

# A Building Model Evaluation Suite Using the CMU Site Exchange Format

George E. Bulwinkle      Jefferey A. Shufelt

December 1, 1998

CMU-CS-98-133

School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213-3890

## Abstract

This document describes a set of tools for evaluating building extraction results for aerial imagery, developed at the Digital Mapping Laboratory at Carnegie Mellon University. These tools compute quantitative measurements of performance by comparing a manually measured reference model of a scene against an automatically generated model for the same scene. Summary metrics are generated for the entire scene, as well as building-by-building coverage statistics. The tools accept reference model and test model descriptions in the CMU Site Exchange format.

This work was sponsored by the Advanced Research Projects Agency under Contracts DACA76-95-C-0009 and DACA76-97-K-0004 monitored by the U.S. Army Topographic Engineering Center, Fort Belvoir, VA. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Advanced Research Projects Agency, the U.S. Army Topographic Engineering Center, or the United States Government.

The Digital Mapping Laboratory's WWW Home Page may be found at: <http://www.cs.cmu.edu/MAPSLab>

**Keywords:** building extraction,evaluation metrics

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Usage and distribution disclaimer	1
<b>2</b>	<b>Overview</b>	<b>2</b>
<b>3</b>	<b>Preparing models for evaluation</b>	<b>2</b>
<b>4</b>	<b>Running the evaluation metrics</b>	<b>4</b>
4.1	2D image space evaluation with <code>eval2d</code>	4
4.2	3D object space evaluation with <code>eval3d</code>	4
<b>5</b>	<b>Interpreting the evaluation results</b>	<b>4</b>
<b>6</b>	<b>Example Evaluation Results</b>	<b>5</b>
6.1	Sample output from <code>eval2d</code> on BUILD + SHAVE results	6
6.2	Sample output from <code>eval3d</code> on BUILD + SHAVE results	8
6.3	Sample output from <code>eval2d</code> on VHBUILD results	9
6.4	Sample output from <code>eval3d</code> on VHBUILD results	10
<b>7</b>	<b>The CMU Site Exchange Format</b>	<b>11</b>



# 1 Introduction

This document describes a set of tools for evaluating building extraction results for aerial imagery, developed at the Digital Mapping Laboratory at Carnegie Mellon University. These tools compute quantitative measurements of performance by comparing a manually measured reference model of a scene against an automatically generated model for the same scene. Summary metrics are generated for the entire scene, as well as building-by-building coverage statistics. The tools accept reference model and test model descriptions in the CMU Site Exchange format.

The user is assumed to have a basic knowledge of UNIX, and is also assumed to have the capability to export building models into the CMU Site Exchange Format, described in “Robust Exchange of Cartographic Models for Buildings and Roads: The CMU MAPSLab Site Exchange Format.” This document is available from the Digital Mapping Laboratory web site (specific URLs can be found in this document, in Section 7).

This document is organized as follows. Section 2 presents a high-level overview of the evaluation tools, and Section 3 describes the additional information required in site exchange files by the evaluation tools. Section 4 illustrates the use of the tools, and Section 5 explains the metrics computed by the tools and their interpretation. Finally, Section 6 provides examples of the tools applied to building extraction results.

Our research on the evaluation metrics package was supported by the Defense Advanced Research Projects Agency (DARPA/ISO) under Contracts DACA76-95-C-0009 and DACA76-97-K-0004 and is monitored by the U.S. Army Topographic Engineering Center. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the U.S. Army Topographic Engineering Center, or of the United States Government.

## 1.1 Usage and distribution disclaimer

The software described in this manual is the by-product and not the primary focus of our program of research. It is not intended to comprise a complete end-to-end building model evaluation capability nor is its use supported by the CMU MAPSLab. No warranty, either expressed or implied, is provided by Carnegie Mellon University for the use of this software. All rights are reserved, including commercial license, by the Digital Mapping Laboratory, Carnegie Mellon University, Pittsburgh, PA.

Distribution of this software is to the U.S. Government to support test and evaluation. Permission for its use must be secured by contacting Mr. Doug Caldwell, U.S. Army Topographic Engineering Center, (caldwell@tec.army.mil). The DARPA point of contact for this research is Mr. George Lukes, Program Manager for Image Understanding, DARPA/ISO (glukes@darpa.mil). The MAPSLab point of contact for questions regarding this software and its distribution is Dave McKeown (dmm@cs.cmu.edu).

Any redistribution of this document and associated software must contain this disclaimer, without alteration.

## 2 Overview

As the capabilities of automated and semi-automated building extraction systems mature, rigorous performance evaluation of these systems becomes increasingly important, to assess the strengths and weaknesses of various techniques. Two traditional difficulties in carrying out evaluation have been the lack of a robust format for interchanging building models, and a set of tools for computing unbiased and meaningful evaluation metrics on site models. The first problem has been addressed by the CMU Site Exchange format, allowing complex generic buildings in geographic coordinates to be represented in an easily exchanged ASCII format, and read/written by a provided API.

