

5-19-2006

# Semi-Automated DIRSIG Scene Modeling from 3D LIDAR and Passive Imaging Sources

Steve R. Lach

*U.S. Air Force*

Scott D. Brown

*Rochester Institute of Technology*

John P. Kerekes

*Rochester Institute of Technology*

Follow this and additional works at: <http://scholarworks.rit.edu/other>

---

## Recommended Citation

S. R. Lach, S. D. Brown, J. P. Kerekes, "Semi-automated DIRSIG scene modeling from 3D LIDAR and passive imaging sources", Proc. SPIE 6214, Laser Radar Technology and Applications XI, 62140I (19 May 2006); doi: 10.1117/12.666096; <https://doi.org/10.1117/12.666096>

This Conference Proceeding is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Presentations and other scholarship by an authorized administrator of RIT Scholar Works. For more information, please contact [ritscholarworks@rit.edu](mailto:ritscholarworks@rit.edu).

# Semi-Automated DIRSIG Scene Modeling from 3D LIDAR and Passive Imaging Sources

S.R. Lach<sup>a</sup>, S. D. Brown<sup>b</sup> and J.P. Kerekes<sup>b</sup>

<sup>a</sup>U.S. Air Force & Rochester Institute of Technology

<sup>b</sup>Digital Imaging and Remote Sensing Laboratory

54 Lomb Memorial Drive, Rochester, NY 14623

Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology

## ABSTRACT

The Digital Imaging and Remote Sensing Image Generation (DIRSIG) model is an established, first-principles based scene simulation tool that produces synthetic multispectral and hyperspectral images from the visible to long wave infrared (0.4 to 20 microns). Over the last few years, significant enhancements such as spectral polarimetric and active Light Detection and Ranging (LIDAR) models have also been incorporated into the software, providing an extremely powerful tool for algorithm testing and sensor evaluation. However, the extensive time required to create large-scale scenes has limited DIRSIG's ability to generate scenes "on demand." To date, scene generation has been a laborious, time-intensive process, as the terrain model, CAD objects and background maps have to be created and attributed manually.

To shorten the time required for this process, we are initiating a research effort that aims to reduce the man-in-the-loop requirements for several aspects of synthetic hyperspectral scene construction. Through a fusion of 3D LIDAR data with passive imagery, we are working to semi-automate several of the required tasks in the DIRSIG scene creation process. Additionally, many of the remaining tasks will also realize a shortened implementation time through this application of multi-modal imagery. This paper reports on the progress made thus far in achieving these objectives.

Keywords: DIRSIG, synthetic imagery, scene generation, hyperspectral, lidar, fusion, building reconstruction

## 1. INTRODUCTION

In recent years, the remote sensing community has seen significant advances in the development of non-traditional imaging techniques. With the advent of commercial imaging spectrometers, hyperspectral image analysis has been near the forefront of this boom, and algorithms exploiting the spectral content of a scene continue to evolve. However, in many cases, testing, validation and training of these algorithms requires the use of well-calibrated datasets, the generation of which can be expensive and time prohibitive<sup>1</sup>. As such, quality hyperspectral synthetic imagery can be a tremendous asset, as often the simulated data is able to reduce or eliminate the need for costly real-world data collection campaigns. Additionally, realistic synthetic image generation (SIG) enables engineers to evaluate the final image products given the parameters of the sensor. This may be done to evaluate an existing sensor under a host of illumination, atmospheric and geometric scene conditions, but it may also be done before the sensor is even built. As such, SIG is often a critical component in the design process of new sensors.

Light Detection and Ranging (LIDAR) systems have undergone a similar maturation in the last decade, and the use of laser-scanning systems to produce three-dimensional imagery is now widespread. However, even with the proliferation of this technology, many researchers are still lacking quality data sets with corresponding truth measurements. This is especially true for those working in the area of multi-modal image fusion, where the LIDAR data is augmented by imagery from an additional sensor. In this field, accurate registration of the imagery is critical, yet without truth data the quality of this registration is often difficult to ascertain. However, with a good multi-modal SIG model, synthetic

combined datasets with associated truth values are easy to obtain. Additionally, synthetic LIDAR images may be used for the characterization and design of sensors, much as they were in the hyperspectral case. For these reasons, the ability to produce accurate simulated imagery can be of great benefit to the LIDAR community as well.

The Digital Imaging and Remote Sensing Image Generation (DIRSIG) model described in the following section is a tool capable of creating this desired synthetic imagery. Given a pre-defined scene, it generates an accurate representation of what a passive electro-optical (EO) or active LIDAR sensor would detect under specified conditions by modeling the relevant physical processes in the imaging chain. Its use has been well documented in the literature (see Ientilucci<sup>1</sup> and Schott et al.<sup>2</sup> for example), and it is currently being used in both academia and industry. However, while it is fairly easy to model different sensor designs, atmospheric conditions and geometries for a given scene, the current process for actually generating a DIRSIG scene is quite involved. As a case in point, the creation of the MegaScene documented by Ientilucci<sup>1</sup> took a team of researchers well over a year to complete.

To shorten the time required for this process, we have begun work that aims at reducing the man-in-the-loop requirements for several aspects of DIRSIG scene construction. Through a fusion of 3D LIDAR data with passive EO imagery, we are working to partially-automate several of the required tasks in the scene creation process. Additionally, many of the remaining tasks will also realize a shortened implementation time through this application of multi-modal imagery. This paper reports on the progress made thus far in achieving these objectives.

