[DC-01-025] Changes in Geospatial Data Capture Over Time: Part 1, Technological Developments

Abstract

Geographic Information Systems (GIS) are fueled by geospatial data. This comprehensive article reviews the evolution of procedures and technologies used to create the data that fostered the explosion of GIS applications. It discusses the need to geographically reference different types of information to establish an integrated computing environment that can address a wide range of questions. This includes the conversion of existing maps and aerial photos into georeferenced digital data. It covers the advancements in manual digitizing procedures and direct digital data capture. This includes the evolution of software tools used to build accurate data bases. It also discusses the role of satellite based multispectral scanners for Earth observation and how LiDAR has changed the way that we measure and represent the terrain and structures. Other sections deal with building GIS data directly from street addresses and the construction of parcels to support land record systems. It highlights the way Global Positioning Systems (GPS) technology coupled with wireless networks and cloud-based applications have spatially empowered millions of users. This combination of technology has dramatically affected the way individuals search and navigate in their daily lives while enabling citizen scientists to be active participants in the capture of spatial data. For further information on changes to data capture, see Part 2: Implications and Case Studies.

Keywords: accuracy, aerial imagery and photogrammetry, aerial photography, digitizing, raster, topographic quadrangles, USGS, vector

Author & citation

Cowen, D. J. and Anderson, K. E. (2021). Changes in Geospatial Data Capture Over Time: Part 1, Technological Developments. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2021 Edition), John P. Wilson (ed.).
DOI: 10.22224/gistbok/2021.1.4.

Explanation

1. Overview
2. Data Capture Fundamentals
3. Data Capture Tools
4. Critical Technological Advancements
5. Conclusion

1. Overview

Geospatial data capture depends on the ability to attach accurate real world coordinates to a feature. Creating today’s spatially enabled environment was only possible through a remarkable series of technological advancements and some enlightened public policies. As was described in 2009 in the Changing Geospatial Landscape report by the National

Geospatial Advisory Committee:

Nearly all the data, technology, and applications we see today can be traced to innovative policies and government practices of the past. As such we require similar innovative policies now to keep pace with this remarkable sea change. Government-based geographic information providers can no longer think of themselves as a player outside of or immune from the community of private sector, state, local or even public stakeholders. (National Geospatial Advisory Committee 2009)

Clearly, the world of geospatial data capture was permanently altered when President Clinton opened the non-degraded GPS constellation for civilian use. Now, even an inexpensive smartphone knows where it is and will geotag a photo with coordinates that are within a few meters of the actual position. In fact, for a huge list of applications, the modern smartphone is a fairly accurate surveying device that can accurately place a user in the context of a detailed map or air photo. GPS technology also powers real-time passive and active sensor systems that inventory and monitor a limitless variety of terrestrial, atmospheric, oceanic, seismic, and hydrological conditions. Fleets of commercial and governmental satellites now provide daily high-resolution coverage of the world capable of finding missile sites or extracting the footprint of a building. The busiest airport in the world has created a complete unified three dimensional interior and exterior geospatial model to support scores of decisions. In other words, the Internet of things has an unlimited geospatial footprint. It is important to place the evolution of geospatial data capture in the wider context of GIS technology and use. As Goodchild (2011) suggested,

the entire Internet is quickly becoming one vast GIS. Over the past two decades, however, the widespread availability of GPS and mapping software has changed the balance in this equation, making it possible to create maps of virtually anything for almost nothing…. In the future, GIS will involve much more real-time situation monitoring and assessment and will need new kinds of tools that treat information as continually changing. (Goodchild 2011).

2. Data Capture Fundamentals

Data for input for GIS applications comes from many sources. The most common format is raster or vector representation of a feature on the Earth’s surface with associated geographic coordinates that are be linked directly or defined relative to fixed positions. A raster is simply a geographically registered matrix of values with a system of registration to real world coordinates. Vector data maintain the fidelity of geographical features as points, lines, or polygons. Historically, the need to integrate multiple data themes resulted in a competition between the two data models. The raster camp generated a matrix of numbers assigned to cells of a fixed size. It is a perfect structure for data acquired from sensors and the representation of continuous surfaces such as elevation. However, considerable information is lost through the generalization process and grid cells are not suited for representation of linear features. Over the past thirty years we have witnessed a coalescence of the raster and vector camps. The choice of data structure and analytical tools is governed by the requirements of the application (Figure 1).