This document addresses the second problem, describing a set of tools for evaluating building extraction results for aerial imagery, developed at the Digital Mapping Laboratory at Carnegie Mellon University for research on building extraction systems [7, 4, 6, 2, 1]. The tools accept reference model and test model descriptions in the CMU Site Exchange format. These tools compute quantitative measurements of performance by comparing a manually measured reference model of a scene against an automatically generated model for the same scene. Summary metrics are generated for the entire scene, as well as building-by-building coverage statistics.

The tools described in this document run on SGI workstations running IRIX 6.2 or greater; memory requirements depend on the size of the site models to be evaluated, but at least 96Mb of memory is recommended.

## 3 Preparing models for evaluation

The evaluation tools described in this document require site model descriptions in the CMU Site Exchange Format, with additional attribution to be described shortly. This ASCII file format is fully described in a technical report, "Robust Exchange of Cartographic Models for Buildings and Roads: The CMU MAPSLab Site Exchange Format," available from the Digital Mapping Laboratory web site in both PDF and compressed PostScript:

<http://www.cs.cmu.edu/~MAPSLab/rcvw/terrainweek98/cmutr-98-134.pdf>  
<http://www.cs.cmu.edu/~MAPSLab/rcvw/terrainweek98/cmutr-98-134.ps.gz>

In addition to providing both a reference model and a test model as files in site exchange format, there are three additional requirements for these files, imposed by the evaluation software:

1. **The reference model and test model site exchange files must use the same names for images specified in the image block of the world block.** Points in site exchange files can contain image measurements from multiple images, so the 2D evaluation software requires that an image be specified to ensure that the correct image measurements are extracted from each site exchange file for comparison. This in turn requires that the two site exchange files use the same names for images.

Section 2.2 of the site exchange technical report describes the format of the image block, where image names are defined.

The following example shows an image block defining one image, with the name 'fhov927'. For the evaluation software to operate correctly, both the reference and test model site exchange files would have to use the name 'fhov927' for the image.

```
Begin images::  
  Number of Images: 1
```

```
Image 0: fhov927
Header 0: fhov927.tec
End images
```

2. **Each point in the site exchange file must have an image measurement for the image in which the 2D evaluation will take place.** These are defined in the Point Block (Section 3.3 of the site exchange technical report) for each point. Since the evaluation metrics do not take point measurement errors into account, **sigma** can be set to 1.0 for every point in the site exchange file.

By requiring the site exchange files to contain image measurements, problems arising from differences in sensor modeling among research groups are avoided, since the world-to-image projections have already been performed.

As a reminder, point coordinates in the CMU Site Exchange format are expressed in a row-column coordinate system. The image origin, (0,0), is taken to be the upper left corner of the raster, with rows increasing down the vertical axis and columns increasing to the right along the horizontal axis. For an image of  $N$  rows by  $M$  columns, the pixel in the lower right corner of the image is indexed by  $(N - 1, M - 1)$ .

The following example shows point measurement information inside a Point Block, in a site exchange file. In this example, the point has been measured on only one image, and is located at row 3852, column 3464 of image 0, with a sigma of 1:

```
Number of Image Measurements: 1
image 0: 3852 3464 1
```

3. **If more than one wireframe volume is intended to comprise a single building, then the “bldg\_group” attribute must be defined in the Attribute Block of each Building Block corresponding to the individual volumes.** The value of this attribute is user-defined; if two building blocks have the same value of `bldg_group`, then they are considered to be the same building for the purposes of evaluation. Building blocks that do not have this attribute will be considered unique buildings for the purposes of evaluation.

Note that the constraint blocks in the site exchange format have more general semantics than the “`bldg_group`” attribute. In particular, the existence of a coplanar constraint between two surfaces does not imply that the surfaces touch; an example might be a series of barracks which have been lined up by the use of coplanar constraints.

Section 1.3 of the site exchange technical report describes attribute blocks and their format; Section 3 illustrates their use inside a building block to store attributes for a building.

The following shows the attribute block for a building block which represents part of a composite structure. The ‘`bldg_group`’ attribute is set to “Complex 1” in each building block which makes up the composite structure.

```
Begin attributes::
  Number of Attributes: 1
  bldg_group: Complex 1
End attributes
```

## 4 Running the evaluation metrics

The evaluation distribution provides two high-level evaluation programs, `eval2d` and `eval3d`, for performing 2D and 3D building evaluations respectively. This section describes both programs and their usage.

### 4.1 2D image space evaluation with `eval2d`

`eval2d` is invoked with the following command line arguments:

```
eval2d <reference.sitexg> <test.sitexg> <evaluation.txt> <image> <-r sr er -c sc ec>
```

where `reference.sitexg` and `test.sitexg` are the reference and test models in site exchange formats, respectively; `image` is the name of an image, indicating which set of image measurements will be extracted from each site exchange file; and `sr`, `er`, `sc`, `ec` specify the bounding coordinates of an image rectangle within which the evaluation will take place, in terms of start/end row and start/end column respectively. `evaluation.txt` is the output of `eval2d`, a text file which contains the evaluation statistics.

