

IDENTIFYING BUILDING CHANGE USING HIGH RESOLUTION POINT CLOUDS – AN OBJECT-BASED APPROACH

Steve du Plessis
Global Product Line Executive Intergraph Corporation | ERDAS,
5051 Peachtree Corners Circle, Suite 100, Norcross, GA 30092 USA
Steve.duPlessis@intergraph.com

KEY WORDS: LiDAR, point cloud, change detection, stereo, image analysis, remote sensing, LAS

ABSTRACT:

High resolution point clouds provide excellent data sources for examining change over time in above-ground features such as buildings and trees. Of particular interest is the identification of illegal construction activity or damage incurred during earthquakes and other disasters. By using multi-date point cloud layers, these types of change can be efficiently identified and mapped. Such analysis is generally not as simple as differencing imagery from the two dates. Variations between the images can be caused by slight geometric mismatches between images from different acquisition dates, errors in the data returns, or natural differences caused by vegetation growth or wind direction. The factors can contribute to the detection of large amounts of inconsequential change throughout the area of interest, resulting in too many false positives for the analysis to be of any practical use. However, by conducting object-based analysis of the data – analysing meaningful objects rather than working point by point – software algorithms can be used to rapidly and accurately detect and map only the changes of interest to the customer.

1. MANUSCRIPT

1.1 Introduction

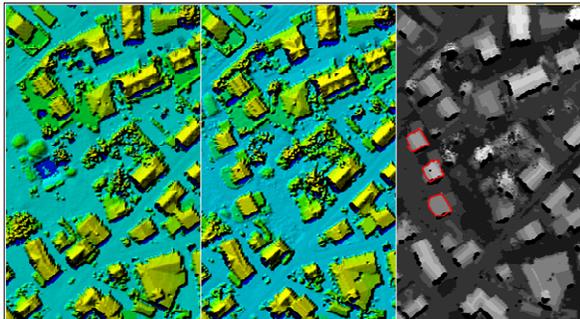


Figure 1. (Left, Center) Hillshaded point clouds from November 2008 and September 2009, (Right) Point cloud indicating new buildings.

High resolution point clouds provide excellent data sources for examining change over time in above-ground features such as buildings and trees. For example, an analyst can use multi-date point cloud layers to efficiently identify and map changes originating from illegal construction activity or damage incurred during earthquakes and other disasters.

Such analysis is generally not as simple as differencing imagery from the two dates. Variations between the images can be caused by slight geometric mismatches between images from different acquisition dates, errors in the data returns, or natural differences caused by vegetation growth or wind direction. The factors can contribute to the detection of large amounts of inconsequential change throughout the area of interest, resulting in too many false positives for the analysis to be of any practical use.

However, by conducting object-based analysis of the data – analysing meaningful objects rather than working point by point

– software algorithms can be used to rapidly and accurately detect and map only the changes of interest to the customer.

Data

The following workflow uses LIDAR data from two dates (November 2008 and September 2009). The data is in the LAS file format.

A very similar workflow could also be conducted using point cloud data from software such as LPS eATE, which auto-correlates stereo imagery to generate point clouds.

1.2 Problem – Mapping New Building Construction

The primary objective of this workflow is to detect and map new building construction that occurred between the two dates. To illustrate that the same process could be used for disaster response mapping, we also performed a secondary analysis to detect and map buildings that were removed between the two dates.

The analysis process consists of three major steps.

First, we processed the two point cloud datasets using a variation of the traditional approach to change detection in which the “before” data is subtracted from the “after” data. This yields a difference file that represents the height difference between the two dates, with large positive height change representing new building construction and negative height change representing building loss.

Second, we converted the “after” point cloud into objects using an image segmentation routine. We chose to segment the “after” data because we are looking for new building construction and those buildings should be in the “after” data but may not be in the “before” data. To detect building damage and destruction as part of a disaster assessment, you would perform the reverse-segmenting the “before” data in which the buildings are most clearly represented as objects.

Third, we combined the objects with the height difference data so that each object reflects the average change in height for that object. Then, we analysed the objects for significant change to remove noise and map the construction activity between the two dates.

The next sections outline how an analyst can conduct this process using ERDAS IMAGINE.