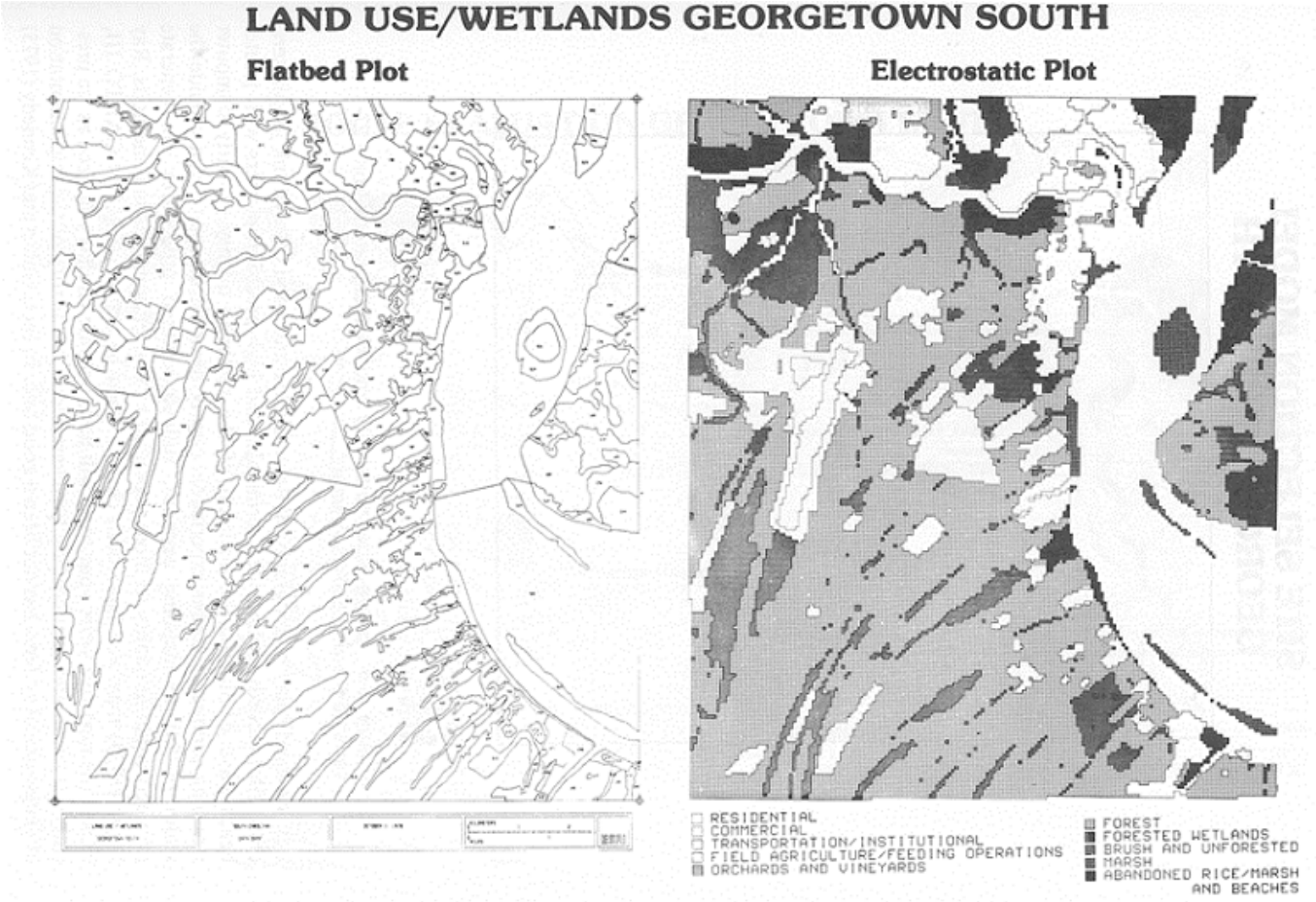


Figure 1. 1978 photo interpreted land use polygons output as vectors on pen plotter and raster on electrostatic plotter. Source: authors.

In the broadest sense geospatial data is either captured directly by an instrument capable of determining its position or indirectly through a manual or automatic procedure. A surveyor's recording of a benchmark's latitude and longitude is an example of GIS ready data. This also means that a smart phone with GPS in the hands of a citizen is also directly capturing GIS ready data. A huge range of devices such as a digital cameras are passive sensors that directly capture raster geospatial data. Other active sensors such as LiDAR and Radar also acquire huge clouds of points that measure a wide range of characteristics.

Indirect geospatial data capture requires that a manual or automated procedure be used to create GIS compatible data. The field of photogrammetry has perfected ways to calculate three dimensional positions for features on air photos. Contour lines and digital elevation models are interpolated from surveyed measurements. Digital orthophotos produced with softcopy (digital) photogrammetry enable a user to capture GIS ready features through heads-up digitizing Alternatively, using georeferencing algorithms traditional analog maps and photos can be adjusted to sources with known coordinates. In the early days of digital data creation, the capture of vector features from source materials required specialized digitizing tables that functioned like a large piece of digital graph paper (Figure 2). An operator mounted an existing map or photo on the digitizing table and used a stylus to trace points, lines, or polygons. Such devices were limited to large organizations or companies that provided commercial digitizing services. As GIS applications expanded and

demand for digital data increased, these tables also became a common fixture in research labs and governmental organizations. Initially, operators traced features without any preview or editing capability. Over time, special workstations enabled the operator to interactively display and edit features on monochrome storage tubes. In today's modern GIS environment, data capture is conducted with desktop tools that support on screen or heads-up digitizing with a mouse on standard color computer displays. Sophisticated editing tools support automated edge following, snapping, polygon completion and enforcement of topological rules.