2D evaluation is relatively fast; on four evaluation runs using an SGI Indigo 2, `eval2d`'s runtimes ranged from 45 seconds to 3.5 minutes.

### 4.2 3D object space evaluation with `eval3d`

`eval3d` is invoked with the following command line arguments:

```
eval3d <reference.sitexg> <test.sitexg> <evaluation.txt> [-r <resolution>]
```

where `reference.sitexg` and `test.sitexg` are the reference and test models in site exchange formats, respectively. `eval3d` discretizes object space into cubic voxels 0.5 meters to a side; the optional argument `resolution` specifies a different size for these voxels. The default of 0.5 meters represents a reasonable tradeoff between evaluation accuracy and evaluation runtime. `evaluation.txt` is the output of `eval3d`, a text file which contains the evaluation statistics.

3D evaluation can be time consuming, due to the intensive nature of the geometric computations for discretizing the site models. On four evaluation runs using an SGI Indigo 2, `eval3d`'s runtimes ranged from 1 to 4 hours.

## 5 Interpreting the evaluation results

In both 2D (image space) and 3D (object space) evaluations, space is divided into pixels and voxels, respectively (hereafter, we just use pixel for brevity). Each pixel is then classified as building or background by the reference and test models, leading to four possible values for each pixel:

- true positive (TP): both the reference and test model classify the pixel as building
- true negative (TN): both the reference and test model classify the pixel as background



- false positive (FP): only the test model classifies the pixel as building
- false negative (FN): only the test model classifies the pixel as background

Three summary statistics are then computed from the counts of each of these four pixel types:

- Building detection percentage:  $100 \text{ TP} / (\text{TP} + \text{FN})$
- Branching factor:  $\text{FP} / \text{TP}$
- Quality percentage:  $100 \text{ TP} / (\text{TP} + \text{FP} + \text{FN})$

The building detection percentage and quality percentage metrics range from 0-100, where 100 is considered to be a perfect score. The branching factor ranges from 0 to infinity, where 0 is a perfect score.

Intuitively, the building detection percentage measures the extent to which buildings have been detected, or covered, by the test models. The branching factor measures the ratio of false positives to correctly detected building pixels; for example, a branching factor of 2 means that a system incorrectly labeled two background pixels as building for every building pixel it correctly labeled. The quality percentage metric measures both missed building pixels and false positives to provide a single performance metric for a scene.

The individual building coverage percentages simply represent the percentage of pixels in each building which have been detected. Note that these numbers do not take into account false positives, which essentially means that precise delineation is not measured by these percentages. As one example, consider a system which generates one building, covering the entire scene. This will produce 100% building coverage percentages for every building in the reference model, although it leaves a great deal to be desired in terms of delineation accuracy. As a consequence, these numbers should be weighed against the global summary statistics (building detection percentage, branching factor, quality percentage) to obtain an accurate picture of overall system performance.

## 6 Example Evaluation Results

This section contains sample results obtained by running the evaluation software on several monocular extraction results. These systems were run on the vanilla area of Fort Hood as shown in Figure 1(a). The reference data for the scene is shown in Figure 1(b). Results from two monocular extraction systems are shown: BUILD + SHAVE (Figure 1(c)) and VHBUILD (Figure 1(d)). BUILD [5] is a line and corner-based analysis system which operates solely in image space. BUILD + SHAVE [3] incorporates shadow analysis to perform structure evaluation, while VHBUILD [4] performs vertical/horizontal analysis in conjunction with the standard BUILD processing flow.

Results for the evaluation of BUILD + SHAVE are given in Sections 6.1 and 6.2. Results for the evaluation of VHBUILD are given in Sections 6.3 and 6.4.

It took approximately 30 seconds to run the 2D evaluations for each of the test scenes and approximately an hour to run the 3D evaluations for each.

The building group “Complex 1” is the L-shaped building group at the upper left hand corner of the scene; “Complex 2” is the L-shaped building located in the top-center of the scene; and “Complex 3” is the building with overhang located in the left-center of the scene.

The remaining buildings are single building objects and are named as follows:

