Landgrebe, "Analyzing High Dimensional Multispectral Data," IEEE Transactions on Geoscience and Remote Sensing, Volume 31, No. 4, pp 792-800, July 1993.

[Lin et al., 1994] C. Lin, A. Huertas and R. Nevatia, "Detection of Buildings Using Perceptual Groupings and Shadows," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, June 1994, pp. 62-69.

[Madhok & Landgrebe] V. MAdhok and D. Landgrebe, "Supplementing Hyperspectral Data with Digital Elevation," To appear in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium," Hamburg, Germany, June 1999.

[Nevatia, 1999] R. Nevatia. "On Evaluation of 3-D Geospatial Modeling Systems," in ISPRS Proceedings of the International Workshop on 3D Geospatial Data Production", Paris, France, April, 1999.

[Noronha & Nevatia, 1997] S. Noronha and R. Nevatia. "Detection and Description of Buildings from Multiple Aerial Images", Proceedings IEEE Conference on Computer Vision and Pattern Recognition, San Juan, PR, June 1997, pp. 588-594.

[Paparoditis et al., 1998] N. Paparoditis, M. Cord, M. Jordan and J.P. Cocquerez, "Building Detection and Reconstruction from Mid- and High-Resolution Aerial Imagery", Computer Vision and Image Understanding Journal, Vol 712, No. 2, November, pp 122-142.