LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Description of Synthetic Model Building and Image Compositing Processes

J. Goforth, L. Gaines, W. T. White III, P. A. Pope

July 2, 2013

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Benchmark Imagery Project
Description of Synthetic Model Building and Image Compositing Processes

John W. Goforth[1], Lucinda R. Gaines[2]
W. Travis White III[1], Paul A. Pope[2]
[1]Lawrence Livermore National Laboratory
[2]Los Alamos National Laboratory

June 27, 2013

INTRODUCTION

The Benchmark Imagery Project is a DOE NA-22-funded effort whose goal is to create a suite of imagery that can be used to test algorithms that extract semantic content from overhead geospatial images of industrial facilities. This report focuses on the portion of the benchmark imagery consisting solely of composite images. It fulfills two specific project requirements for Task 1.1 of the project:

- Team delivers documentation describing the model building process (due June 2013).
- Team delivers documentation describing the image compositing process (due Sept. 2013).

The document is organized into four main sections. First, we provide background information about the use of composite images in the project. Next, we describe the model building process. Then we describe the image compositing process. Finally, we conclude with a summary of the key points and some comments about possible future directions.

1. BACKGROUND

Composite images are images that use elements of both real images (such as photographs) and synthetic objects (such as computer-generated three dimensional models). They have potential advantages over actual images of facilities, especially sensitive sites. First, they can be customized to contain features of high importance to users. For example, many different types of facilities can be generated. Moreover, the facilities can be drawn in different stages of construction and with various components and layouts. Images can be "captured" at different angles and different times of day. Effects such as fog, clouds, and lighting can be simulated. Sensor and mission characteristics such as over-exposure, distortion, blur, and flight patterns can be modeled. Second, composite images can be totally fictitious and therefore unclassified. This makes such images potentially distributable to academic researchers. Third, they can have complete ground truth. Since the scene is constructed in a computer, all elements of the world — the exact locations, shapes, sizes and function of all objects in the scene — are completely known.   Finally, composite images can be segmented and labeled automatically to create posterized, pre-segmented versions that can be used for measuring the accuracy of automated

annotation algorithms and for testing geospatial semantic processing algorithms. These features are all helpful to algorithm developers for testing and development.

The Benchmark Imagery Project focuses almost exclusively on visual-band, red-green-blue (RGB) images. This is done for a combination of reasons, both managerial (*e.g.*, to bring the scope of work in line with the resources available) and technical (*e.g.*, to focus the research effort). Although limiting the imagery to RGB precludes the use of data fusion algorithms that exploit other helpful modalities such as infrared or radar, enough different images of enough different facilities (over 200) have been compiled to enable one to demonstrate how the benchmark images could be used to test semantic extraction algorithms as applied to several classes of industrial facilities. Additionally, for composite imagery, a great advantage of restricting attention to RBG images is, it permits the use of commercially available, synthetic 3D models, as discussed in the next section.

The composite images in the Benchmark Imagery Project consist of commercially procured synthetic objects like buildings, towers, and vehicles placed on a background scene that is an actual overhead photograph of an area 1 to 2 km^2. The orthorectified background scene is draped over its actual 3D base terrain constructed from the corresponding digital elevation mapping (DEM) data. Thus, the base terrain is constructed to be geospatially accurate. Using real overhead imagery and DEM data to create the background increases the realism of the final product and decreases effort. (To construct all elements with purely synthetic textures would require extensive effort and still might not look as realistic as a real image.) As explained in Section 3, we use real time 3D rendering software "SceneWorks"[1][2] for making the composite images.

For each composite image a posterized image counterpart is generated. A posterized image is an image with every pixel classified according to the specific object in which it falls. Posterized images serve as answer keys for object segmentation algorithms by providing a pixel-to-object conversion. They also serve indirectly as input for semantic processing algorithms that need labeled, segmented images as input.

2. MODEL BUILDING PROCESS

By a *model* we mean a computer object that represents a physical structure such as a building. Our software, SceneWorks, requires three-dimensional models. These models are 3D spatially accurate geometric meshes which have been textured with RGB images. We obtain them in two different ways, as we now describe. We buy them, or we fabricate them by starting with two-dimensional overhead images.