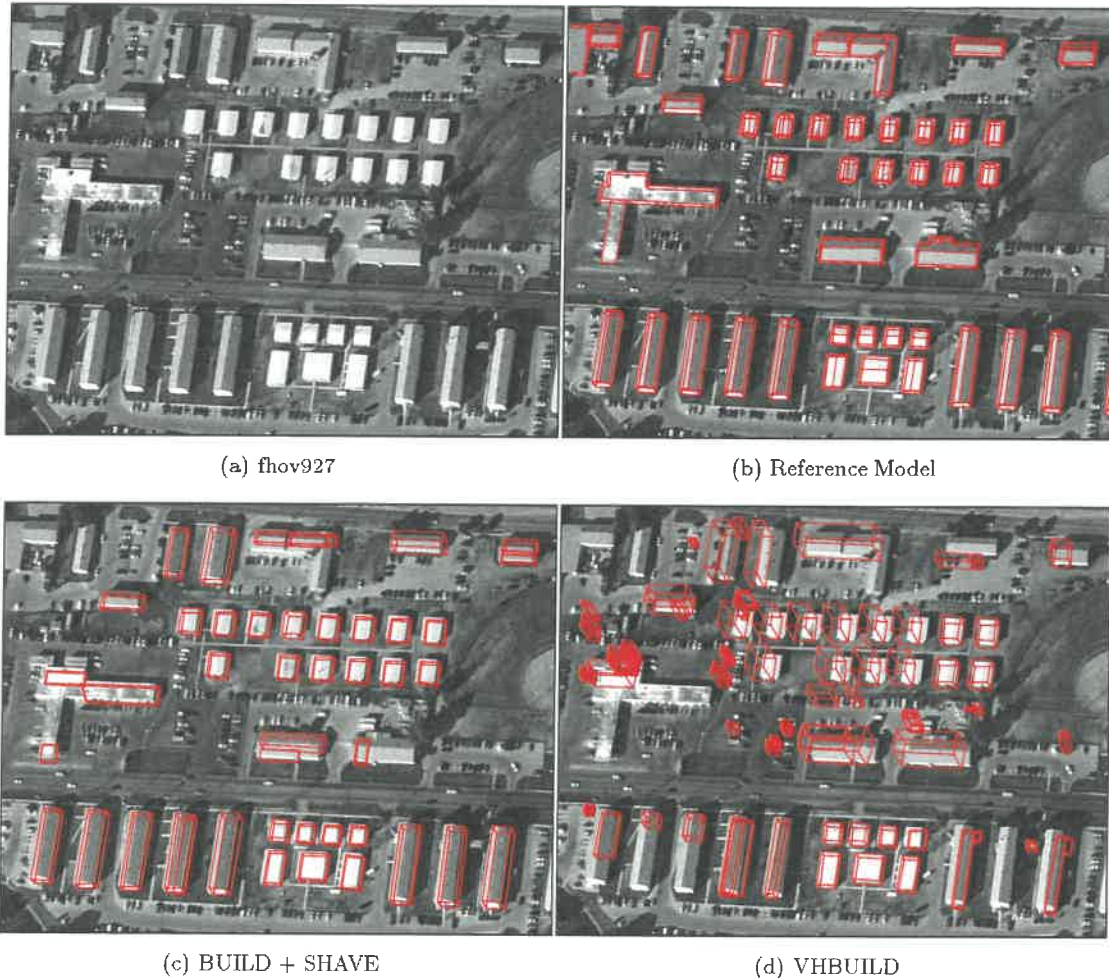


Figure 1: Vanilla scene

There are a group of large peak roof buildings along the bottom of 1(b) interrupted by a cluster of smaller peak roof buildings. The large peak roof buildings are named “Long Peak 1” through “Long Peak 9”. The smaller peak roof buildings are named “Short Peak 1” through “Short Peak 7”. Numbering of the buildings is done from left to right, top to bottom in the figure.

There are a cluster of overhang peak roof buildings in the center of 1(b). These are named “Overhang Set 1” through “Overhang Set 14”. Again, numbering is done from left to right top to bottom.

The remaining buildings are labeled from left to right as they are encountered in 1(b) by building type: “Rectilinear 1”, “Peak 1” through “Peak 7”, “Generic Peak 1”, and “Overhang Peak 1”.

## 6.1 Sample output from eval2d on BUILD + SHAVE results

The following is output produced by eval2d on BUILD + SHAVE results (Figure 1(c)).

```
# of pixels: 254666 # of background: 209365 # of building: 45301
Quality: 0.600443
```

```

% of scene that is buildings: 0.177884
Correctly detected background pixels: 200889
Correctly detected building pixels: 32290
Building pixels missed: 13011
Background pixels incorrectly classified: 8476

% correctly classified pixels: 91.5627
% incorrectly classified pixels: 8.43733
% buildings detected: 71.2788
% buildings missed: 28.7212
% background detected: 95.9516
% background missed: 4.04843
% incorrect pixels that are false positives: 39.4471
building pixel branching factor: 0.262496

