

Vector Migration to Image Basemap

Timothy L. Haithcoat, Wenbo Song, and Derek Smith
Geographic Resources Center (GRC) – Department of Geography
University of Missouri-Columbia

Introduction

Due to rapid advances in information technology, many local governments have begun to use geographic information systems (GIS) and millions of dollars have been invested to transfer paper parcel maps and other data into digital format. Unfortunately, many of these historic parcel mapping systems lack the positional accuracy required to create a seamless data base. In the situation when the local government's vector data are overlain to remotely sensed imagery with higher accuracy, the problem of positional accuracy becomes obvious. For example, parcel lines cutting through houses. This becomes an embarrassment to the local government officials and the geographic information technology groups responsible for these data. The inaccurate vector data have to be adjusted. Traditional transformation methods are believed to be ineffective because the process of mapping often pieces together data from multiple sources, scales, and accuracies, each with their own distortion characteristics. Therefore, the maps do not allow one function to transform the entire map space effectively. One solution is extensive manually editing. This is very labor intensive and time-consuming. Another possible solution is re-mapping the entire area. That however is too expensive and the abandonment of existing work and investment cannot be justified.

In the case of Boone County and the City of Columbia, the parcel map had a variable accuracy of up to 40 ft plus or minus once the various tiles were combined. This is the result of data being built by hand historically and remaining un-edgematched between tiles within a mylar mapping system. The investment to convert this (the only base map widely used) was made (\$200,000) and the sheets were scanned and vectorized by the private sector, which very accurately reproduced the inherent errors of this mapping approach. With the increase in usage of GPS within the City and County to map specific point features, as well as the associated problems discovered when trying to edge-match the tiles into a seamless database, the consortium was stymied by the range of variability exhibited within the data bases. This led to the development of a series of image base maps within Synergy I that included DOQQs (1995 vintage) as well as IKONOS (2000). In Synergy II we are developing a 1m aerial multispectral base map obtained in 2001. The Synergy I imagery allowed for the creation of a single image base. This base is now being used by users and developers of numerous data sets being created and compiled within the county, city, and private sector. This single base allows all future data to be spatially integrated and thereby maximizing their investment. This still left however their entire historic or legacy data base structure in CAD and GIS as positionally unusable.

This is a very common problem. Unaware of these inaccuracies or not being able to easily or cost effectively rectify the positional inaccuracy, many local governments continue to digitize into the parcel base. This map is often used as a base to compile other data such as zoning, utilities, etc. In this way the local government keeps the relative relationship between features (i.e. relative accuracy) but still is unable to integrate GPS work or imagery into its standard business or operating model. It was this issue of enabling and leveraging these historic / legacy investments in data and data bases, and their inherent relative accuracies, that this research addresses.

Objectives

The objectives of this activity were:

- To improve the positional accuracy of existing, lower positional accuracy digital vector data utilizing an imagery base map such as aerial photography, USGS DOQ (Digital Orthophoto Quadrangle), or high-resolution satellite imagery such as 1 meter IKONOS.
- To assemble the process so as to limit the amount of manual labor required.
- To allow for visualization of the vector data displacement necessary to match the imagery source.
- To create a user interface for this procedure within ArcInfo / ArcGIS software.

User Community of Focus

User clients included: Ross Short, City of Columbia and Boone County GIS Coordinator; David Storvick, City of Columbia, Engineering Department; and J.R. Richardson, Boone Electric Cooperative. These people represent the core of the local consortium and we the ones most concerned with the development of a single base map and the migration of the various data layers to a more known and accurate coordinate system. From this group, the parcels and other data layers will be distributed through the consortiums server to all parties for their use in the various applications to which their mandate defines.

Product Development

Many transformation techniques have been used in GIS. Piecewise transformation have been developed and used by computer cartographers over the last 30 years (Christensen 1996). They are characterized by the subdivision of the maps in pieces and by the use of different transformation parameters in each piece. The aim is to adjust maps showing deformations that do not follow a clear trend. Such deformations are easily found in parcel maps compiled piecemeal from non-contemporary survey documents of diverse origins and scales.