1.2 Point Cloud Differencing

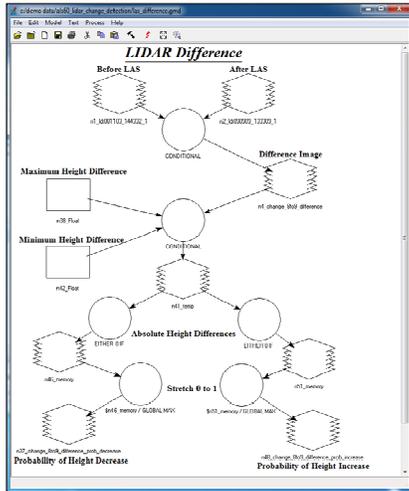


Figure 2. The graphical model used to produce Probability of Height Increase and Probability of Height Decrease images.

Input Point Cloud Data

The first two objects at the top of the model represent the input point cloud datasets. To change the dataset simply double-click the relevant object and select your input data in the resulting File Chooser dialog.

ERDAS IMAGINE can directly read the LAS file format for LiDAR data as well as several other formats used for point cloud information. Alternatively, you can first use one of the import or surface interpolation routines (in the Terrain Preparation dialog) to convert your source data. This is useful if you want greater control over parameters such as the spacing of resulting raster surfaces or the use of different LIDAR returns.

Difference Calculation

The first function object in the model, fed by the two input point clouds, performs the primary differencing by subtracting the “before” values from the “after” values to generate the difference in height.

However, point clouds generally have areas of Null or NoData in them, for example where the LIDAR failed to return values or in areas outside of the collection footprint. These are often represented by using a NoData value in the file itself, such as -32767. So as not to bias the height differencing at these values the differencing does not apply the calculation at these locations (instead passing through the value of -32767).

Difference Image

The model does create a file of the raw differencing results even though they are not directly utilised, it is generated simply as an aid for Quality Assurance if needed.

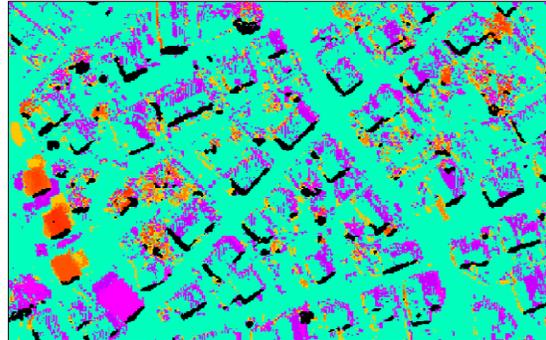


Figure 3: A portion of the difference image.

To the lower left you will see three orange rectangular areas of height increase which are new buildings. There is also a larger violet rectangular area showing a building that has been demolished.

While these are clear to the human eye, you can also see the “noise” which generally confuses any attempts to map just features of interest, such as new buildings. There are strong linear effects caused by slight mismatch of building edges between the two dates, there are clumps representing strong tree growth (top right) as well as other scattered effects representing bad points in the LiDAR data, minor vegetation height changes, variations caused by vegetation sway, leaf-on / leaf-off variability, etc. All these would result in “false positives” using more traditional approaches – areas identified as change which, while they may represent actual change in the data, do not represent the types of change we wish to map.

Minimum and Maximum Height Difference Constraints

One way to reduce false positives is to define the range of valid heights that would occur with the type of change being mapped. For example, when looking at new building construction work we can assume that the minimum valid height difference for a new building or new floor on an existing building is 1.5 meters. Height differences of more than 40 meters are more likely the result of bad height returns, or represent significant new buildings which probably do not need to be automatically identified.

The more you can constrain the valid height range you wish to automatically map, the fewer errors of commission (false positives) you introduce into the final results.

Consequently, the graphical model includes two input scalar objects on the left side representing the maximum and minimum height differences that are allowed. These are set at 40 meters and 1.5 meters, respectively, but can be easily changed.

These scalars feed into a further function which takes in the difference image and sets any values outside the valid height range (including NoData values) to zero (since those locations represent areas of zero probability of being a valid new building construction).

Absolute Height Differences

As mentioned above, the primary objective is to identify new building construction. For this, we are really only interested in the height difference values which are greater than zero. Consequently, the right-hand branch at the bottom of the model takes only the positive height difference values.

The branch on the left takes the values less than zero and calculates their absolute value. This is done so that when looking for the probability of buildings being demolished we still have positive values (i.e. a positive value of height decrease).