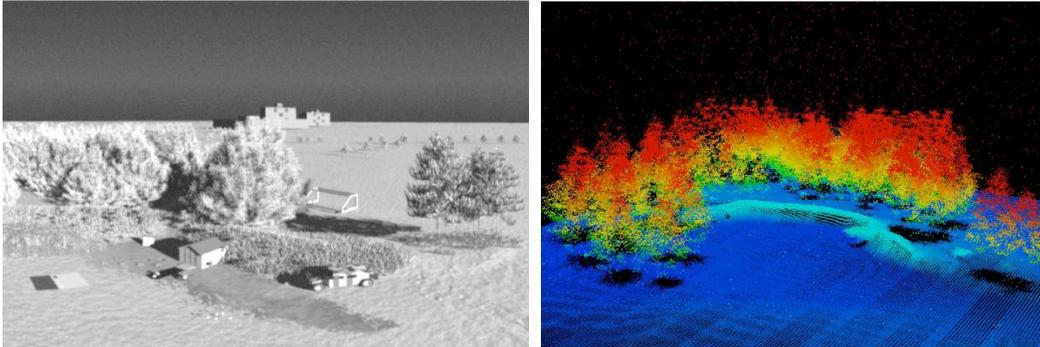
## 2. THE DIRSIG MODEL

The DIRSIG model is a complex synthetic image generation utility that was developed at the Digital Imaging and Remote Sensing (DIRS) Laboratory at the Rochester Institute of Technology (RIT) over the last 20 years. The tool was originally designed to model the thermal infrared region of the electromagnetic spectrum, but was expanded several years ago to cover the full visible to long-wave infrared (0.4 to 20 micron) range. It effectively models broadband, multispectral and hyperspectral imagery using a suite of first-principles based radiation propagation modules. These modules perform specific tasks such as predicting bi-directional reflectance functions (BRDF), computing time and material dependent surface temperature values, and computing the dynamic viewing geometries of scanning instruments on a moving platform. In addition to these DIRSIG-specific modules, the tool leverages several utilities (such as MODTRAN and FASCODE) that have been used extensively by the remote sensing community.

In 2002, a first-principles-based LIDAR model was incorporated into the passive radiometry framework enabling the model to calculate arbitrary, time-gated photon counts at the sensor for atmospheric, topographic, and volumetric backscattered returns. The DIRSIG LIDAR model was first developed by Burton<sup>3</sup>, and it was recently expanded by Blevins<sup>4</sup> to handle a wide variety of complicated scene geometries, diverse surface and participating media optical characteristics, and a variety of sensor models. The LIDAR module now includes the effect of multiple scattering, multiple bounces from topographical targets, and absorption within non-homogenous finite volumes. Additionally, several noise sources are extensively modeled, such as speckle from rough surfaces and atmospheric turbulence phase effects. As such, DIRSIG is now able to help researchers evaluate design trades for topographic systems and the impact that scattering constituents (e.g. water vapor, dust, sediment, soot, etc.) may have on a Differential Absorption LIDAR (DIAL) system's ability to detect and quantify constituents of interest within volumes including water and atmospheric plumes<sup>4</sup>.

Although the passive and active models used by DIRSIG are valuable in their own right, in many applications the ability to integrate passive EO and LIDAR simulations using a common scene (and potentially common viewing geometries) is of even greater utility. The multi-modal capabilities of the DIRSIG model allow designers to evaluate both passive and active approaches to solving specific imaging problems, as well as potential fusion techniques using both image modalities. Additionally, the upper performance limit for a given approach may be more easily determined when using a common baseline scene.

Two representative synthetic images produced by the DIRSIG model are included in Figure 1 (Courtesy of Blevins<sup>4</sup>). The first depicts a near-infrared (passive EO) image of MicroScene1 (a small test range on the RIT campus), and the second is a topographic LIDAR product of a portion of the same area.



**Figure 1:** A near-infrared simulation of MicroScene1 produced by DIRSIG (left) and Simulated Topographic LIDAR Image (right)<sup>4</sup>

### 3. CURRENT PROCESS FOR SCENE CREATION IN DIRSIG

The current method of creating a scene for use in DIRSIG simulations is straightforward, but unfortunately it is also quite labor and time intensive. The detailed steps described below have been taken from Brown<sup>5, 6</sup> and Ientilucci<sup>7</sup>, and represent the techniques used at the Rochester Institute of Technology. Although the specific methods employed in creating scenes may be different at other organizations, the fundamental principles and basic steps remain the same.

The first step in the generation of a new scene has typically been to acquire a digital terrain model (DTM) of the area to be modeled. For a real scene, we have used USGS Digital Elevation Models (with a 10 meter ground sample distance) with success, but other sources of this data are also available. For fictional scenes, the terrain model may be designed as desired. Once the terrain geometry has been determined, it is usually up-sampled and registered with any relevant aerial imagery, which facilitates several of the other steps required in the scene specification. Additionally, the terrain model needs to be facetized, which is often done using a triangular irregular network (TIN) generation program for terrains with a near-constant slope. It should be noted that in many cases, the terrain must be smoothed via some form of spatial averaging before the facetized version may be created.