2.1 USING EXISTING 3D COMPONENTS

Creation of highly detailed facility components should be avoided whenever possible. It is very time consuming and there are tremendous existing resources on the internet where entire 3D models of facilities can be purchased for only a few hundred dollars (see Turbosquid.com [3]) There are also free resources (see Trimble 3D Warehouse [4]) where 3D models can be freely downloaded but in our experience the quality of these models is not sufficient.

High quality models acquired from existing resources are typically downloaded in a format compatible with the 3DSMax software application [5], such as the '.MAX' extension. They also may be complete facilities. The 3DSMAX software application with the OSG exporter plug-in[6] installed can be used to break up and export the models in a format optimized for SceneWorks. Complete facilities should be broken up into components. Each component should be moved to the origin with the ground at its base (i.e. 'ground-clamped') and the origin at its center. There is no need to create multiple copies of identical components. For example, if the complete facility has two cooling towers, only one cooling tower needs to be exported as an individual component. It can be replicated in SceneWorks. The models should be exported in OSG format. This is a text based format that is read directly in SceneWorks.

3D models have textures which have different material lighting properties. Each texture has an ambient, diffuse and specular lighting component. Sometimes these material lighting properties of 3D components acquired make the lighting to look unrealistic compared to real satellite images.   These components can be changed in either 3DSMAX before export, or by modifying the exported OSG file directly. We prefer modifying the OSG file because this can be done in bulk using python scripts to change all textures in all models in a given directory at once. Results can be viewed in the context of the scene in SceneWorks almost instantly, with no modification of any other files. Visible satellite images typically have very harsh lighting compared to the lighting typically found on 3D models. Model lighting can be corrected to be more like satellite (harsh lighting) simply by increasing the diffuse lighting, and reducing the ambient lighting.

2.2   CREATING 3D COMPONENTS FROM 2D IMAGES

Some 3D components that are hard to find can be created very quickly. Components which have no height such has railroad tracks, basins and parking lots are easily created by texturing a basic 2D shape with imagery. These objects do not need 3D geometry because they are flat, therefore do not cast significant shadows and do not have significant perspective.

The procedure starts with finding an example of the object in high-resolution orthorectified imagery. This imagery is acquired from the USGS as our other imagery used in creating base terrains. The object's dimensions are measured in an online program such as Google Earth [14], or from the image geospatial data directly using a GIS program like GlobalMapper[8]. Once the dimensions are known, then the object is created in using a basic 3D modeling program, like Trimble Sketchup[7]. Even though the object is 2D, it must exist in a 3D world, so it must have some nominal height. Using Sketchup we extrude the object vertically (along the Z axis) so that the height is approximately 1/2 meter. SceneWorks will automatically disable shadow casting on objects that are less than 1 meter in height to prevent any visual artifacts.

The top of the 3D object should be textured with the image of the object in Sketchup. Before that, the texture image must be cropped out of the large orthorectified satellite image. We are using a free Photoshop-like tool called "GIMP"[9]. Once the texture image of the object is cut out and saved in PNG format[10] using GIMP, the object can then be textured on the top face of the 2D object with the texture image. The 3D object is then exported from Sketchup as a DAE (Collada) [11] file. This file can then be converted to OSG using the OSGConv tool[12] which is included with SceneWorks.

3. IMAGE COMPOSITING PROCESS

The goal of the image compositing process is to construct a 3D scene in our rendering software, SceneWorks, so that realistic visible images can be generated along with posterized counterparts. The main steps in the image compositing process are building a base terrain, laying out the facility, generating the composite image, and generating the posterized counterpart image.

3.1   BASE TERRAIN

Our goal is to create composite images which are similar to real images in the Benchmark Imagery suite. Composite images contain geospatial metadata in the same format as the real world images in the suite (in the GeoTIFF [13] file format.) The composite images should be generated with a variety of base terrain environments to test algorithms over a range of inputs. Different base terrains are essential to preventing algorithms from being mistakenly "tuned" to specific locations or environments. However, there is a trade-off between having a large variety of backgrounds and the effort required to find and populate those backgrounds with synthetic 3D models of industrial facilities. As a tractable compromise, we have decided to use 4 separate base terrains.   Each base terrain will support multiple facility types, including nuclear power plants and coal-fired power plants, the two types of facility on which our synthetic images focus.