```

Individual building coverage percentages:

```

Complex 1: 0 (0/1877 pixels)
Complex 2: 42.3945 (1165/2748 pixels)
Complex 3: 52.6964 (2140/4061 pixels)
Generic Peak 1: 18.859 (357/1893 pixels)
Long Peak 1: 87.851 (1533/1745 pixels)
Long Peak 2: 88.255 (1578/1788 pixels)
Long Peak 3: 83.3068 (1572/1887 pixels)
Long Peak 4: 88.9898 (1665/1871 pixels)
Long Peak 5: 87.4934 (1651/1887 pixels)
Long Peak 6: 87.3844 (1607/1839 pixels)
Long Peak 7: 84.3972 (1547/1833 pixels)
Long Peak 8: 91.7481 (1690/1842 pixels)
Long Peak 9: 0 (0/363 pixels)
Overhang Peak 1: 91.2707 (826/905 pixels)
Overhang Set 1: 88.0716 (443/503 pixels)
Overhang Set 10: 82.9868 (439/529 pixels)
Overhang Set 11: 84.9341 (451/531 pixels)
Overhang Set 12: 81.6162 (404/495 pixels)
Overhang Set 13: 76.8657 (412/536 pixels)
Overhang Set 14: 79.8893 (433/542 pixels)
Overhang Set 2: 82.1429 (414/504 pixels)
Overhang Set 3: 79.5019 (415/522 pixels)
Overhang Set 4: 91.791 (492/536 pixels)
Overhang Set 5: 79.1339 (402/508 pixels)
Overhang Set 6: 83.0476 (436/525 pixels)
Overhang Set 7: 78.5441 (410/522 pixels)
Overhang Set 8: 79.2593 (428/540 pixels)
Overhang Set 9: 79.3738 (431/543 pixels)
Peak 1: 0 (0/953 pixels)
Peak 2: 77.7393 (674/867 pixels)
Peak 3: 85.084 (810/952 pixels)
Peak 4: 89.7705 (1369/1525 pixels)
Peak 5: 85.6818 (1508/1760 pixels)
Peak 6: 96.5829 (961/995 pixels)
Peak 7: 0 (0/0 pixels)
Rectilinear 1: 0 (0/0 pixels)
Short Peak 1: 90.9524 (382/420 pixels)
Short Peak 2: 78.5388 (344/438 pixels)
Short Peak 3: 77.0115 (335/435 pixels)
Short Peak 4: 84.8341 (358/422 pixels)
Short Peak 5: 80 (656/820 pixels)
Short Peak 6: 86.2069 (850/986 pixels)
Short Peak 7: 82.2978 (702/853 pixels)

```

As this example shows, a variety of summary statistics are produced by `eval2d`. The building detection percentage, branching factor, and quality percentage in this example were given by the following lines in the output:

- *building detection percentage*: % buildings detected: 71.2788

- *branching factor*: building pixel branching factor: 0.262496
- *quality percentage / 100*: Quality: 0.600443

Note that the quality number produced by `eval2d` is a fraction, and must be multiplied by 100 to obtain a percentage. In this case, the actual quality percentage would be 60.0443.

The individual building coverage percentages are given at the end of the output, in alphabetical order by building name: for example, the building coverage percentage for the building entitled “Complex 3” is 52.6964. The ratio of building pixel coverage is given after the percentage. Using the same example, the ratio of building pixel coverage for the “Complex 3” building is 2140/4061 pixels.

## 6.2 Sample output from `eval3d` on BUILD + SHAVE results

The following is output produced by `eval3d` on BUILD + SHAVE results (Figure 1(c)).

```

True positive voxels: 244314
True negative voxels: 9061248
False positive voxels: 157160
False negative voxels: 274531
(tp / (tp + fn)) Detection pct: 47.0881
(fp / tp) Voxel branch factor: 0.643271
(fn / tp) Voxel miss factor: 1.12368
(tp / (tp + fp + fn)) Quality pct: 36.1409
((fn + fp) / tp) : 1.76695
(tp / (tp + fp)) : 0.608543

Individual building coverage percentages:
Complex 1: 0 (0/50311 voxels, 0/6288.88 m^3)
Complex 2: 9.20782 (2148/23328 voxels, 268.5/2916 m^3)
Complex 3: 25.3571 (7260/28631 voxels, 907.5/3578.88 m^3)
Generic Peak 1: 1.94416 (337/17334 voxels, 42.125/2166.75 m^3)
Long Peak 1: 62.7758 (14085/22437 voxels, 1760.62/2804.62 m^3)
Long Peak 2: 76.4605 (17446/22817 voxels, 2180.75/2852.12 m^3)
Long Peak 3: 75.2383 (17842/23714 voxels, 2230.25/2964.25 m^3)
Long Peak 4: 86.4716 (21157/24467 voxels, 2644.62/3058.38 m^3)
Long Peak 5: 92.3396 (24048/26043 voxels, 3006/3255.38 m^3)
Long Peak 6: 61.8008 (14517/23490 voxels, 1814.62/2936.25 m^3)
Long Peak 7: 68.5039 (15760/23006 voxels, 1970/2875.75 m^3)
Long Peak 8: 94.1543 (23854/25335 voxels, 2981.75/3166.88 m^3)
Long Peak 9: 0 (0/24675 voxels, 0/3084.38 m^3)
Overhang Peak 1: 60.5084 (6308/10425 voxels, 788.5/1303.12 m^3)
Overhang Set 1: 84.6277 (2637/3116 voxels, 329.625/389.5 m^3)
Overhang Set 10: 60.1354 (1866/3103 voxels, 233.25/387.875 m^3)
Overhang Set 11: 62.4538 (1690/2706 voxels, 211.25/338.25 m^3)
Overhang Set 12: 33.3333 (1020/3060 voxels, 127.5/382.5 m^3)
Overhang Set 13: 11.4895 (361/3142 voxels, 45.125/392.75 m^3)
Overhang Set 14: 11.3591 (341/3002 voxels, 42.625/375.25 m^3)
Overhang Set 2: 48.0052 (1480/3083 voxels, 185/385.375 m^3)
Overhang Set 3: 37.5 (930/2480 voxels, 116.25/310 m^3)
Overhang Set 4: 100 (2885/2885 voxels, 360.625/360.625 m^3)
Overhang Set 5: 22.3276 (612/2741 voxels, 76.5/342.625 m^3)
Overhang Set 6: 33.678 (912/2708 voxels, 114/338.5 m^3)
Overhang Set 7: 22.9204 (642/2801 voxels, 80.25/350.125 m^3)
Overhang Set 8: 23.6296 (694/2937 voxels, 86.75/367.125 m^3)
Overhang Set 9: 70.7692 (2576/3640 voxels, 322/455 m^3)
Peak 1: 0 (0/7963 voxels, 0/995.375 m^3)
Peak 2: 40.0974 (3705/9240 voxels, 463.125/1155 m^3)
Peak 3: 39.5609 (2865/7242 voxels, 358.125/905.25 m^3)
Peak 4: 73.9415 (16049/21705 voxels, 2006.12/2713.12 m^3)
Peak 5: 46.8917 (7543/16086 voxels, 942.875/2010.75 m^3)