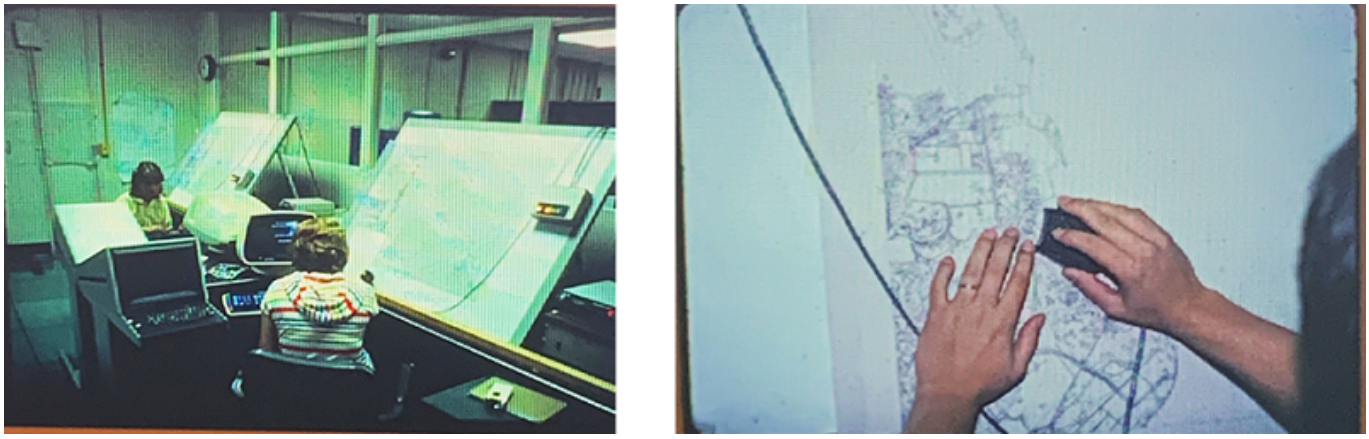


Figure 2. 1987 Manual digitizing process at the National Wetlands Research Center, USFWS Slidell LA, Image Source: James D. Scurry, used with permission.

The more than 55,000 scanned USGS topographic quadrangles) are an example of a geographically rectified source material. Digitizing can be performed by mounting the material on a digitizing table or working with a scanned version. This procedure usually requires air photo interpretation to identify and select features. In effect, scanned versions of air photos have the same properties as any raster data acquired by a passive sensor such as a multispectral Earth observation satellite. That means the values of the pixels can be classified into categories such as land cover. These raster data sets can also be used to extract features such as roads and rooftops. Today, they can be analyzed with sophisticated machine learning tools to perform automated feature extraction. GIS software is ready to incorporate them as either raster or polygons.

Other forms of indirect capture are also critical to creating geospatial data. In many cases new features are linked to existing rectified features. This process can be performed interactively by snapping to relevant features and adding new attributes. It can also be performed through a spatial search that transfers attribute information from the closest point or within a buffer. Reverse geocoding that finds all the people who need to be notified in an emergency is an example of the latter. The surveyor's records of metes and bounds can also be used to create polygons of land ownership and generate an authoritative digital tax map. Using the site address these parcels also provide a basis for automated address geocoding that links coordinate points to an address. In a similar manner, the US Census Bureau interpolates the coordinates of a street address from the address ranges in its TIGER line files. NOAA's Coastal Mapping Program demonstrates how data from active and

passive sensors can be fused to create a comprehensive representation of the coast. For example, a handheld Analytical Spectral Device that collects eighty spectral signatures is used in the field to capture coastal features that are integrated with LiDAR and multispectral imagery to create a comprehensive depiction of the coast. This integrated multi-layer data is used to monitor conditions and can be updated on demand.

3. Data Capture Tools, Platforms, and Approaches

Over the past half a century the process of data capture has changed dramatically, both through introduction of novel technologies and significant improvements in existing ones. The evolution of GIS as essential technology benefited directly from the general advancements in the power of processors, the magnitude of storage systems and the capabilities of IT network infrastructure. In terms of software tools cartography, surveying, geodetic surveying, photogrammetry, computer aided design, remote sensing, and machine learning can all be accommodated under the umbrella of GIS. That means that there are pathways between these applications that enable a user to capture geospatial data from many sources, integrate them with other themes, convert them between data structures, use analytical procedures to create new data from existing features such as interpolation or spatial search. In addition, numerous specific advancements have directly impacted the way geospatial data are now created, including:

- GNSS (Global Navigation Satellite System). Examples include Europe's Galileo, the USA's NAVSTAR, Russia's GLONASS, and China's BeiDou. These have enabled even smart phones to independently record coordinates, and has revolutionized surveying and enabled crowd sourced data capture.
- Digital photography (passive sensors). These directly generate raster data by assigning reflective and emitted light values to pixels. They can be mounted on a range of platforms from small UAVs to Earth Observation satellites.
- LiDAR & other active sensors. These create huge point clouds of values returned from emitted pulses and have revolutionized the way that elevation data is created and manipulated.
- Digital orthophotographs produced from softcopy photogrammetry. The availability of such imagery means it is the preferred photo source for heads-up manual digitizing. They have eliminated the need for digitizing tablets.
- Oblique air photos. These multiple directional images provide 3D perspectives and support and enhance data extraction.
- Classification algorithms for raster data. These include supervised and unsupervised procedures for generating land cover and other characteristics. They provide a way to update maps and automatically create new data themes.
- Automated feature recognition and extraction. Image processing algorithms that identify and separate edges of features and extract polygons.
- Address geocoding. These procedures assign coordinates to a textual string such as a street address, achieved through a direct database join or by interpolation from address ranges. Improvements in accuracy and completeness mean this approach has become an essential capability for data capture.
- Creation of accurate global datums and software to handle geographic and projected coordinates. Improved measurements and integration of these into GIS has facilitated the development of data themes.

- Spatial metadata standards. Advances in the development and deployment of these had improved interoperability, development of new applications, and assessment of suitability of data use.
- City Geographic Markup language (CityGML). This open data model provides a mature semantic information model for the representation of 3D urban objects at different levels of detail (LOD).
- GeoTif and Geopdf file structures. These common formats facilitate the transfer of traditional digital data. They have provided GIS and non-GIS users tools to calculate coordinates, separate themes, and make measurements.
- Geographic Search Engines. These web-based tools provide easier search, discovery, access, and dissemination of geospatial data
- Robust networks. The existence of robust networks has facilitated virtual connections rather than relying on more cumbersome file transfer methods.
- Web 2.0. This has enabled sharing and collaboration for extensive applications, including geospatial ones.

4. Critical Technological Advancements

Certain advances in technologies have enabled GIS in particular to flourish. This section elaborates on a few of these innovations.

4.1 Direct Capture in the Field / Surveying

For centuries, surveyors have utilized specialized measurement and optical devices to estimate coordinates in the field. They created an elaborate framework of benchmarks for registration of additional data themes. Fortunately, we have witnessed a revolution in the way coordinates are now associated with features. Even before the establishment of the Global Positioning System (GPS) laser range finders and total stations had modernized the surveying profession. Theodolites and electronic distance measurement devices enabled surveyors to measure long distances and calculate angles based on line-of-sight rather than direct access. However, the deployment of the GPS network was a game changer. This was a remarkable achievement that opened the door to the geographically enabled world that we now enjoy. The GPS network enabled a receiver to interpret latitude and longitude without linking to a network of existing control points. However, it took President Clinton's executive order in 2000 to remove selective availability for the civilian sector. The impact was immediate. Usery et. al. (2018) provide an excellent discussion of the impact of GPS.

However, with the initial availability of GPS location signals in the 1980s, and particularly with the decryption of the signal and public availability in the 1990s, GPS and its counterpart global positioning systems in Russia and Europe and regional systems in a host of countries (China, India, and others) have become the standard of geolocation and have launched thousands of new location-based services, many of which changed the fabric of our social and business systems. (Usery, Varanka & Davis 2018, 386).

Over the past twenty years, the GPS network became part of an international Global Navigation Satellite System, (GNSS). This umbrella organization includes Russia's GLONASS and the European Union's Galileo constellations. The larger constellation greatly increases the opportunity to locate the desired number of satellites.

While GPS receivers can independently acquire latitude and longitude positions anywhere on Earth, they do not all generate the same level of accuracy. Improvements in precision are related to whether the receiver can accept more than one signal, can synchronize with ground-based stations and the length of time the receiver collects data. The US government, foreign partners and the private sector have provided the tools to refine the accuracy of the coordinates. These devices are used by licensed surveyors as well as public employees and scientists for direct field data capture. These systems often refine their data to within a meter by linking to a land based Continuously Operating Reference Stations (CORS). When precision work in construction and engineering is required surveyors collect at least two hours of data on a dual frequency receiver and send it to the National Geodetic Survey's (NGS) Online Positioning User Service (OPUS). This service adjusts the data from three National CORS sites to refine the horizontal and vertical coordinates. Every month, users voluntarily upload tens of thousands of these highly accurate data files. Via email the user receives a location information with centimeter level position. This system is used thousands of times each month by surveyors who also contribute their data to densify the National Spatial Reference System. This system is designed to tie together all geographic features to common, nationally used horizontal and vertical coordinate systems" or to make sure that "everything matches up."

