Lidar-Derived Building Footprints vs. Aerial Photographs

A Kopale

University of Iowa

Lidar: Principles and Applications

Feb 2019

Contents

I. Introduction

a. Since the late 1980s and early 1990s, lidar technology has become one of the primary methods for developing accurate, detailed spatial data (Beraldin et al., 2010; Ackerman, 1994).
    i. Lidar, or light radar, relies on a laser range finder, scanning optics, some type of platform, an inertial measurement unit (IMU), and a scanning device to take time-of-flight and reflectance measurements, which are point data that can then processed to create 3-dimensional models (Linderman, 2019).
    ii. The point data obtained through lidar scans contain x, y, and z data and are called returns.
    iii. The collection of points from a lidar scan are called point clouds and can be used to create things like detailed digital elevation models.
b. The commercial development and release of accurate lidar systems have allowed various sectors to utilize this state-of-the-art technology.
    i. Environmental scientists and geographers can use lidar for hydraulic modelling.
    ii. Archeologists can use lidar to search for and model structures built by ancient civilizations.
    iii. Lidar can even be used by programmers and developers to create accurate representations of existing terrain for virtual reality videogames or automated vehicles.
c. This project looks at building footprints created from lidar data and compares the accuracy of the lidar-derived footprints to an aerial photograph of the same study area.
    i. While lidar technology has improved significantly over the last few decades, it is still just a model of the real world, which means there is plenty of room for error.
        1. This analysis will determine what errors, if any, are present in the lidar-derived building footprint model.
        2. Once we know the errors, we can begin to fix them, in this case by using an aerial photograph as a guide.
        3. Fixing errors and inconsistencies in data is one of the most important steps in any project, because a project with bad data almost always leads to a bad result.

II. Data

a. The data used to complete this project include an aerial photograph, a lidar point cloud file, and a shapefile of building footprints.
    i. The lidar is from GeoTree, a University of Northern Iowa lidar distribution project that provides USDA NRCS lidar data for the state of Iowa to the public (2019).
        1. The specific grid of lidar data used for this is 06224612.
        2. The data were downloaded as an .las file so that all returns measured could be viewed and processed, rather than GeoTree’s other option – a pre-processed ascii file that only provides bare earth estimates.

ii. The aerial photograph is from the Iowa DNR GIS data website (2017).
    1. The aerial photograph is a digital ortho mosaic with a target resolution of 1 meter and was published on October 22, 2017.
    2. The photograph's geographic coordinate system in NAD 83, projected to UTM Zone 15 North.
iii. The building footprints shapefile was provided by the instructor, Marc Linderman.
    1. This shapefile is vector polygon data of building footprints, generated from a tile of lidar data.
iv. All analysis and processing was done in ArcMap version 10.6, a GIS platform published by ESRI.

III. Methodology

a. The first thing I did was visually explore the lidar data in ArcMap.
    i. The data can be explored using las tools, an ESRI ArcMap plugin that allows the data to be shown by classification type or return.
        1. Classification codes classifies each return into one of 19 pre-defined categories, which have been defined by the American Society for Photogrammetry and Remote Sensing (ESRI 2019).
            a. The classification codes include things like building, water, unassigned, and high vegetation.
        2. Returns are simply what we call the pulse that returns to the sensor, but some pulses sent out by the laser can record multiple returns.
            a. These might include first return, second return, and last return, but usually does not exceed five returns in discrete scans.
            b. The returns also may have an intensity associated with it, which is the strength of the return relative to the generated pulse (ESRI).
    ii. The data can also be visualized by surfaces, which include TIN, slope, and aspect.
        1. When the data is visualized as "TIN," or triangular irregular networks, it uses triangulation to create nonoverlapping triangles from the point cloud.
            a. This can be useful in visualizing and determining things like ridgelines of surface flows.
        2. Slope is a lidar filter that shows how sharp/flat a change in elevation is, while aspect shows the direction the slope is facing.
        3. Contours is an las tool that visualizes the elevation of the surface using contour lines, which can be used to create an isoline elevation map for a hiking trail, for example.
    iii. Another useful way of exploring lidar data is by using the profile view.
        1. This shows the data as if it were sliced from the sky to below the surface of the earth.