```

```

Peak 6: 94.2201 (8004/8495 voxels, 1000.5/1061.88 m^3)
Peak 7: 0 (0/16686 voxels, 0/2085.75 m^3)
Rectilinear 1: 0 (0/10818 voxels, 0/1352.25 m^3)
Short Peak 1: 93.0108 (2768/2976 voxels, 346/372 m^3)
Short Peak 2: 55.8959 (1825/3265 voxels, 228.125/408.125 m^3)
Short Peak 3: 50.1737 (1444/2878 voxels, 180.5/359.75 m^3)
Short Peak 4: 47.6378 (1452/3048 voxels, 181.5/381 m^3)
Short Peak 5: 91.3355 (5608/6140 voxels, 701/767.5 m^3)
Short Peak 6: 71.0483 (5978/8414 voxels, 747.25/1051.75 m^3)
Short Peak 7: 56.6286 (3665/6472 voxels, 458.125/809 m^3)

```

Again, a variety of summary statistics are produced by `eval3d`. The building detection percentage, branching factor, and quality percentage in this example were given by the following lines in the output:

- *building detection percentage*: (tp / (tp + fn)) Detection pct: 47.0881
- *branching factor*: (fp / tp) Voxel branch factor: 0.643271
- *quality percentage*: (tp / (tp + fp + fn)) Quality pct: 36.1409

The individual building coverage percentages are given at the end of the output, in alphabetical order by building name: for example, the building coverage percentage for the building entitled “Complex 3” is 25.3571. The ratio of building voxel coverage as well as the ratio of building volume coverage in cubic meters is given after the building coverage percentage. Using the building entitled “Complex 3” again, the respective ratios are 7260/28631 voxels and 907.5/3578.88 meters cubed.

### 6.3 Sample output from `eval2d` on `VHBUILD` results

The following is output produced by `eval2d` on `VHBUILD` results (Figure 1(d)). Refer to Section 6.1 for help interpreting the data.

```

# of pixels: 254666 # of background: 209365 # of building: 45301
Quality: 0.363917
% of scene that is buildings: 0.177884
Correctly detected background pixels: 186208
Correctly detected building pixels: 24913
Building pixels missed: 20388
Background pixels incorrectly classified: 23157

% correctly classified pixels: 82.9011
% incorrectly classified pixels: 17.0989
% buildings detected: 54.9944
% buildings missed: 45.0056
% background detected: 88.9394
% background missed: 11.0606
% incorrect pixels that are false positives: 53.1795
building pixel branching factor: 0.929515

Individual building coverage percentages:
Complex 1: 0 (0/1877 pixels)
Complex 2: 67.8675 (1865/2748 pixels)
Complex 3: 16.4245 (667/4061 pixels)
Generic Peak 1: 87.4802 (1656/1893 pixels)
Long Peak 1: 49.4556 (863/1745 pixels)
Long Peak 2: 18.0089 (322/1788 pixels)
Long Peak 3: 31.4255 (593/1887 pixels)
Long Peak 4: 88.5623 (1657/1871 pixels)
Long Peak 5: 63.964 (1207/1887 pixels)