The GNSS is also improving our ability to accurately measure the shape of the Earth. This impacts the establishment of local and global datums and reference frames. From the perspective of the GIS community, access to these datums established by the National Geodetic Survey (NGS) and software to handle more than 200 map projections has greatly simplified the integration of geospatial data.

While improvement of in surveying technology has had an enormous impact on commercial and scientific endeavors, the real game changer was the incorporation of a GPS receiver on billions of smartphones. In 2009 Apple turned the iPhone 3gs into a surveying device. Instantaneously, iPhone users are never lost, and they could find and navigate to positions anywhere in the world. While a smartphone with a single receiver can estimate a location within a few meters, the positional accuracy improves to a few centimeters when collecting many records on a modern dual-frequency phone. These new phones can compensate for interference from buildings and connect to the full range of GNSS satellites. This means that for some applications, the next generation of smartphone may compete with dedicated GPS instruments. Even today a common smart phone can directly capture the location of an unlimited number of features such as trees, trash bins, fire hydrants, urban furniture, bike racks, signs, poles, and meters.

For example, even the simple tools on an iPhone and Google maps can capture very useful data such as an inventory of fire hydrants (Figure 3). It is interesting to note that HazardHub, a property risk company, has outsourced the capture of fire hydrants to a group in southeast Asia, nine analysts have identified thousands of hydrants by manually scanning Google Street view images. They found that this manual process was better than an Artificial Intelligence Bot (Foust 2020).

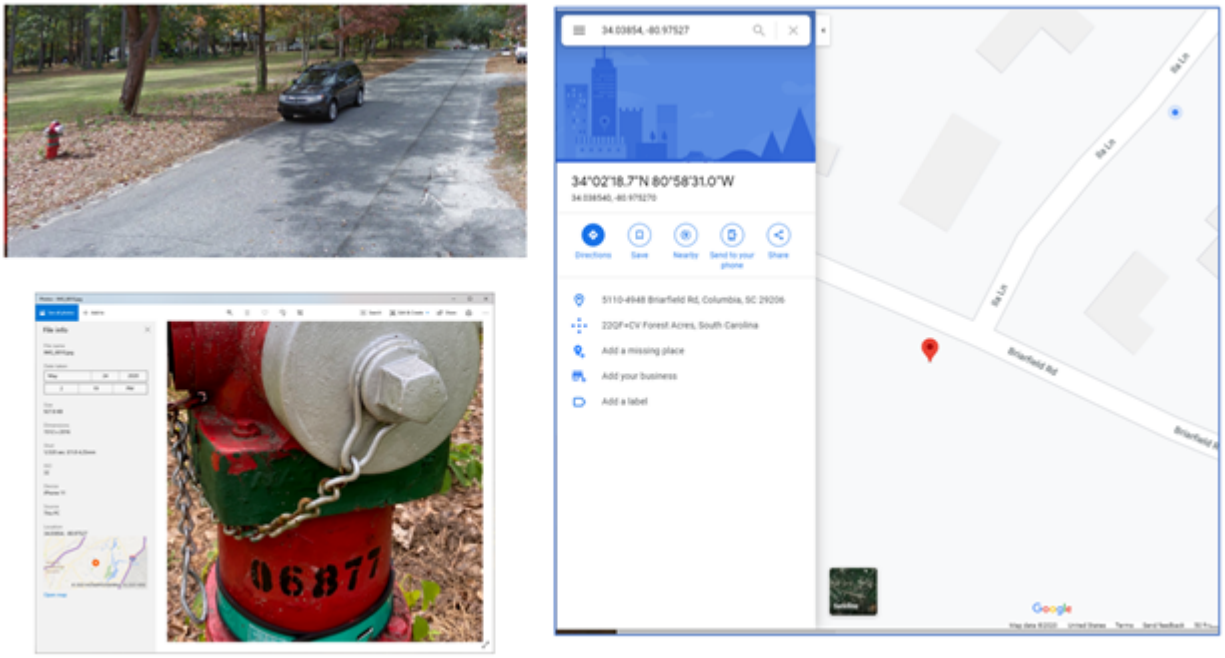


Figure 3. Latitude and Longitude of a fire hydrant with information from iPhone photo displayed on Google Maps. Source: authors.

These same devices are being used by citizens to crowd source the location of potholes, vandalism, and many other areas of concern. A good example of volunteered geographic information occurred in 2017 when 700 volunteers set a Guinness world when they used their phones to “survey” the King Tide in Virginia (Virginia Institute of Marine Science, 2017. However, the most remarkable application has been the creation of Open Street Map (OSM) created by volunteers who capture traces of streets and points of interest during mapping parties. In many areas OSM is the best available map for civilian use. Furthermore, the USGS uses a a group of volunteers called the National Map Corps, that “have made more than 500,000 edits to over 400,000 structure points” (Usery, 2019).

4.2 Georeferencing Photos

The history of aerial photography includes a wide range of options from hot air balloons to orbiting satellites. The importance of imagery as the foundation for data capture cannot be overstated.