Stretch

So far, the values being calculated represent the actual difference in height. For the next stage of the analysis, we are only interested in the probability of change, so the greater the height difference, the higher the probability that change has truly occurred. The final calculation stretches the differences in heights so that they range from 0 to 1.

Probability of Height Difference

The two output raster objects represent the probability of height decrease (lower left) and the probability of height increase (lower right).

1.3 Object-based Image Analysis using IMAGINE Objective

IMAGINE Objective provides tools for feature extraction and update and change detection, enabling geospatial data layers to be created and maintained using remotely sensed imagery. IMAGINE Objective combines inferential learning with expert knowledge in a true object-oriented feature extraction environment.

Segmentation

The 16-bit “after” point cloud will be segmented in the Raster Object Creator node to create objects. The Raster Object Creator has been set as the start point of the analysis thereby skipping the Raster Pixel Process (RPP). Normally, a per-object probability would initially be calculated by the RPP. However, here we already have our “probability” values from the graphical model.

So, instead of running a RPP, we have selected the Segmentation node, changed to the I/O (Input / Output) tab and specified the input raster layer as being the Probability of Height Increase layer. This is not the layer which will be segmented (the 16-bit “after” point cloud will be segmented to produce spatial objects) – this layer is instead used to determine the initial probability value that will be associated with each object.

Remember that the “after” point cloud was stretched to a full 16-bit data range. So, the properties for the segmentation process need to be correspondingly large / loose. The Minimum Value Difference was set to 100 and, in the Advanced Settings, the Edge Detection Threshold was set to 50.

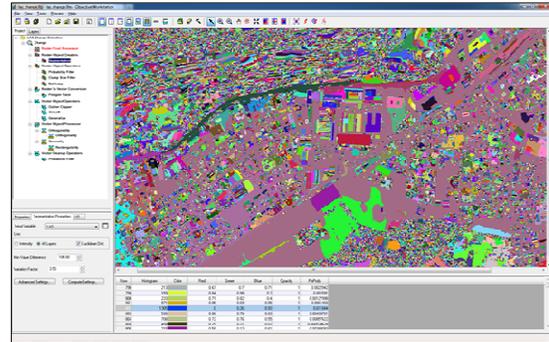


Figure 4. Segmentation results. Note that each segment is attributed with a pixel probability derived from the Probability of Height Increase layer.

Raster-Object Operators

An initial **Probability Filter** is applied to the segments to remove any segment with a probability of representing significant height increase less than **0.1**:

A Clump Size filter is then applied to remove any groups of segments which are smaller than 50 square meters, the smallest building that would be of interest:

Remaining segments are then reclustered if contiguous.

Raster to Vector Conversion

The IMAGINE Objective software then converts the raster clumps to vector polygons so that they can be further analysed using object-based cue metrics such as area:

Vector-Object Operators

To assist in analysing the shapes of the remaining objects they are first processed using an outlier clipper (set to a mild 0.1) to remove any erroneous spikes:

Polygons are then smoothed to remove the stair-step effects of converting pixel-based clumps to vector polygons using a mild **0.1** smoothing factor:

Finally, **Generalize** is applied at 0.4 meter thinning tolerance to ameliorate the stair-stepping effects:

Vector-Object Processor

At this stage we have numerous polygons remaining which represent objects with a high degree of a positive height change. Since we are interested in identifying new building construction these objects will then be measured for characteristics which would be found in buildings and which will therefore help distinguish new buildings from other types of height difference.

The characteristics used include orthogonality and rectangularity. We could also use other cues, such as area, since buildings tend to have constrained area footprints. However, it is best to use only orthogonality and rectangularity because the valid ranges for the cue do not need to be specified by the user.

Since we have already thinned the objects to represent only those with relevant height changes, the per-pixel probability measure is no longer important in comparison to the object shape measures. So, we also have reduced the Pixel Probability Weight from the default 50 percent to only 1 percent. This way the final probabilities are mainly derived from the shape cues, not from the height differences.

Vector Clean-up Operators

After the polygon objects have been updated with geometry-based object cues, the final step is to filter out all low probability objects. Objects with low final probabilities are more likely to be trees than they are buildings because of the shape measures we applied.

A final probability filter is applied at 0.5.

1.4 Quality Assurance

Running the IMAGINE Objective feature model identified 19 objects representing new building construction between the two dates of LiDAR collection. To check the validity of these results, the following procedure was followed.