The four different environments we selected to use as background images include sites in arid, coastal, rural/forest, and urban/suburban areas. The arid environment was previously selected from a region near Kettleman City, CA. We found our coastal background images by performing a broad area search of the United States coast line. We selected a region near Galveston, TX. For the forest background we searched the main forested areas of the continental U.S. and chose a region near Ellington, MO. For the urban background, we searched the Environmental Protection Agency website for superfund/brownfield areas that would provide an adequate clearing for a facility in an urban setting, and we selected an area near Commerce City, CO.

When selecting a location for building a base terrain there are several desirable and undesirable characteristics that must be considered. The location must contain a flat clearing of land large enough for multiple facility types (at least 100m x 100m.) The **clearing** should be **flat** in elevation (DEM) and have imagery **clear** of any discernible objects that would prevent a facility from being placed (built) there. The clearing should have no trees, buildings, or vehicles. Base imagery may contain shadows being cast from nearby trees and small buildings, but the shadows simulated in SceneWorks should match the underlying imagery. Finding the date that the imagery was taken can help in generating shadows that match, as SceneWorks generates shadows from facility 3D models using accurate ephemeris modeling. With the imagery date, real-time shadow rendering in SceneWorks is used to find the correct time of day that the generated shadows match the underlying imagery.

Google Earth [4] is very useful tool for seeking out and evaluating different locations. It displays high resolution imagery and DEM, and it is quickly navigated. Once a potential location is found in Google Earth and appears to be viable, the source data (DEM and imagery) should be acquired and

evaluated. Most imagery in Google Earth is commercial and restricted in distribution. We acquire our source data using the USGS National Map Viewer and Download Platform [5]. Using the USGS website we acquire a 5km x 5km region of imagery and DEM centered on the desire location of the facility. For examples of images reviewed and more information on selecting base imagery please see "Selecting Real Imagery for Generating Composite Images" Goforth, et al. [16].

The source data (imagery and DEM) must be converted into a format that optimized for 3D real-time rendering. Using SceneWorks' terrain generator, the input files are selected and the 3D terrain is generated in only several minutes. This 3D real-time format is an OpenSceneGraph terrain format (IVE format.) SceneWorks also uses the geospatial metadata from the input files to configure its internal GIS engine and coordinate system. The SceneWorks scene map is now saved and ready for facility components to be added.

3.3   FACILITY LAYOUT

There is a moderate amount of effort in creating a realistic layout of a facility, which includes reviewing that layout with SMEs. The Benchmark Imagery Suite contains two types of facilities - a nuclear power plant and a coal-fired power plant. These two facilities are each placed in 4 different environments: arid, coastal, urban/suburban, and rural/forest. The facilities must be modified slightly for each environment. For example, cooling towers are not needed when a power plant is positioned on the ocean in a costal environment.

Our approach for facility layout is to select one or more real facilities from our benchmark imagery set as a template for our layout. The selection of these facilities is based largely on finding a real facility that contains components that are similar to the 3D models we have acquired and created.

The coal power plant layout is based on:

- Bailly Generating Station, Chesterton, IN (Benchmark Image HP52)

Comparison of real and synthetic images can be seen in figure 1 and 2.

Figure 1 - Bailly Generating Station (HP52)

Figure 2 - Synthetic coal power plant in coastal environment, based on layout in Figure 1

The nuclear power plant layout is based on:

- Bellefonte Nuclear Plant, Hollywood, AL (Benchmark Image HP104)
- Duke Energy Mcguire Nuclear Station, Huntersville, NC (Benchmark Image HP119)

Comparison of real and synthetic images can be seen in figures 3-5.

Figure 3 - Bellefonte Nuclear Plant (HP104)

Figure 4 - Duke Energy Mcguire Nuclear Station (HP119)

Figure 5 - Synthetic nuclear power plant in arid environment based on layout in figures 3 and 4.