Air photos were a critical asset in World War I. Following the war several aerial photography programs were initiated by the Department of Agriculture and the USGS to map and monitor changes on the Earth’s surface. Over time, photographs evolved from black and white, to color taken from manned aircraft, to output from multispectral data sensors captured by satellites. From the perspective of GIS data capture vertical photographs taken directly below an aircraft are the most useful format. However, the scale on the photo changes outward from the nadir position. Special optical devices such as a zoom transfer scope were used to manually align features on a photo to the base map. In a digital environment, photos and maps can be georeferenced to compensate for differences in scale and orientation. The term “rubber sheeting” is often used to describe the process. The accuracy of this georeferencing process depends on the selection of a good set of photo identifiable control points with known coordinates. The process involves

creating a series of links between the features on the image and corresponding features on the ground. The best sources of these ground control points are surveyed targets such as a white X in an open field or road. In other cases, street corners, hedges, building edges, roads, docks, and other features have been used. A set of links spaced across the image is used to calculate the parameters for a transformation that shifts and warps each pixel to a new position. Georeferencing is an iterative process that allows the operator to adjust the control points, links, and choice of transformation. Most software provides a choice of transformations such as polynomial, spline, projective, or similarity. An error table is generated to evaluate the accuracy of the transformation. Ultimately, the operator will visually inspect the results of the transformation and accept a level of error. The final step is to rectify the georeferenced image to standardized coordinate systems. The rectified source material is then ready for display or to serve as a base map for extraction or update of features. Over several decades, countless photos have been georeferenced.

The tools to georeference images are so straight forward that the Leventhal Map & Education Center of the Boston Public Library operates a web-based application that allows users to interactively georeference historical maps and photos. These georeferenced air photos and maps are ready for a analysis and visualization with the popular swipe tool that graphically portrays two views of the same area. They are a quick way to illustrate change (Figure 4).

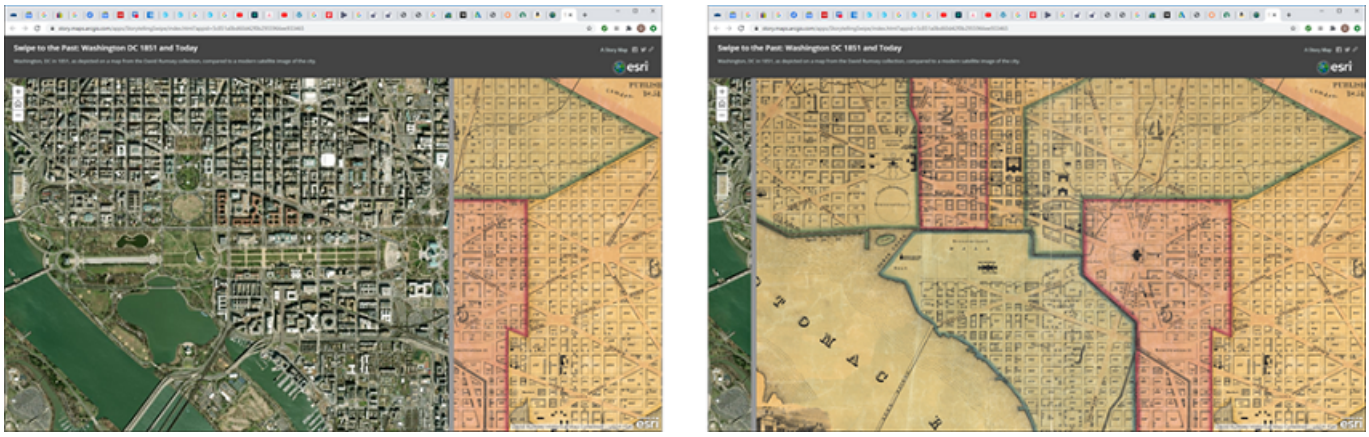


Figure 4. Comparing a georeferenced historical map of Washington, D.C., over aerial imagery, via Esri's swipe application. Image source: authors.

4.3 Orthophotography

The orthophoto process warps the source image so that distance and area correspond with real-world measurements. This allows photos to be used directly in mapping applications as it mitigates distortion otherwise present in aerial photography. Developed by photogrammetrists in the 1960s the process takes overlapping stereo images and a digital elevation model to adjust for variations in terrain and tilt of the aircraft (Figure 5).