The piecewise transformation is achieved by constructing a surface model through Delaunay triangulation. The vertices of the Delaunay triangulation are the locations where control points should be (shift to) and their Z values represent the amount of distortion in x or y direction that the original x, y coordinates need to be adjusted (shift from). Delaunay triangulation is a proximal method that requires that a circle drawn through the three nodes of a triangle will contain no other point. This means that all adjustments are localized to the area of these triangles. Another point of consideration is that these triangles be 'fat' in that there is equity in the 'weights' associated with the displacement factors calculated for these points.

After the transformation, the coordinates of features located at the from-location will be migrated to the to-location. For other points, their new location will be interpolated based on their distance and direction from the known points forming the triangle. For this interpolation there are two approaches: linear and bivariate quintic.

The linear interpolation method considers the surface as a continuous faceted surface formed by triangles. The normal, or imaginary line existing perpendicular to the slope of the surface, is constant throughout the extent of each triangle facet. However, when an edge separating two

adjacent triangles is crossed, the normal changes abruptly to that of the next triangle leaving steps and sharp breaks within the surface.

Quintic interpolation also considers the surface model to be continuous. In addition, quintic interpolation considers the surface model to be smooth, that is, the normal to the surface varies continuously within and between each triangle. Thus, there are no abrupt changes in the normal as it crosses an edge between triangles. This smoothing characteristic is accomplished by considering the geometry of the neighboring triangles when interpolating the shift value (z) of a point in a triangle. Quintic interpolation employs the bivariate quintic interpolation algorithm using a bivariate fifth-degree polynomial in x and y. The quintic interpolation method is used to minimize abrupt changes between the shift surfaces resulting from these triangles.

Successful implementation of the piecewise transformation requires a dense tessellation of triangles created through a good distribution of control points. These control points need to be easily identified from both the vector database and imagery. Since most property corners are not visible in imagery, road intersections are used as the basis for the tessellation. If the local government GIS database does not have a road centerline layer, it can be automatically created from dual-line features such as road casings or parcel data, as well ICREST has developed a road feature extraction tool to extract estimated road centerlines from imagery.

After developing the road centerline, it is graphically overlaid to the remotely sensed imagery being used as the base map. The human operator must then identify and tag each corresponding intersection from the imagery. Next the corresponding point pairs are found and their coordinates are used to create two shift surfaces, one for the 'X' coordinate and one for 'Y' coordinate. Finally, the parcel and any other vector layers built on that base will be migrated to their new positions using the constructed X and Y shift surfaces.

We tested the piecewise transformation method using real world data. The city of Columbia, Missouri has already digitized their parcel map tiles. The DOQQ / IKONOS base map imagery was used as the base map for the migration. It must be noted that the geo-referenced and ortho-rectified imagery base had a much better positional accuracy than the parcel map.

When the parcel data were originally overlaid on the digital base map (Figure 1), it did not line up with aerial photograph. Many parcel lines cut through houses, and some parcel lines even fall in another blocks.

All the road intersections are used as control points. The road centerline and parcel layers are displayed on top of aerial photograph (Figure 2). They act as reference for locating the corresponding road intersection points on imagery. A human operator collects these points manually and saves it to a points file. The corresponding point pairs (from, to) are obtained by finding the nearest distance between to-point and from-point layers. Therefore, the x, y coordinate attribute of from points can be transferred to the to-point layer. The to-point layer is used to create two TINs whose vertex are to-points location and Z values are x, y coordinate of from-points respectively.



Figure 1. Overlaying original parcel lines on aerial photograph (part of test area).



Figure 2. Finding the road intersection pairs: x symbol is location in vector database, and circle symbol is where it should be in imagery.

Figure 3 show the x coordinate shift distribution between from and to control points and Figure 4 show the coordinate shift in y coordinate. It can be clearly seen that the coordinate shift varies over space and there is no clear trend. Also the shifts in x & y direction have no correlation. That verifies the piecewise transformation is the most appropriate method that deals with the positional accuracy problem of parcel data.