Once the geometric properties of the terrain have been determined, the 3D objects that will appear in the scene must be created. These objects include buildings, trees, vehicles, and any other structures that are not a part of the terrain model. This is often a very time-intensive process, especially if many unique object types are desired. The geometric description of these objects can be generated using a variety of computer aided design (CAD) tools, as long as the specific utility is able to output the final facetized structure as a Alias/Wavefront OBJ file. At RIT, most of the geometric models have been manually drawn with the Rhinoceros<sup>8</sup> CAD package, which has proven to be well suited to the task. The one exception is the modeling of tree structures, for which we have chosen to use Tree Professional<sup>9</sup>. This software allows the user to render a host of trees, palms, and other plants of various species and sizes and output the result as a facetized model.

Next, the spectral reflectance and emissivity data must be acquired. This is typically accomplished either through a direct collection (using an instrument such as a field spectrometer) or by accessing a spectral data library. This data is then included in a materials file that will be used when accessing the scene.

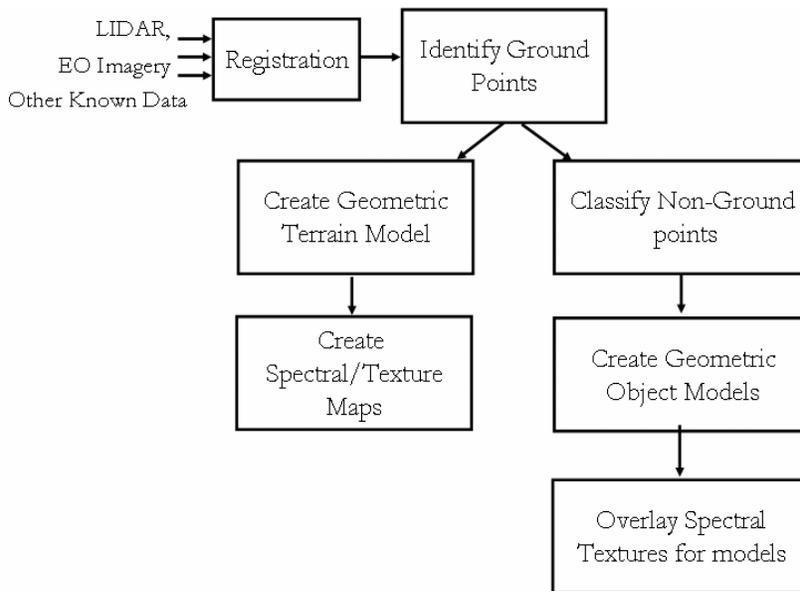
The geometric objects are then attributed material characteristics using DIRSIG's Bulldozer tool. This utility allows the user to import the Alias/Wavefront OBJ files created earlier and assign material properties to each facet. This task is also quite time consuming, as each portion of an object must be attributed manually.

Road, material, texture, and other background maps (scene layers) are created next<sup>10</sup>. Road maps have traditionally been created in Adobe Photoshop, due to artifacts left behind when segmenting real imagery. These maps are created by opening the scene file, and adding a new layer to the image. The roads are then drawn on the new layer with the paintbrush tool, using the scene image as a guide. The scene layer is then removed, and a black background is added to the road map layer. A texture map may also be created that drives in-material spatial-spectral variations. These maps usually take the form of traditional classification maps, and are often generated through traditional segmentation algorithms on real aerial imagery.

Finally, the Bulldozer utility is used to place the attributed objects onto the facetized terrain model. If aerial imagery is available, a "tracing paper" layer may be employed help with object orientation. Once the objects are placed in the scene as desired, a DIRSIG configuration file is created to link together all of the previous files.

#### 4. PROPOSED PROCESS

Although the current process for creating DIRSIG scenes has proven to yield excellent resultant images, a streamlining of the construction of large scenes would be beneficial to many users. To this end, a new process is proposed where several of the scene design tasks have been replaced by semi-automated methods. The basic concepts for automating much of the DIRSIG scene generation process are depicted below in Figure 2, and will be explained in greater detail throughout the remainder of this section.



**Figure 2:** Proposed Process for Scene Construction

#### 4.1 Registration of the Passive Imagery with LIDAR Data

A critical step when working with multi-modal imagery is the proper registration of all data sources, as a common frame of reference must be achieved before data may be combined effectively. Although this typically has been done by determining the deformation parameters associated with mapping homologous points from one image to another, recent literature indicates that a better approach is to map extracted features such as edges and surface patches.<sup>11</sup> Using explicitly defined features in the registration process has the additional benefit of contributing to the geometric reconstruction of objects in the scene.

Also, due to the differing phenomenologies associated with the various sensor types, the standard correlation metric for evaluating the quality of registration is often inappropriate. Instead, a better choice may be to use a maximization of mutual information (MMI)-based approach<sup>12</sup>. MMI techniques use the mutual information found between two images to extract the transformation parameters and evaluate the quality of the registration. It may be shown that such approaches are significantly more robust with respect to phenomenological changes than basic correlation methods.

Additionally, it should be noted that when dealing with hyperspectral images, one should usually avoid working with a transformed spectral image cube. Although each image modality needs a transformation describing how its geometry relates to all of the others, processing of the spectral data should be performed on the non-interpolated image cube.