        2. This is particularly useful in understanding how topography affects the lidar data in a way that isn't always apparent from an overhead view.
            a. For example, where there are buildings, there are no surface level returns.
            b. Where there are trees, there is some of what appears to be surface level returns, presumably the places where a pulse was able to find a gap in the canopy.

b. After visually exploring the data and learning some of the functionality of las tools, I opened the aerial photograph and overlaid the lidar-derived building footprints shapefile.
    i. Once these were in ArcMap and properly projected to UTM Zone 15 North, I created a new polygon shapefile.
        1. I edited the newly created polygon shapefile to the size I wanted, then used the Clip tool in the geoprocessing toolbox to clip the building footprints to the extent of the polygon.

c. Next, I explored how the lidar derived building footprints compared to the aerial photograph of the study area.
    i. I zoomed in and out of certain buildings to look at how the building footprint matched the image.
    ii. I turned off and on the lidar layer and increased its transparency to see how things that from the aerial photograph were clearly not buildings were considered buildings in the building footprint shapefile.

d. Finally, I adjusted the building footprint outlines to match the buildings in the aerial photograph as closely as possible.
    i. I used the editing tool bar for this, relying significantly on the editing vertices tool to make minor adjustments, the delete tool to get rid of polygons that were not buildings at all, and the create new feature tool to create footprints, when necessary.

IV. Results

a. The original building footprints shapefile, which was derived from lidar data, did not match up perfectly to the actual building footprints shown in the aerial image.
    i. Some areas matched up well, while others were significantly off.
        1. The full study area extent showing building footprints on top of the lidar layer is shown in Figure 1 in the Appendix.
        2. Figure 2 and Figure 3 show the unedited and edited building footprints overlaid on the aerial photograph.
        3. Small portions of the study area that were particularly problematic are shown in Figure 4 in the Appendix as before and after images.
    ii. It seems like areas that had trees around the same height as the buildings were often off, as the building footprints extended past the structures to include returns from the canopy.
        1. This can be seen very clearly in Figure 4A, where surrounding trees are considered part of the building.
    iii. Another issue that was common was footprints that were slightly off in terms of rotation.

        1. This can be seen in Figure 4B.

V. Discussion

- a. One thing that may have caused inaccuracies in the lidar-derived building footprints is the nature of the lidar data itself.
    - i. Because of the way lidar data is acquired, there is almost never a perfectly uniform distribution of points across a surface area.
        1. Certain areas are scanned twice (overlap) while other areas are scanned only once.
        2. The angle of the laser as well as the angle of the returning pulses relative to the sensor can also lead to variations in point density across an area.
        3. These variations in point density lead to variations in the quality and accuracy of the data
- b. Topography also has a big influence on the accuracy of the lidar data.
    - i. If there is something completely obstructing a roof, for example and overhanging tree with a very dense canopy, returns will not be recorded for the roof.
        1. In a true-color aerial photograph, we can see that it is clearly a tree that is blocking the roof, so we can use our common knowledge of geometry and what homes look like in a particular area to adjust the footprint to better match what we think the building looks like under the tree.
- c. Some other potential influences include errors on the human side.
    - i. Maybe the GPS unit was attached to a block a few inches away from the sensor and not the sensor itself, and this distance was not corrected for.
    - ii. Maybe the pilot was a little shaky that day.
    - iii. Maybe the atmospheric corrections for the lidar data were not spot-on.
        1. All three of these are reasonable explanations for error.
- d. What this project showed is that geographic data, whether it be lidar, aerial imagery, or georeferenced buildings, are all just projections of the real world…even the really high-tech, dense geographic data.
    - i. How the data are collected, conditions when they are collected, the tech used to collect, and so on all affect the data, in turn affecting the quality and accuracy of the building footprint, and the overall project.
    - ii. This project proved that correcting geospatial data is a necessary step in ensuring the accuracy and integrity of the data, and in turn improving the accuracy and integrity of future analysis using that data.

VI. Works Cited

Ackerman, F., 1994. “On the status and accuracy performances of GPS photogrammetry.” *Proceedings of Mapping and Remote Sensing Tools for the 21st Century*, 26-29 August, Washington DC, 80-90.

GeoTree. The Iowa Lidar Project. Iowa Lidar Consortium, Jan 1 2007-Feb 4 2019. GeoInformatics Training Research Education and Extension. January 28 2019. http://www.geotree.uni.edu/lidar/

USDA NRCS FSA. 2017 NAIP Aerial Photography of Iowa. Iowa State University. Iowa State University GIS Facility, 2018-03-06.

Beraldin, J., F. Blais, and U. Lohr. “Chapter 1: Laser Scanning Technology.” *Airborne and Terrestrial Laser Scanning*, edited by G. Vosselman and H. Maas, Whittles Publishing, 2010, pp. 1-39.

Linderman, Marc. “Week 2-2,” *Lidar: Principles and Applications*. January 2019.