3.2   IMAGE GENERATION

Once the facility layout is established, the 3D models are placed in the scene using the SceneWorks STAGE tool. The models must be ground clamped to ensure there are no gaps between the building and ground. If model is not flush with the ground, the shadows will not cast directly from the base of the object because the object is above the ground. Simulated time of day and date should be set in SceneWorks so that shadows match the underlying base imagery.

In SceneWorks, the simulated camera is positioned with nadir view at an altitude high enough to allow for a large amount of lens magnification using the zoom feature. This removes a large amount of perspective, simulating an image taken from a satellite platform. At this point the scene has been constructed and the settings configured, so generating the composite image is trivial. The image capture button is pressed and the composite image is automatically generated and saved to disk in the GeoTIFF format.

3.3   POSTERIZATION

In a posterized image all pixels have been classified by object type. For example, all the pixels that make up a cooling tower are labeled as 'cooling tower'. Posterized images are classified using colors, and have corresponding text files (CSV format) which key colors to objects. Posterized images also contain geospatial metadata in the GeoTIFF format.

To allow for posterization abstraction levels, the objects are named (using the STAGE tool,) according to our specific naming convention. The naming convention is the object class first, then the specific object number (i.e. "OBJECT_CLASS-OBJECT_NUMBER.") This enables the posterization engine to create multiple levels of abstraction. For example "Hangar-2" would be colored a singled color in the posterization, and would be keyed with three different levels of specificity:
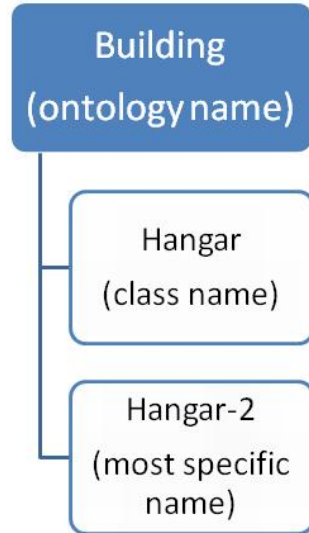
Having different levels of object abstraction allows for algorithms to using the posterized images at different levels of detail. For example, if two buildings are connected to each other, some algorithms just mark the complete structure as a "building," while others segment those two buildings apart and define the specific type of building (e.g. Hanger-1 and Hanger-2.) The posterized image has complete ground-truth detail, but it may be abstracted out using the CSV file if needed.

SceneWorks will automatically posterize the image with a single button click. The code for posterization in SceneWorks is based on a few key OpenGL calls which are trigged through the scene graph. The posterization is achieved through force overriding the OpenSceneGraph (OSG) node of each object to disable texturing of the object. Lighting and shading is also disabled in the node so that there is no shading. A color is selected for the object by iterating through the objects and selecting a different hue in the RGB color space for each object. Finally, the object is colored by replacing the color arrays on the node and all of its children. The base terrain is rendered all black.

When generating a posterized image it is important to disable antialiasing in the computer's graphics card settings. Antialiasing softens the edges by blending them but the result is additional colors in the posterized image that are not defined. Antialiasing is very important in producing realistic visible imagery, so it should only be disabled when producing posterized images.

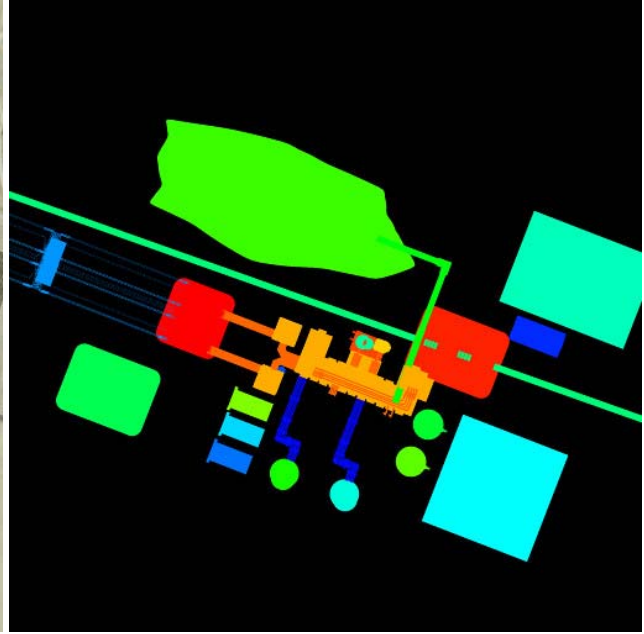**Figure 6. Synthetic Image of Coal-fired Power Plant in a Forest Environment**