#### 4.2 Extraction of the DTM from LIDAR Imagery

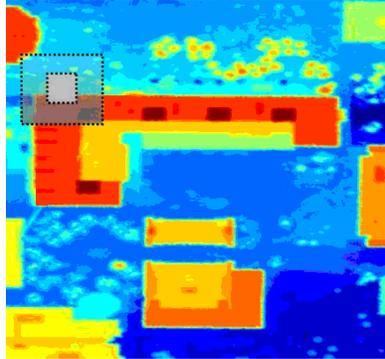
In creating the terrain model, the irregularly-sampled point cloud may be processed directly, or the data may first be interpolated onto a regular grid to form a range image. In either case, the fundamental concept is to separate the ground points from the non-ground points, then to remove the non-ground points (and potentially interpolate across them). Ma<sup>13</sup> describes several techniques that may be applied directly to the 3D point data, but we found that satisfactory results could be obtained easily by applying image processing techniques to the range image.

One potential technique for determining non-ground pixels is based on fitting polynomial splines to each row and column of the range image. The height of each pixel above these curves is then determined, and through a simple thresholding points may be classified as being ground-level or not. Although this technique was successful in several cases, we found it difficult to determine parameters that were robust with respect to different sampling densities and different sized scenes. Additionally, in many cases this method failed to produce accurate results at the scene edges. As such, this technique was replaced by one with better performance.

An alternate approach is to use a variation of the spatial sliding window median filter in order to identify points that are significantly higher than their neighbors. Although in many cases a standard median filter could be used, care must be taken to ensure that the kernel is large enough to span the roof structures of the largest buildings in the scene. If it is not, the central points of large objects may not be flagged as being non-ground, and more complicated processing would be required. However, when the kernel is large, this technique may fail in regions where there is a low ratio of ground to non-ground points. We avoid these issues by computing the median value for a small region (typically 10m x 10m), and then flagging points in a larger region that are significantly higher than this median value. This modified median filter also has the advantage of being much more computationally efficient, and a similar technique may be employed directly on the point cloud, if desired. Figure 3 illustrates this concept for a single location of the filter's sliding windows.

Once the points that are significantly higher than their surroundings have been flagged as being non-ground locations, a second filtering operation must be performed in order to remove additional objects from the terrain. These include large tree canopies, isolated small trees, vehicles, and other man-made objects, which would all be ignored by median filter processing. In order to effectively remove these points, the LIDAR range image is high-pass filtered, thereby highlighting regions with rapid changes in elevation. Bright pixels in this second filtered image (those with maximum rate of change) are also flagged as being non-ground.

After the non-ground pixels have been identified, they are removed from the data set, and an initial terrain model is obtained by interpolating across the removed points. If desired, a smoothing filter may be applied to remove remaining high-frequency content in the resultant data, and the output is then facetized to produce the final DTM.



**Figure 3:** Depiction of the Modified Median Filter. The median value of data in the smaller square region is computed. Data points in the larger square region more than a given threshold above this median value are then flagged as “non-ground”

### 4.3 Segmenting Buildings from Trees

Through the use of the modified median filter, we extracted a set of points that were significantly higher than their neighbors (the initial set of “non-ground” points). In many cases, it may be assumed that these points represent buildings and trees, and the next logical task is to identify to which class each of these points belongs. With finely registered multi-modal imagery, many features are available for addressing this problem.

The spectral reflectance of plants has a distinct signature, and therefore multispectral and hyperspectral classification techniques excel at differentiating vegetation from most man-made objects. Through the use of simple spectral clustering techniques applied to the building/tree pixels, one may easily determine to which class each object belongs. This classification problem has been well documented in the literature, and a host of supervised and unsupervised algorithms may be used for this task. Recently, Rex et al demonstrated that a similar method of fusing LIDAR data with a multispectral imagery was quite successful in extracting tree and building pixels in a complex scene<sup>14</sup>. In general, use of spectral information is the most robust method that we considered for separating building and tree regions, although in many cases images with the required spectral information may not be available.

In the event that spectral information is not available, spatial-based approaches may also be used. We have successfully implemented a range-image based morphological technique in which each region was “opened” by a specific amount. By sizing the structuring element to preserve just a few pixels from the smallest building in the scene, tree regions were effectively be removed. This held true even for large areas containing many trees, as several ground pixels are present in almost all tree canopy regions. Objects that contained pixels in common with the opened image were specified as being buildings, and the remaining objects were classified as trees. It should ne noted that the user must ensure that the building regions are fully filled in when using this approach, or else those regions may be misclassified.

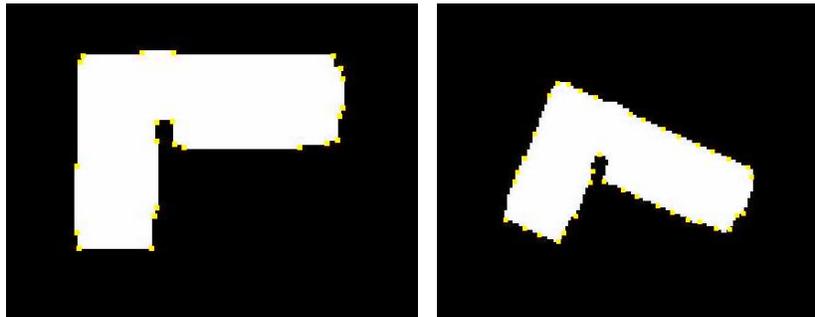
Additional approaches exploiting different features are also found in the literature. Ma<sup>13</sup> lists several references addressing this problem from a geometric standpoint, each applicable to scenes meeting a few simple assumptions. These include techniques for finding buildings based on homogeneity of height and LIDAR return intensity, using the average length of surface edges as a distinguishing parameter, and categorizing objects made up of planar patches as buildings. Taking a cue from the first of the methods, we were able to show that the average magnitude of the high-pass filter output (from the DTM extraction stage) may be used value to distinguish trees from buildings.