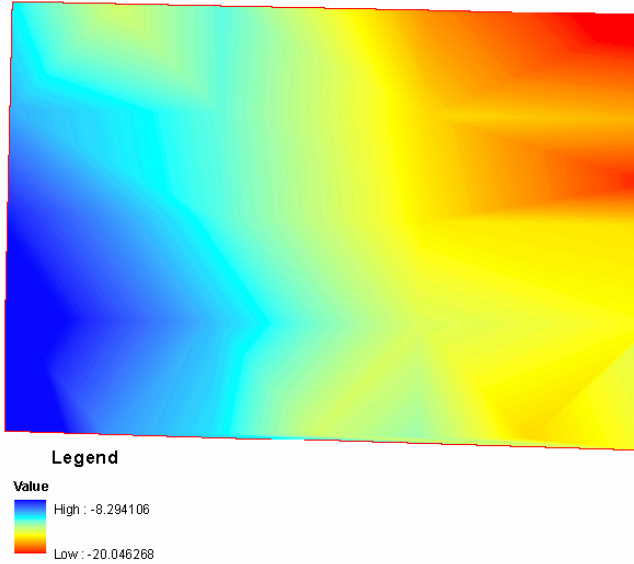


Figure 3. The coordinate shift in x direction.

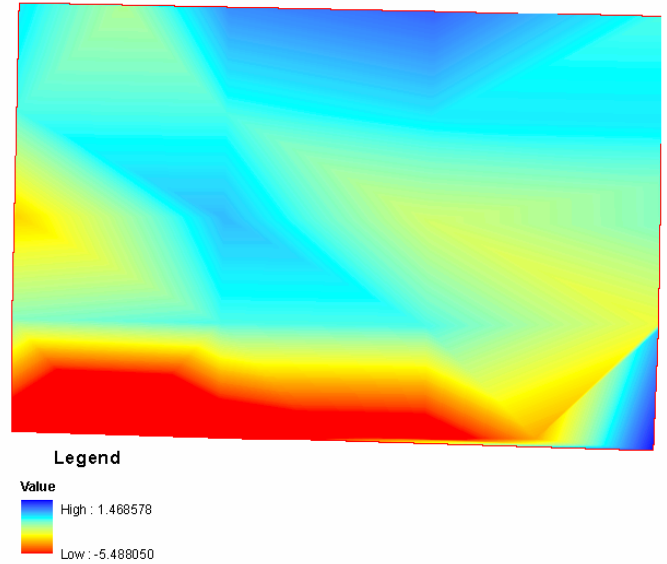


Figure 4. The coordinate shift in y direction.



Figure 5. The parcel data after piecewise transformation.

The results after piecewise transformation are shown in Figure 5. The parcel lines are within the block and most lines line up well with the imagery. There is no problem of parcel lines cutting through houses.

Experience of User Community

This process was tested within our lab for the entire Boone County / City of Columbia parcel database. This provided our staff an opportunity to Beta-test the process, develop a series of ‘tips-and-tricks’, as well as develop a full understanding of the process and how a user interface would need to be compiled for this tool and process. Parameters were set by the City-County GIS Coordinator and other users through the Synergy I QFD process as to what positional accuracies were desired for this migration. At an earlier time the Consortium attempted to edgematch these same data using a standard approach within the existing GIS software. The development took over two person-months. While frustrated in that when they fixed one area they broke another, they did create an edgematched file for much of the area. The development of the vector migration surfaces and the completion of the migration took one-half of a person-month. These two data sets were just recently compared and discussed within the Consortiums Technical Committee. They found that many areas were represented the same but in other areas the vector migration was superior. In their preliminary analysis they found no areas where their process was better. As well, an item of accountability came into play in that the original edgematched file did not retain any information as to how far any vector, point, or polygon was moved. This created a general feeling of mistrust in that no one knew where this data base was more faithful to the input or less faithful. It was unmeasurable. For the assessors office and others this was not well received. A product of the vector migration methodology however is the X and Y shift surfaces that can be viewed or visualized under the resulting coverage so that for every point or location within the migrated database the user knows how far the original had to be shifted to match the imagery base. This is an extremely valuable metric for allowing the integration of these techniques and protocols into the mainstream office processes.

A second Beta-test was just initiated where we sent a trainer with the protocols and knowledge to Springfield, Missouri. In this test we wanted to deliver the methodology and

determine how well the process could be implemented by a team of technical persons from the local government itself. In this case a consortium of again City, County, and Utility staff were brought together in a single lab setting.