Figure 7. Posterized Counterpart to Synthetic Image in Figure 6

Table 1 - Posterization Key (CSV file) with different levels of pixel classification and corresponding colors for Figure 2

Red	Green	Blue	Ontology Name	Object Class	Object Name
255	34	0	Building	hangar	hangar-1
255	68	0	Building	stackbldg	stackbldg-1
255	102	0	Building	bldgpipes	bldgpipes-1
255	174	0	Building	smallbldg	smallbldg-1
255	209	0	Stack	stacksmall	stacksmall-1
255	243	0	Conveyor	pipes	pipes-1
128	255	0	Building	hangarsmall	hangarsmall-2
93	255	0	Pond	pool	pool-1
59	255	0	Pile	ashpile	ashpile-1
21	255	0	Cooling Tower	coolingtower	coolingtower-2
0	255	12	Conveyor	conveyor	conveyor-1
0	255	46	Pond	pool	pool-2
0	255	81	Building	hangar	hangar-2
0	255	119	Railline	rail	rail-1
0	255	153	Stack	stacklarge	stacklarge-1
0	255	187	N/A	parkinglot	parkinglot-1
255	0	0	Electrical Substation	substation	substation-1
0	255	221	Cooling Tower	coolingtower	coolingtower-1
0	255	255	Pond	basin	basin-1
0	217	255	Building	hangarsmall	hangarsmall-1
0	149	255	N/A	powerlines	powerlines-1
0	115	255	Building	hangarsmall	hangarsmall-3

Red	Green	Blue	Ontology Name	Object Class	Object Name
0	76	255	Water Tower	watertower	watertower-1
0	42	255	Building	bldg	bldg2-1
0	8	255	Conveyor	pipes	pipes-2

4. CONCLUSION

4.1 Summary of Processes

We have outlined the processes used in the Benchmark Imagery Project to create synthetic models of facilities and the corresponding composite imagery. Imagery and Digital Elevation Maps (DEM) are the input for generating the base terrain. The base terrain is a 3D model which has been optimized for real-time 3D rendering in which the imagery has been textured to the elevation data. 3D models of facility components are built or acquired from internet public sources. Facility components are placed on the base terrain consistent with real facilities and input from subject matter experts (SMEs.) Basic parameters such as camera position, time of day, and model material lighting are configured to match the underlying imagery. Time of day and date is set so that shadows generated in SceneWorks match the underlying imagery. Allowing only a single time of day and date is a current limitation of our process, and removing this limitation is an area for improvement.

Once the scene configuration is complete, visible high-resolution images are generated with resolution and GSD (Ground Sample Distance) approximately equal to that of real images in the benchmark imagery set. Camera (e.g. sensor) elevation along with basic lens modeling is used in SceneWorks to simulate the reduction in perspective which is apparent in real images taken from satellite platforms. The visible images are encoded with geospatial data to match the base terrain real location in the GeoTIFF format. Posterized counterpart images are then generated to provide "answer key" pixel-to-object classification.

4.2 Composite Images for the Benchmark Imagery Set

The goal of the Benchmark Imagery Project is to create a suite of imagery (both real and composite) that can be used to test geospatial algorithms used for extracting semantic content from overhead images of industrial facilities. Our goal this year is to create composite images that will appear realistic to both human and algorithm. To augment realism in our composite images, the real world facilities served as a template for our layout. The selection of these facilities was based largely on finding a real facility that contained components that were similar to the 3D models we had acquired and created.