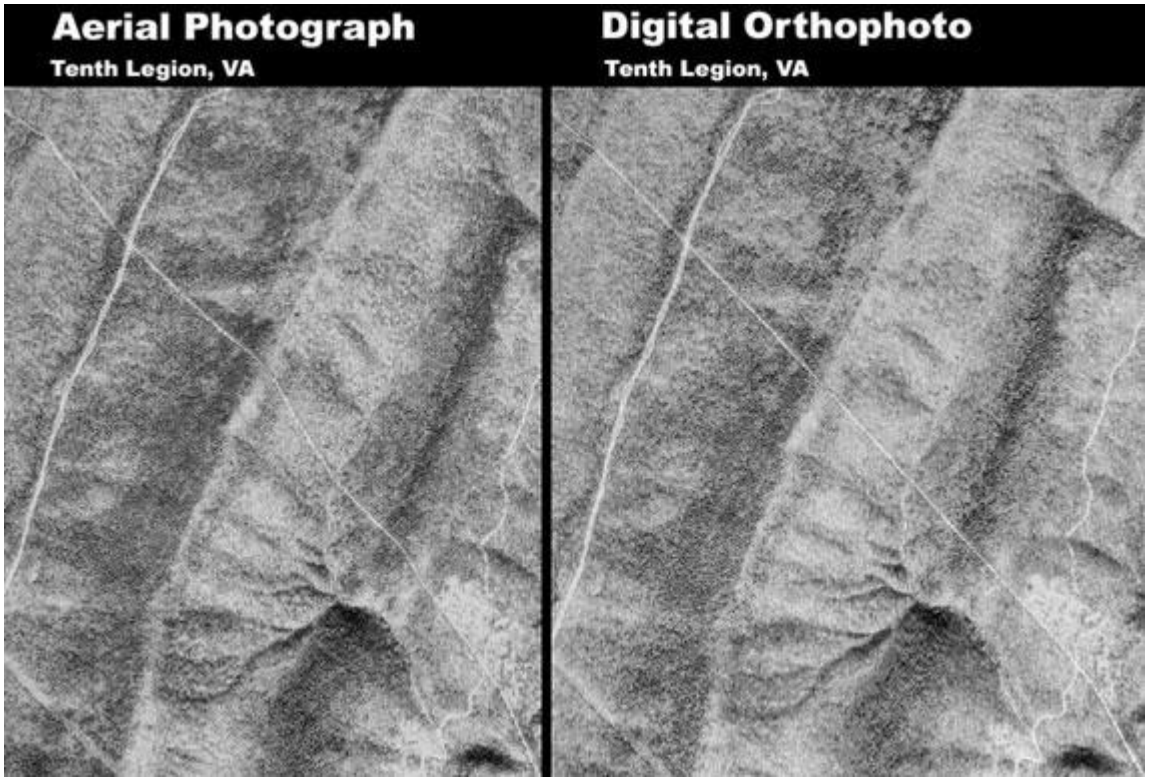


Figure 5. Comparison of georeferenced aerial photograph and an orthophoto. Note the difference in the representation of the straight pipeline. Source: USGS.

4.4 Automated Classification and Feature Extraction

The evolution of data capture has been highlighted by the conversion from manual to automated procedures. In terms of raster data from cameras remote sensing researchers have devoted their careers to finding the best combination of spectral signals for identifying different types of land cover and monitoring conditions on the earth. The classification algorithms are critical tools for finding new features e.g. wetlands, measuring characteristics e.g. crop yields or detection change e.g. urban sprawl. These new tools have evolved from the broader field of image processing. Much of the current work on pattern recognition, artificial intelligence and machine learning focuses on automated feature extraction. In surveillance applications machine learning tools can be used to find specific targets. Radiologists employ the same tools to find tumors. In domestic applications these tools can find specific objects like swimming pools, additions to buildings, or areas of unhealthy vegetation. In a similar fashion LiDAR elevation and intensity values can also be analyzed to isolate trees, powerlines, structures, and other features on the surface (Figure 6).