VII. Appendix

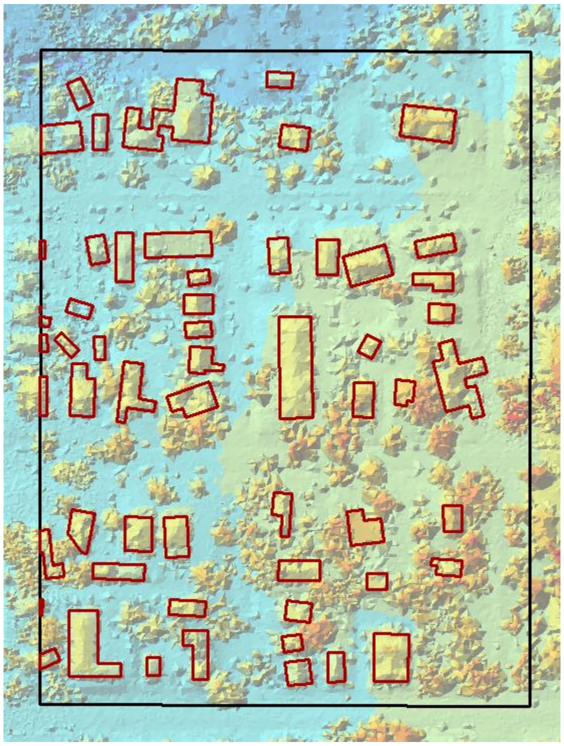

Figure 1 shows the original, unedited building footprints on top of the lidar layer

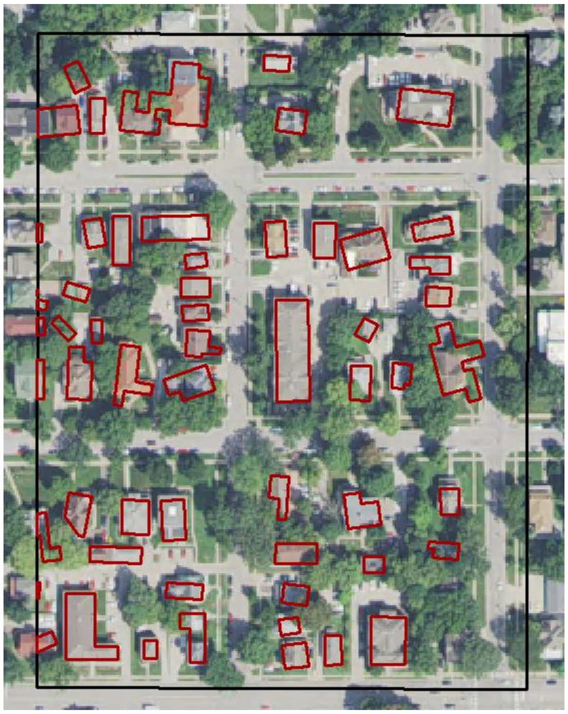

Figure 2 shows the unedited building footprints on top of the aerial photograph

Figure 3 shows the corrected building footprints on top of the aerial photograph

Figure 4A shows how trees included in the lidar-derived building footprints were edited using the aerial photograph to create a more accurate building footprint.

Figure 4B shows how building footprints that were incoorectly sized or rotated were fixed to create more accurate footprints.

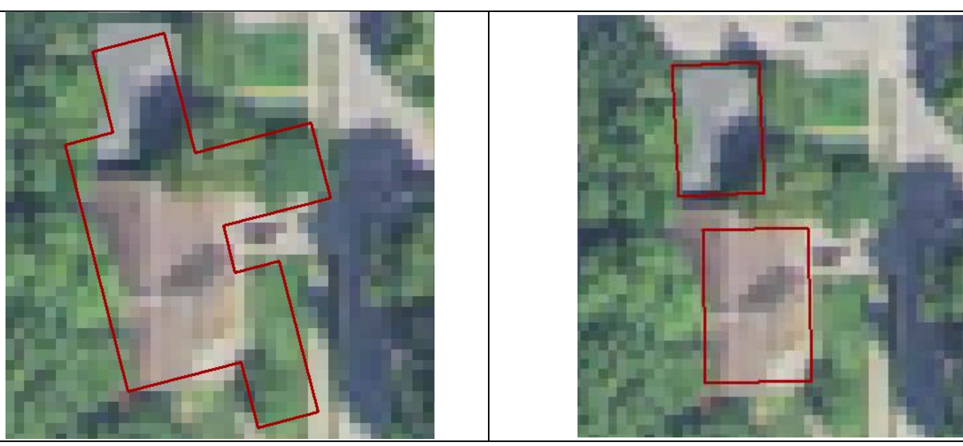

Figure 4C shows another problem area in the original footprints shapefile, likely the result of the trees being a similar distance from the senor as the homes, and how it was fixed.