```

Long Peak 6: 60.2501 (1108/1839 pixels)  
 Long Peak 7: 0 (0/1833 pixels)  
 Long Peak 8: 45.494 (838/1842 pixels)  
 Long Peak 9: 0 (0/363 pixels)  
 Overhang Peak 1: 49.8343 (451/905 pixels)  
 Overhang Set 1: 76.3419 (384/503 pixels)  
 Overhang Set 10: 72.5898 (384/529 pixels)  
 Overhang Set 11: 69.6798 (370/531 pixels)  
 Overhang Set 12: 60.6061 (300/495 pixels)  
 Overhang Set 13: 99.8134 (535/536 pixels)  
 Overhang Set 14: 99.262 (538/542 pixels)  
 Overhang Set 2: 74.8016 (377/504 pixels)  
 Overhang Set 3: 75.6705 (395/522 pixels)  
 Overhang Set 4: 66.0448 (354/536 pixels)  
 Overhang Set 5: 72.0472 (366/508 pixels)  
 Overhang Set 6: 61.7143 (324/525 pixels)  
 Overhang Set 7: 99.2337 (518/522 pixels)  
 Overhang Set 8: 99.2593 (536/540 pixels)  
 Overhang Set 9: 72.3757 (393/543 pixels)  
 Peak 1: 0 (0/953 pixels)  
 Peak 2: 94.233 (817/867 pixels)  
 Peak 3: 73.7395 (702/952 pixels)  
 Peak 4: 46.4262 (708/1525 pixels)  
 Peak 5: 89.6023 (1577/1760 pixels)  
 Peak 6: 5.02513 (50/995 pixels)  
 Peak 7: 0 (0/0 pixels)  
 Rectilinear 1: 0 (0/0 pixels)  
 Short Peak 1: 97.619 (410/420 pixels)  
 Short Peak 2: 97.032 (425/438 pixels)  
 Short Peak 3: 87.5862 (381/435 pixels)  
 Short Peak 4: 72.5118 (306/422 pixels)  
 Short Peak 5: 97.561 (800/820 pixels)  
 Short Peak 6: 99.4929 (981/986 pixels)  
 Short Peak 7: 88.5111 (755/853 pixels)

## 6.4 Sample output from eval3d on VHBUILD results

The following is output produced by eval3d on VHBUILD results (Figure 1(d)). Refer to Section 6.2 for help interpreting the data.

True positive voxels: 169450  
 True negative voxels: 17559349  
 False positive voxels: 675546  
 False negative voxels: 348883  
 (tp / (tp + fn)) Detection pct: 32.6913  
 (fp / tp) Voxel branch factor: 3.9867  
 (fn / tp) Voxel miss factor: 2.05891  
 (tp / (tp + fp + fn)) Quality pct: 14.1932  
 ((fn + fp) / tp) : 6.04561  
 (tp / (tp + fp)) : 0.200533

Individual building coverage percentages:  
 Complex 1: 0 (0/46184 voxels, 0/5773 m<sup>3</sup>)  
 Complex 2: 68.4279 (16357/23904 voxels, 2044.62/2988 m<sup>3</sup>)  
 Complex 3: 14.2545 (4690/32902 voxels, 586.25/4112.75 m<sup>3</sup>)  
 Generic Peak 1: 74.2041 (13146/17716 voxels, 1643.25/2214.5 m<sup>3</sup>)  
 Long Peak 1: 31.8122 (7264/22834 voxels, 908/2854.25 m<sup>3</sup>)  
 Long Peak 2: 13.588 (3144/23138 voxels, 393/2892.25 m<sup>3</sup>)  
 Long Peak 3: 28.2094 (6322/22411 voxels, 790.25/2801.38 m<sup>3</sup>)  
 Long Peak 4: 66.1914 (16434/24828 voxels, 2054.25/3103.5 m<sup>3</sup>)  
 Long Peak 5: 47.1976 (12446/26370 voxels, 1555.75/3296.25 m<sup>3</sup>)  
 Long Peak 6: 9.46528 (2110/22292 voxels, 263.75/2786.5 m<sup>3</sup>)  
 Long Peak 7: 0 (0/23356 voxels, 0/2919.5 m<sup>3</sup>)