In this case each organization was charged with certain record keeping responsibilities that did not reflect the technology available or differing technological approaches. The use of a single basemap across the consortium was the goal. The incorporation of GPS and the associated problems of a seamless, integrated basemap with unknown positional accuracy, had caused the consortium to reach an impasse for the development of individual applications on a common basemap. The use of high-resolution imagery with high positional accuracy (± 10 ft RMS) provided the common ground for migration of all files and allowed for their integration and maintenance. The only way the group achieved full participation to produce and maintain this file was through the use of an imagery-based solution. Once the GIS data are migrated to a more accurate position, imagery can be used for more accurate assessment, planning, growth management, infrastructure assessment/reporting, and change monitoring.

The consortium has invested approximately \$13.5 million dollars and 194 person years in the development of 175 additional GIS data layers that referenced an inaccurate basemap. The consortium had, in collaboration with the University of Missouri – ICREST, access to the needed tools and techniques for the development of a high resolution, high accuracy image basemap and the accurate migration and maintenance of the large number of vector data layers to the image base. These issues had to be resolved in order to operationalize remote sensing data into these various local government departments.

A single day of training was held using their own vector data and an imagery base from 2001. The training was conducted by a graduate student funded through Synergy II. The technology transfer went well. This local government group is now in the process of generating the vector migration protocols for their data. They have adopted the tools and protocols and are now applying them to their data. We are now awaiting any questions or notification that the process has been completed. The initial response has been very positive.

The participating departments have mandated responsibilities from the City Charter, City Council ordinances, County Commission, Board of Public Utilities and State Legislature. These mandates relate to record keeping in many forms, map/drawing productions and maintenance, economic development, and ultimately public safety. The Consortium's GIS participants (City of Springfield, Greene County and City Utilities of Springfield) are in constant need of accurate geospatial information.

The application will increase the positional accuracy of these files. In this way resolving the discrepancies in the two parcel base layers that are the foundation of their GIS user applications. The application also develops a shared responsibility workflow enabling the Consortium to eliminate redundancy and encourage an integrated and cooperative effort to more efficiently meet the mandates for record keeping that each respective organization is charged with. The application provides future decision makers and the public with a much needed level of confidence in the information they are using. The application allows city, county, and private sector users to maximize the use of spatially integrated remote sensing and GIS vector data.

The benefits realized from this integrated application are both tangible and intangible. These include:

1. the use of these data in areas such as desktop flood plain evaluation, emergency vehicle route mapping, automated vehicle locating, planning and zoning, etc. ;
2. better and broader access to imagery with base reference layers (parcels, roads, etc.) for tracking change and growth as well as increased speed of access to information (for many first access);
3. set procedures for use of future imagery data sets;
4. personnel time savings in daily work flow and increased staff productivity;
5. allowing for consensus to be reached on the first comprehensive basemap to be created within and across their user constituencies (city, county, utility);
6. creating an environment of data sharing and leveraging of resources within the consortium;
7. instilling a belief in the necessity of imagery to the future growth management and infrastructure planning within the cooperators;
8. providing a foundation upon which their spatial data base conversion and development can occur;
9. allowing the local citizenry access to the information upon which the city and county are basing their decisions which itself provides a forum for better understanding of these decisions.

It should be noted that in both of these cases the consortiums assembled by the local entities predated our participation. In each of these cases the consortium was beginning to self-destruct due to the lack of a consistent base map that was usable with the various imagery bases that were being generated for these jurisdictions. The Information Technology departments as well as the GIS staff were frustrated. This ability to create a bridge from their legacy data and linework into a more positionally accurate coordinate system using imagery was seen as the saving point for these cooperatives.

This is critical work as it is this migration and spatial referencing capacity which will allow the legacy systems and historic data bases to be integrated with the remote sensing data available today. The tool kits and interfaces to allow for this conversion process will continue to be pursued and refined within in Synergy III to allow for incorporation of certain CAD components (or the development of work-arounds) and other spatial object types. As well, marketing, training and education in the use and utility of this approach and tool is needed to ensure proper implementation and integration within local government applications. Development of a more robust integration of road centerline extraction from imagery with the vector migration tools will also be pursued in Synergy III. As well, a compilation of testimonials and accuracy assessments from the Beta-tests will be completed. Finally, cost/benefit analyses need to be completed (and understood) so as to present to the user the range of possible methods of achieving the migration of their vector data. This would include imagery options, base map options, and training or service provision.

Funding Support

Sources for funding of this research were Raytheon Synergy (60%) and NASA (40%).