Errors of Commission

Using ERDAS IMAGINE 2011, two 2D views were open alongside each other. The “after” LiDAR data was loaded into the left view and the “before” into the right view. The IMAGINE Objective results were overlain into the left view and their symbology changed to an unfilled, outlined polygon. The two views were linked and scales equalized.

We started the attribute table for the shapefile and selected all records.

In the Table tab, the **Zoom to Item** controls were used to drive to each polygon one by one. By visually comparing the height information at the same location, it was easy to determine whether each polygon was correctly identified as new building construction or if it was a false positive.

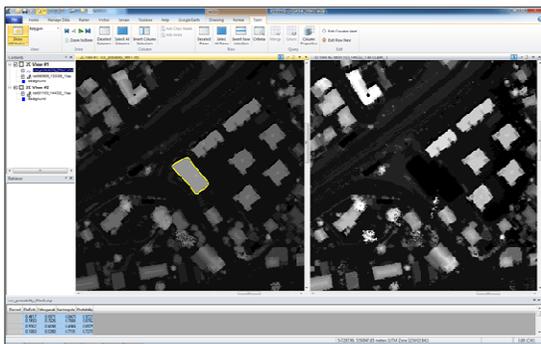


Figure 5. This polygon is a correct detection.

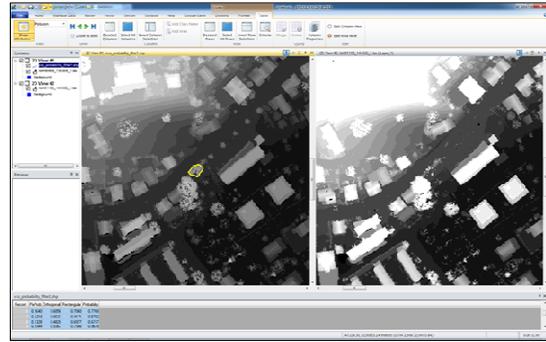


Figure 6. This polygon appears to be a false detection (and would be easier to confirm if colour or false-colour infrared imagery were available as a simultaneous or near-simultaneous capture).

False detections can easily be deleted from the vector layer as they are reviewed.

Errors of Omission

To find new building construction missed by the IMAGINE Objective change detection routine, both LiDAR point clouds were loaded into a single 2D View with the results shapefile overlain. The Swipe (or Blend) tools were then used to peel away the “after” data and visually compare with the “before.” This enables the human eye to detect other locations of height change which might be buildings and which can be investigated more closely.

Results

Of the 19 detected objects, 5 appear to be false positive (errors of commission) resulting from locations of height change which are not actually new building construction. Only two false negatives (errors of omission) were identified in the area.

1.5 Discussion

Spectral Information

Analysing the intermediary results of the IMAGINE Objective model (a capability which is a significant advantage of IMAGINE Objective), the two errors of omission are objects which came very close to meeting the probability cut-off for inclusion into the identified set.

On the other hand, the errors of commission appear to generally represent specific trees (with a large height difference, whether for new planting, leaf on/off conditions, or other reasons) which resulted in objects with a high degree of similarity in shape to rectangular buildings due to either having high rectangularity measures or high orthogonality measures.

If the analysis included a source of spectral information such as natural colour or preferably 3- or 4-band data with red and near-infrared wavelengths, it could discriminate between vegetation objects and buildings. In this manner, the majority of false positives would be rejected and it would be far more likely that the relative probability of the few false negatives would increase thereby including them into the correct detections without increasing the number of false positives unduly.

Inclusion of multispectral data would also help reduce the dependency on the presence of straight edges to identify an object as a building (rather than a tree). This way, we would also be able to detect other shapes of new buildings such as the oval building in the northeast side of this test dataset.

1.6 Conclusion

IMAGINE Objective appears to provide a robust capability for detecting specific types of change between point cloud data from two different dates, even in the absence of other ancillary information such as co-registered and contemporaneous multispectral data or even pre-classification of the height points.

Additionally, the approach outlined requires no training of the classifier. The user does not need to specify locations to serve as examples of the change of interest. Traditionally, having to provide training samples slows down the analysis process and results in approaches which can't easily be transported between data sets. Instead the user might have to make minor adjustments to some of the parameters in the IMAGINE Objective project to better reflect the local conditions, but that is all.