Long Peak 8: 0 (0/24091 voxels, 0/3011.38 m<sup>3</sup>)  
 Long Peak 9: 0 (0/25156 voxels, 0/3144.5 m<sup>3</sup>)  
 Overhang Peak 1: 39.8188 (4220/10598 voxels, 527.5/1324.75 m<sup>3</sup>)  
 Overhang Set 1: 45.7311 (1323/2893 voxels, 165.375/361.625 m<sup>3</sup>)  
 Overhang Set 10: 45.9807 (1430/3110 voxels, 178.75/388.75 m<sup>3</sup>)  
 Overhang Set 11: 43.4907 (1216/2796 voxels, 152/349.5 m<sup>3</sup>)  
 Overhang Set 12: 36.6787 (1016/2770 voxels, 127/346.25 m<sup>3</sup>)  
 Overhang Set 13: 97.0112 (3116/3212 voxels, 389.5/401.5 m<sup>3</sup>)  
 Overhang Set 14: 93.981 (2873/3057 voxels, 359.125/382.125 m<sup>3</sup>)  
 Overhang Set 2: 47.0074 (1343/2857 voxels, 167.875/357.125 m<sup>3</sup>)  
 Overhang Set 3: 43.5484 (1080/2480 voxels, 135/310 m<sup>3</sup>)  
 Overhang Set 4: 39.7758 (1171/2944 voxels, 146.375/368 m<sup>3</sup>)  
 Overhang Set 5: 45.1841 (1276/2824 voxels, 159.5/353 m<sup>3</sup>)  
 Overhang Set 6: 39.2009 (1089/2778 voxels, 136.125/347.25 m<sup>3</sup>)  
 Overhang Set 7: 97.2096 (2787/2867 voxels, 348.375/358.375 m<sup>3</sup>)  
 Overhang Set 8: 100 (2997/2997 voxels, 374.625/374.625 m<sup>3</sup>)  
 Overhang Set 9: 44.4642 (1498/3369 voxels, 187.25/421.125 m<sup>3</sup>)  
 Peak 1: 0 (0/7211 voxels, 0/901.375 m<sup>3</sup>)  
 Peak 2: 83.9588 (7898/9407 voxels, 987.25/1175.88 m<sup>3</sup>)  
 Peak 3: 46.6389 (3469/7438 voxels, 433.625/929.75 m<sup>3</sup>)  
 Peak 4: 23.0549 (5073/22004 voxels, 634.125/2750.5 m<sup>3</sup>)  
 Peak 5: 81.1118 (13351/16460 voxels, 1668.88/2057.5 m<sup>3</sup>)  
 Peak 6: 0 (0/7705 voxels, 0/963.125 m<sup>3</sup>)  
 Peak 7: 0 (0/16948 voxels, 0/2118.5 m<sup>3</sup>)  
 Rectilinear 1: 0 (0/10818 voxels, 0/1352.25 m<sup>3</sup>)  
 Short Peak 1: 87.6694 (2652/3025 voxels, 331.5/378.125 m<sup>3</sup>)  
 Short Peak 2: 84.4287 (2505/2967 voxels, 313.125/370.875 m<sup>3</sup>)  
 Short Peak 3: 96.4019 (2840/2946 voxels, 355/368.25 m<sup>3</sup>)  
 Short Peak 4: 42.0851 (1316/3127 voxels, 164.5/390.875 m<sup>3</sup>)  
 Short Peak 5: 91.6032 (5771/6300 voxels, 721.375/787.5 m<sup>3</sup>)  
 Short Peak 6: 94.9101 (8186/8625 voxels, 1023.25/1078.12 m<sup>3</sup>)  
 Short Peak 7: 91.2814 (6041/6618 voxels, 755.125/827.25 m<sup>3</sup>)

## 7 The CMU Site Exchange Format

The CMU Site Exchange Format is described in the following report:

Bulwinkle, G., Cochran, S., McGlone, J. C., McKeown, D., and Shufelt, J., "Robust Exchange of Cartographic Models for Buildings and Roads: The CMU MAPSLab Site Exchange Format," Technical Report CMU-CS-98-134, School of Computer Science, Carnegie Mellon University, December 1998.

This report is available via the web in PDF and compressed PostScript:

<http://www.cs.cmu.edu/afs/cs/usr/maps/www/rcvw/terrainweek98/cmutr-98-134.pdf>  
<http://www.cs.cmu.edu/afs/cs/usr/maps/www/rcvw/terrainweek98/cmutr-98-134.ps.gz>

## References

- [1] Y. Hsieh. Design and evaluation of a semi-automated site modeling system. In *Proceedings of the ARPA Image Understanding Workshop*, Palm Springs, California, February 1996. Advanced Research Projects Agency, Morgan Kaufmann Publishers, Inc.
- [2] Y. Hsieh. Sitecity: A semi-automated site modelling system. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 499–506, San Francisco, California, 18–20 June 1996.
- [3] R. B. Irvin and D. M. McKeown. Methods for exploiting the relationship between buildings and their shadows in aerial imagery. *IEEE Transactions on Systems, Man & Cybernetics*, 19(6):1564–1575, 1989.

Also available as Technical Report CMU-CS-88-200, Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213.

- [4] J. C. McGlone and J. A. Shufelt. Projective and object space geometry for monocular building extraction. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 54–61, Seattle, Washington, 19–23 June 1994.
- [5] D. M. McKeown. Toward automatic cartographic feature extraction. In L. F. Pau, editor, *Mapping and Spatial Modelling for Navigation*, volume F 65 of *NATO ASI Series*, pages 149–180. Springer-Verlag, Berlin Heidelberg, 1990.
- [6] J. Shufelt. Exploiting photogrammetric methods for building extraction in aerial images. In *International Archives of Photogrammetry and Remote Sensing*, volume XXXI, B6, S, pages 74–79, 1996.
- [7] J. A. Shufelt and D. M. McKeown. Fusion of monocular cues to detect man-made structures in aerial imagery. *Computer Vision, Graphics, and Image Processing*, 57(3):307–330, May 1993.