We've generated eight composite images comprising of two facility types (nuclear power plant and coal fired power plant). The two types of facilities are placed in four different environments (arid, coastal, forest/rural, and urban/suburban). By attaining different base terrains will prevent algorithms from being mistakenly "tuned" to specific locations or environments. Additionally, we've generated

eight corresponding posterized images, one for each composite image. Posterized images can serve two different purposes, one as "answer keys" for object segmentation algorithms. Secondly, they can offer guidance for semantic algorithms which need segmented images are input. Lastly, all images include geospatial metadata in the same format (GeoTIFF) as the real benchmark imagery set. Currently, the benchmark imagery suite includes over 200 real world facilities, 8 composite images, and 8 corresponding posterized images.

Composite images have the advantage of being customizable, distributable, and they contain complete ground truth. This is extremely helpful to algorithm developers for testing and development. The model building and image compositing processes used in this project have proven effective, in terms of being straight forward and yielding realistic results, at creating a small suite of composite images for training and testing geospatial, semantic-extraction algorithms.

4.3   Path Forward

We will explore the ability to create suites of images by varying the layout quickly. Using a layout generation algorithm, site layouts can be randomized while obeying semantic rules that are derived from both real world facilities and subject matter expertise. Our rendering software, SceneWorks, has the ability to ingest layout files in the form of KML[17] and position facility components automatically. This workflow enables the generation of many composite images with different facility layouts quickly. Having a large image set with different layouts of the same facility type would help algorithm developers identify and characterize problems in both object segmentation and semantic extraction.

Based on our discussions with algorithm developers, the Benchmark Imagery Suite would certainly benefit from simulated images of other modalities such as hyperspectral, infrared or radar. Different approaches will have to be used to generate these images because commercially available 3D models have textures which are for the most part in the visible modality.

Finally, the dissemination of the benchmark imagery suite to developers is essential to establishing a feedback loop to improve the benchmark imagery set. Ultimately the goal of the benchmark imagery set is to drive and support the improvement of existing algorithms and the creation of new algorithms in the areas of both object segmentation and semantic extraction.

References

1. **Peterson, Jeffery.** (2010) *"Predicting Pollutant Dispersion and Exposure in Mixed Outdoor-Indoor-Vehicle Environments"* US Army Dugway Proving Ground. Chemical and Biological Defense Science, and Technology Conference, Orlando Florida.
   [Predicting Pollutant Dispersion and Exposure in Mixed Outdoor-Indoor-Vehicle Environments](#)

2. **Gaurang, Dave.** (2010) *"Joint Expeditionary Collective Protection System Performance Model (JECP SPM) Overview"* Naval Surface Warfare Center, Dahlgren. Chemical and Biological Defense Science and Technology Conference, Orlando Florida. [Joint Expeditionary Collective Protection System Performance Model (JECP SPM) Overview](#)

3. Turbosquid online 3D model market place. http://www.turbosquid.com

4. Trimble 3D Warehouse. (Formerly Google 3D Warehouse.)
   http://sketchup.google.com/3dwarehouse/

5. 3DSMax software application http://www.autodesk.com/products/autodesk-3ds-max/overview

6. OSG Exporter Plugin from Delta3D
   http://www.delta3d.org/forum/viewtopic.php?showtopic=17151

7. Trimble Sketchup software application. http://www.sketchup.com/

8. GlobalMapper software application. http://www.bluemarblegeo.com/products/global-mapper.php

9. GNU Image Manipulation Program (aka "GIMP") http://www.gimp.org/

10. PNG format. https://en.wikipedia.org/wiki/Portable_Network_Graphics

11. COLLADA format. https://collada.org/

12. OSGConv Tool from OpenSceneGraph.
    http://trac.openscenegraph.org/projects/osg//wiki/Support/UserGuides/osgconv

13. **Ruth, Mike**. (2011). *"GeoTIFF FAQ Version 2.4"* http://remotesensing.org/geotiff. Last date accessed May 29, 2013.

14. **Author Unknown**. (2013) "Google Earth" http://www.google.com/earth/index.html. Last date accessed May 29, 2013.

15. **Author Unknown.** (2013) "The USGS National Map Viewer and Download Platform" http://nationalmap.gov/viewer.html. Last date accessed May 29, 2013.

16. **Goforth, John, et al**. (2013) "Selecting Real Imagery for Generating Composite Images"

17. KML file format, https://developers.google.com/kml/documentation/