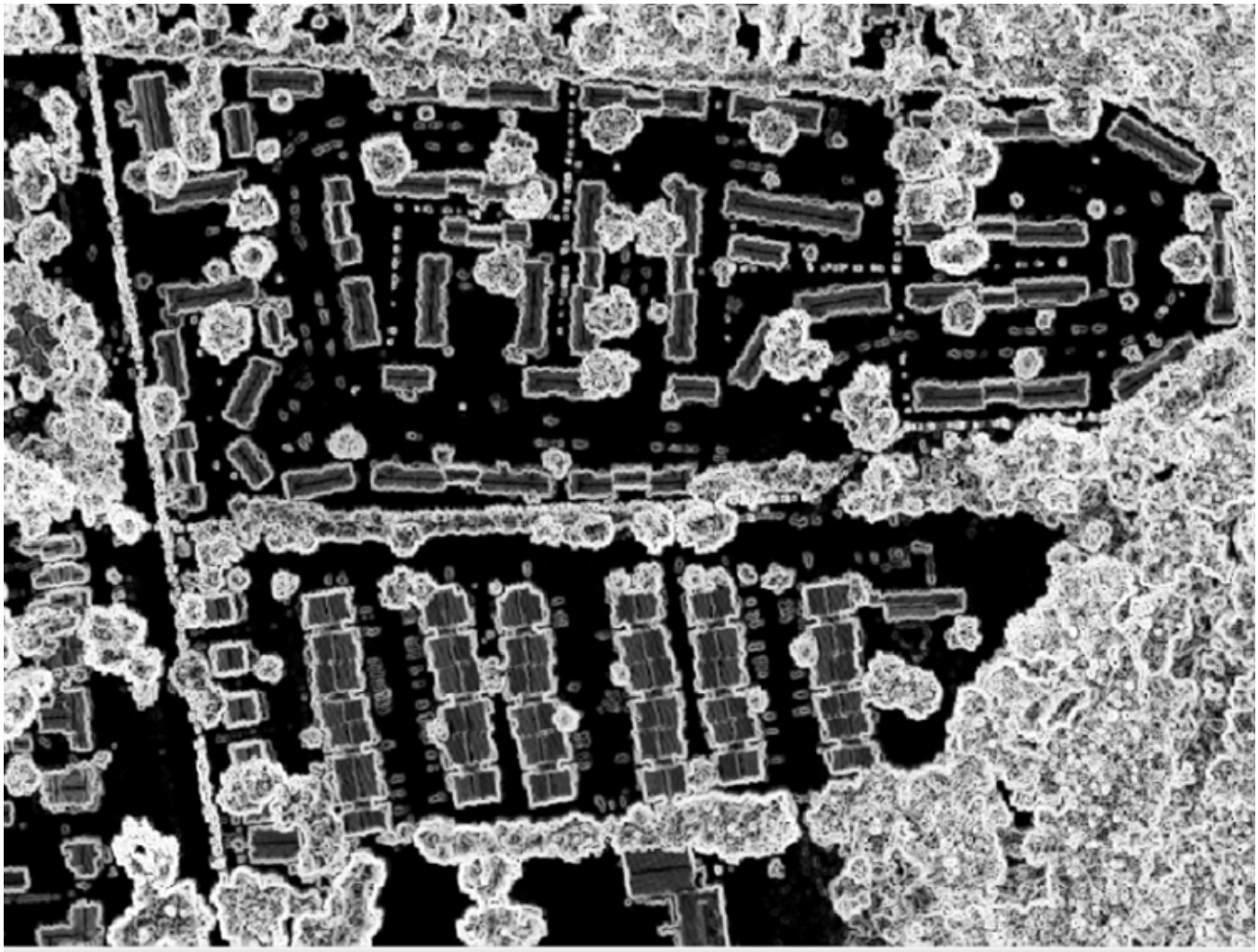


Figure 6. Extraction of individual structures and trees from LiDAR. Image source: authors.

Many image processing routines will generate polygons from the classification images. A recent analysis by Microsoft highlights the status of automated feature extraction. Microsoft analysts accessed five million satellite images to capture the edges of buildings and create polygon footprints for 125,192,184 buildings in the United States (Figure 7). It is noteworthy that a major IT company wanted that information and then shared it. From a technical viewpoint the analysis puts a new marker on the phrase “even visible from space” and certainly indicates that geospatial data capture is ready to handle big data.



Figure 7. Microsoft building footprints for a section of Columbia, South Carolina. Data source: Esri's Living Atlas. Image source: authors.

5. Conclusion

The Microsoft rooftop example illustrated above demonstrates the current state of the art in terms of the supply and demand for geospatial data. However, it is only part of a larger movement that impacts the way individuals live, work, and play in a geospatially-enabled society. As a result of technological advancements, demands for and investments in many programs exists, as they are able to draw upon a remarkable storehouse of geospatial data. Numerous geospatial platforms have emerged that provide easy-to-use tools for locating existing geospatial resources based on standardized metadata. These platforms include the Esri Living Atlas that hosts thousands of authoritative layers contributed by its user community. Many of the nation's major research institutions have built and shared geographic data discovery platforms using GeoBlacklight, a "multi-institutional open-source collaboration building a better way to find and share geospatial data." These search engines provide a experiences for users to discover, browse, and download geospatial data for a specific area of interest. Part 2 of this entry on the changes to geospatial data capture will focus on the implications of these changes, including the existence of vast new collections of data.

References

Foust, B. (2020). Personal communication.

Goodchild, M. F. (2011). Looking Forward: Five Thoughts on the Future of GIS.

Esri ArcWatch.

National Geospatial Advisory Committee (NGAC). (2009). NGAC Report: The Changing Geospatial Landscape.

Tomlin, C. D. (1990). Geographic Information Systems and Cartographic Modeling,1st Edition. Englewood Cliffs, New Jersey: Prentice-Hall (now Pearson Education).

Usery, E. L. (2019) U.S. Geological Survey Accomplishments in Cartography, 2015–2019. US National Committee for the International Cartographic Association US National Report.

Usery, E. L., Varanka, D. E., and Davis, L. R. (2018). Topographic Mapping Evolution: From Field and Photographically Collected Data to GIS Production and Linked Open Data. The Cartographic Journal. 55:4, 378- 390.