#### 4.4 Object Reconstruction (Trees and Buildings)

Once the building and tree points have been categorized, we need to determine the geometrical properties of each object and represent them as a set of facetized models. For the autonomous geometric reconstruction of trees, our method is an extension of previous work done at RIT. Gray et al.<sup>15</sup> developed a feature extraction tool that partially automated the construction of forested areas in a DIRSIG scene. In that method, high resolution aerial imagery was blurred in order to emphasize the lower-frequency features of the trees in the image. By then performing correlations with different sized circle functions, regions with high radial symmetry were identified. An additional comparison of the NIR-band response to that of the red channel helped to further isolate individual tree boundaries. We propose to augment this process by also using the height information provided by the LIDAR data. Tree centers may be found by extracting local maxima of a low-pass filtered version of a more finely sampled range image, and the combined use of passive and LIDAR imagery is able to yield a more accurate measure of the spread of the individual trees. Once tree sizes and locations are determined, a specific tree model is selected from a library of objects previously created with the Tree Professional software.

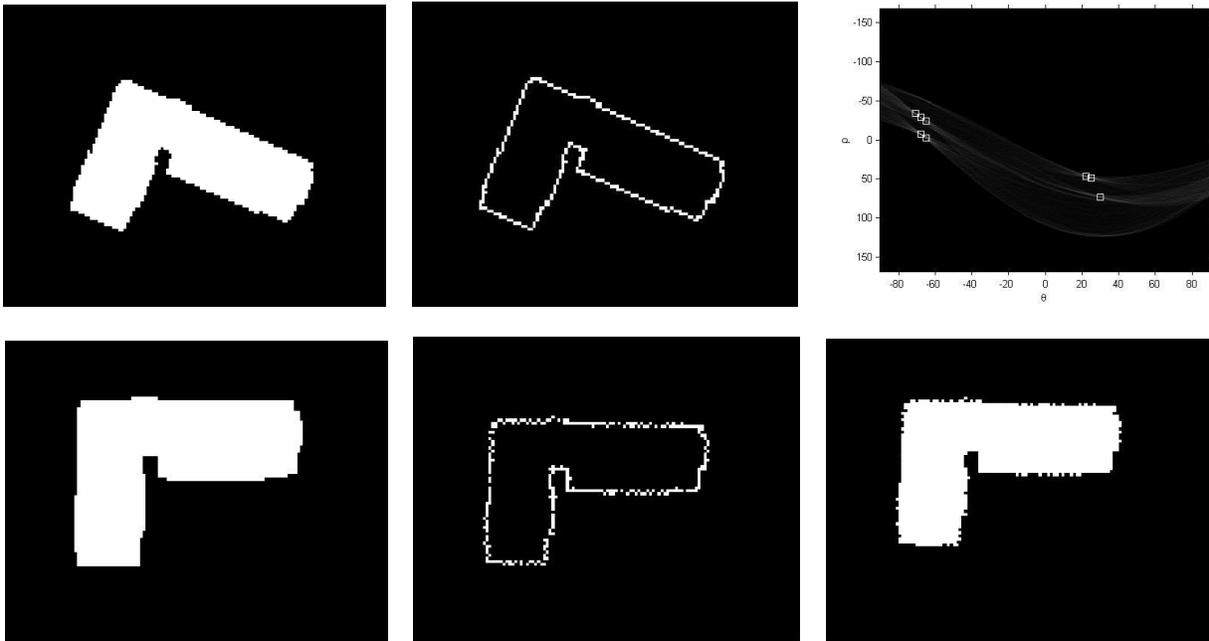
For buildings, a different process is proposed. For our initial models, we chose to extract features strictly from the LIDAR range image, although future work aims at fusing this data with features found by traditional photogrammetric techniques. Additionally, we made the assumptions that each building was of uniform height, all surfaces were either horizontal or vertical, and the side facets were likely to be oriented in integer multiples of  $45^\circ$ . By using these assumptions in combination with a Harris feature (corner) detector on a binary image representation of each building region, we were able to extract enough vertex detail to recreate the basic shape of most buildings.

For range images with lower sampling frequencies, we did run into difficulty when the buildings were aligned so that the dominant edge angles were not at  $0^\circ$  and  $90^\circ$  relative to the underlying image grid. In these cases, the diagonal lines describing the building outline are actually composed of noticeable horizontal and vertical line segments, and the Harris detector finds additional spurious corners. This may be seen in Figure 4, where the underlying image grid has pixels centers spaced every two meters.



**Figure 4:** Harris feature points (corners) for two building orientations

In order to alleviate this problem, we opt to rotate each building until its dominant edge angles are at  $0^\circ$  and  $90^\circ$  before applying the Harris detector. The angular rotation required is determined by performing a boundary extraction of the original building (using a Canny edge detector or morphological techniques), and then transforming this boundary image into parameter space via a Linear Hough Transform (LHT). The bright points in Hough space represent the length and angle from the origin or a line normal to the dominant lines in the original image. By extracting this angular information, one is able to rotate the building with enough accuracy for the corner detection process to be successful. This process is depicted in Figure 5.



**Figure 5:** Reorientation of a Building: Building with primary angle at  $-22.5^\circ$  (top left), border pixel extraction (top center), linear Hough parameter space representation (top right), and reoriented building (lower left). Also shown is the case of rotating the border pixels and filling the resultant region (lower center, right), which produces a final image with significantly noisier edges

#### 4.5 Creating Background Maps

As noted in Section 3, a variety of background characterization methods have been added to the DIRSIG model in the form of a collection of map layers. Each layer assigns various attributes to surfaces within a scene, and in general, we propose to extract these map layers from multispectral or hyperspectral data sharpened by higher resolution passive imagery. In most cases, this spectral data will be used to geometrically segment the image, and the desired properties will then be assigned via a look-up table that accesses previously created attributes. In nearly all cases, the mapping images should be available at spatial resolutions significantly higher than the resolution of the sensor to be simulated, so that DIRSIG can sufficiently oversample the scene to allow for modeling of the detector point spread function<sup>10</sup>.

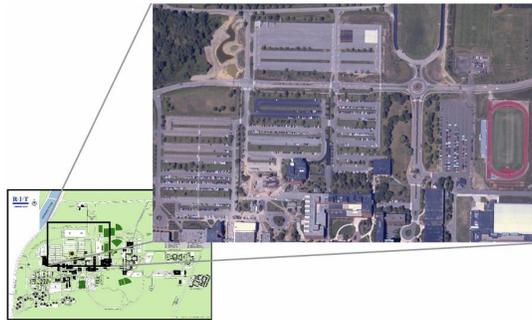
As a case in point, road maps in DIRSIG have traditionally been created in a drawing utility such as Adobe Photoshop. An improvement may be realized by segmenting the scene using traditional spectral classification techniques, then refining the road regions through use of high-resolution panchromatic imagery (see Wang<sup>16</sup>, for an example of performing this refinement via edge extraction). Although this often works well without additional processing, we have found that industrial roof-tops are frequently categorized in the same class as roads. Therefore, augmenting the spectrally-derived classification map with the LIDAR-derived building locations provides a means of improving the final roadmap. Similar techniques may be used to create material, texture, mixture, and reflectance maps.

Although this method of map generation will produce adequate results in many cases, in others a lack of data limits the degree of automation that can be incorporated into this process. For example, a significant portion of a road may be hidden under a tree canopy. Although the technique described above can extract road outlines in the case where small segments are blocked by isolated trees, a more complete hiding of the road will cause erroneous results. In these cases, the extracted layers will still need to be modified by the user before inclusion in the final scene. As such, we anticipate retaining a certain level of manual intervention in the creation of background maps for the foreseeable future.

## 5. INITIAL RESULTS

In this section, we briefly demonstrate the feasibility of the proposed approach by applying these techniques to a portion of the RIT campus. The imagery used in this analysis came from a commercial LIDAR collection augmented by a single image from the Wildfire Airborne Sensor Program (WASP). WASP is an imaging system designed and built at RIT for the primary purpose of wildfire detection and mapping, and it augments standard visible-band imagery with sensors operating in the SWIR, MWIR, and LWIR regions.

The area of interest is shown in Figure 6, and as can be seen it contains several buildings, a region with densely populated trees, isolated trees, roads, grass and dirt. There is also a 10 meter change in ground elevation throughout the region, and a significant amount of truth data for this locale has been collected. As such, this has proven to be an excellent test scene for the proposed techniques.



**Figure 6:** Region of Interest for Initial Efforts

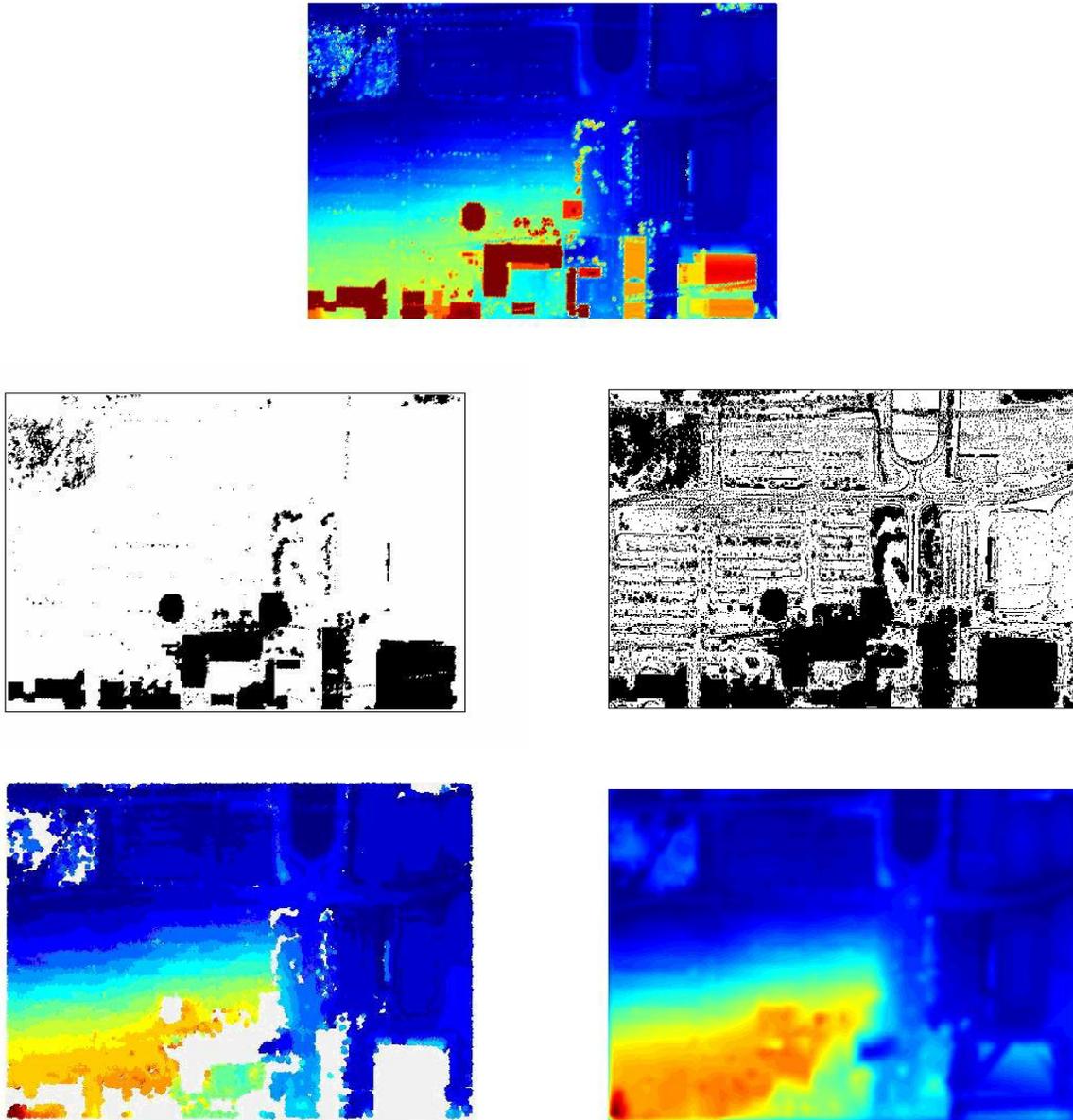
We began the process by performing a simple point matching transform to register the LIDAR data to the WASP multispectral image. Although we were able to find matching points in both datasets, we were unable to autonomously match them. Therefore, for the purposes of this paper, this simple registration was done manually. Figure 7 depicts the quality of the registration where the passive imagery has been converted to HSV space, with the LIDAR height values scaling the value channel. Although we believe better registration is certainly possible, we still achieved a matching to within a few meters across the image.



**Figure 7:** Region of Interest for Initial Efforts

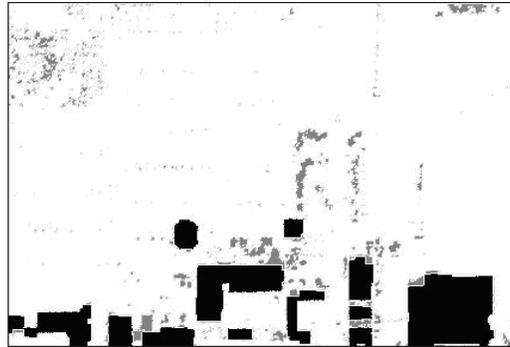
Once the registration was completed, we processed the LIDAR data as specified above to extract the DTM. Both the modified median filter and high-pass filter techniques were used to flag points for removal, and a simple linear interpolation was used to reconstruct the full terrain image. A 5x5 averaging filter was applied after interpolation to smooth out some of the high-frequency content before facetization. The various stages of this DTM extraction process are depicted below in Figure 8.

One problem area we noticed when analyzing the final terrain model was that holes adjacent to buildings were not handled well. There is a basement-accessible loading dock on the eastern side of the L-shaped building, and once the building points were removed, this hole caused local problems with the subsequent interpolation. Future research will focus on a method to address this issue



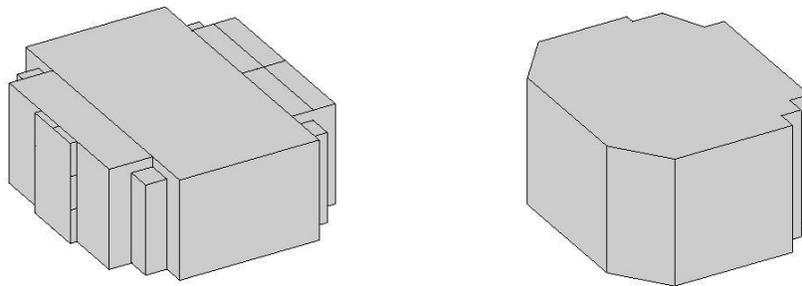
**Figure 8:** Original scene (top), points identified for removal using MMF only (upper left), points identified for removal using MMF and High-Pass Filter (upper right), scene after removal of points (bottom left), interpolated and smoothed DTM (bottom right)

In creating the DTM, a set of combined building/tree points was also produced (by the median filtering). We then used a morphological opening on the binary image of combined tree/building points, and were able to determine which regions corresponded to each object type. The result of this object segmentation is shown in Figure 9.



**Figure 9:** Building/tree map obtained using only morphological techniques

Finally, we attempted to recreate one of the buildings in the scene using corner points taken from the LIDAR-derived building map. We forced the roof to be a horizontal planar surface, and the building height was set to the maximum roof to ground distance plus eight meters. This additional height will be inconsequential when the building is placed in the final scene, as it will be submerged into the terrain by the proper amount. Figure 10 shows a manually drawn CAD model of this building alongside the extracted model. Although a few corners were missed or misrepresented in the extracted model, the overall exterior building measurements match truth data quite well.



**Figure 10:** Manually Drawn (left) and Autonomously Extracted (right) CAD Models of the Carlson Center for Imaging Science at RIT

## 6. SUMMARY

This work presented an approach to reducing man-in-the-loop requirements for several aspects of synthetic hyperspectral scene construction. Through a fusion of 3D LIDAR data and passive imagery, we were able to partially automate several of the required tasks in the DIRSIG scene generation process. These included extraction of a bare-earth digital terrain model, identification of buildings and trees, object reconstruction, and the generation of background maps. Through the proper application of these techniques, we hope to enable the creation of synthetic scenes where truth data is not available, as well as to significantly reduce the time required by current scene-building methods.

After describing the proposed process for scene construction, we demonstrated the feasibility of the techniques by applying them to a portion of the RIT campus. The DTM extraction, identification of buildings and trees, and low-level geometric reconstruction of buildings was shown to be successful.

The overall scope of this research is quite large, and the work presented here represents the initial steps of a much larger effort. Future research will focus on improved methods of automating the data registration, as well as additional ways to integrate traditional photogrammetric techniques into the geometrical object reconstruction. Additionally, the identification and specification of spectral signatures from registered hyperspectral data will be explored further.

## DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of the U.S. Air Force, Department of Defense, or U.S. Government.

## ACKNOWLEDGEMENTS

This work has been supported in part by the NGA under University Research Initiative HM1582-05-1-2005, "Automated Imagery Analysis and Scene Modeling."

## REFERENCES

1. Ientilucci, E. J., Brown, S. D., "Advances in Wide Area Hyperspectral Image Simulation", *Proc. of SPIE Conference on Targets and Backgrounds IX: Characterization and Representation*, Vol. 5075, April 2003.
2. Schott, J.R., Brown, S. D., Raqueno, R. V., Gross, H. N., and Robinson, G., "An Advanced Synthetic Image Generation Model and Its Application to Multi/Hyperspectral Algorithm Development", *Canadian Journal of Remote Sensing*, Vol. 25, No. 2, June 1999.
3. Burton, R., "Elastic LADAR Modeling for Synthetic Imaging Applications", Ph.D. thesis, RIT, NY, 2002.
4. Blevins, D., "Modeling Scattering and Absorption for a Differential Absorption LIDAR System", Ph.D. thesis, Rochester Institute of Technology, Rochester, NY, 2005.
5. DIRSIG User's Manual Release 4, Digital Imaging and Remote Sensing Laboratory, RIT, 2005.
6. DIRSIG homepage, <http://dirsig.cis.rit.edu/>
7. Ientilucci, E. J., Ewald, K., Marcin, J., Spivey, A., "Guide to Building Large Scale DIRSIG Scenes", Digital Imaging and Remote Sensing Laboratory, RIT, 2003.
8. "Rhinoceros: NURBS Modeling for Windows" software, information available online at <http://www.rhino3D.com>
9. "Tree Professional" software, information available online at <http://www.onyxtree.com>
10. Brown, S. D., and Schott, J. R., "Characterization Techniques for Incorporating Backgrounds into DIRSIG", *Proc. of SPIE* Vol. 4092, 2000.
11. Schenk, T., Csatho, B., "Fusion of Lidar Data and Aerial Imagery for a More Complete Surface Description", Commission III, Working Group 5 Report, Ohio State University.
12. Inglada, J. Giros, A., "On the Possibility of Automatic Multisensor Image Registration", *IEEE Trans on Geoscience and Remote Sensing*, Vol. 42, No. 10, 2104-2120, Oct 2004. – Registration
13. Ma, R., "DEM Generation and Building Detection from Lidar Data", *Photogrammetric Engineering and Remote Sensing*, vol. 71, No. 7, July 2005.
14. Rex, B., Fairweather, I., Halligan, K., "Flight Rehearsal Scene Construction from LIDAR and Multispectral Data Using ARC Spatial Analyst and 3D Analyst", *Earth Observation Magazine*, Vol 14, No 6., August 2005.
15. Gray, R., Brown, S., Schott, J., "Scene construction methodologies and techniques for simulating forest areas", 11th Annual Ground Targets Modeling and Validation Conference, August 2000.
16. Wang, R., Zhang, Y., "Extraction of Urban Road Network Using Quickbird Pan-Sharpned Multispectral and Panchromatic Imagery by Performing Edge-Aided Post-Classification", *ISPRS Joint Workshop on Spatial Temporal and Multi-dimensional Data Modeling and Analysis*, Quebec, Canada, October, 2003.