Comparison of Two Sign Inventory Data Collection Techniques for Geographic Information Systems

ALLEN POLING, JIM LEE, PATRICK GREGERSON, AND PAUL HANDLY

Global positioning system (GPS) units are fast becoming powerful tools for collecting data for use with geographic information systems (GIS). The data collection process for a sign inventory conducted by Lee Engineering in Washington, D.C., is described. In the conduct of the sign inventory project, GPS was used where possible. However, the GPS was determined not to be the best method of data collection in all instances. In sections of the study area near the downtown core, taller buildings effectively blocked satellite signals. In these areas GPS collection was set aside for a more manual method of locating sign positions using a measuring wheel. Data collection efforts for the sign inventory using GPS typically required more data collection time; however, the data were quickly and easily exported to a compatible GIS format. Conversely, data collection using manual methods minimized data collection time but required much more effort to enter into the GIS data base. The collection of data using both GPS techniques and manual techniques allows for comparison of both techniques to determine which was more cost-effective.

Global positioning system (GPS) units are fast becoming powerful tools for collecting data for use with geographic information systems (GIS). This type of equipment was recently used in Washington, D.C., to collect sign inventory data for input into a GIS for the National Park Service. However, in areas where taller buildings masked the skyline, creating an "urban canyon" effect, the GPS units were abandoned for more manual methods of inventory data collection. This paper documents the data collection process using the GPS equipment and manual techniques and provides a comparison of the time required to collect, process, and input the sign inventory data collected by each process into the GIS.

STUDY AREA

The sign inventory activities were conducted as part of larger data collection efforts being conducted within two study areas of Washington, D.C. These areas, the White House and Memorial Core study areas, are shown in Figure 1.

The White House study area is bounded by 17th Street to the west, H Street to the north, 15th Street to the east, and Constitution Avenue to the south. This area is basically composed of two land uses. The southern portion, containing the Ellipse, is primarily open park land. The remainder of the study area is more oriented toward the central business district (CBD), with large multistory buildings lining the roadways.

A. Poling and J. Lee, Lee Engineering, Inc., Suite 310, 2701 East Camelback, Phoenix, Ariz. 85016. P. Gregerson and P. Handly, Center for Urban Ecology, National Park Service, National Capital Region, 1100 Ohio Drive, S.W., Washington, D.C. 20242.

The Memorial Core study area is bounded by Ohio Drive to the south and west, Constitution Avenue to the north, and 14th Street to the east. This area is primarily open park land and contains the East and West Potomac Parks, the Jefferson and Lincoln Memorials, and the Washington Monument.

The sign inventory activities conducted as part of this project included collecting traffic and pedestrian signage information on all roadways within and one block adjacent to these study areas.

INVENTORY PREPARATION

To conduct the sign inventory, a four-step process was followed. The first step of the process involved project planning or inventory preparation. During this step, several key tasks for the data collection process were conducted. These tasks included

- Develop preliminary list of signs,
- Identify attributes to be collected,
- Develop a data dictionary (data collection program), and
- Conduct mission planning.

Before conducting the sign inventory, an effort was made to identify the types of signs that would be encountered in the field during data collection activities. This was done to obtain a better understanding of the numbers, types, and conditions of signs that would be included in the inventory. To do this, personnel spent several hours in the study areas photographing and identifying the signs that would be encountered. Each unique sign type was then assigned a unique five-digit code for identification. In addition, a computer-aided design (CAD) drawing was made of each sign that included the five-digit code.

The second step of the inventory planning process was the identification of the attributes, or characteristics of the signs, to be collected during the inventory efforts. This task was conducted with much input from the National Park Service. As a result of this task, the following attributes were identified to be collected as part of the inventory efforts:

- Number of signs on the assembly;
- Sign number: from 1 (top sign on assembly) to the total number of signs on the assembly;
- Direction the sign is facing: north, east, south, west, northwest, northeast, southeast, or southwest;
- Mounting height: in feet;
- Sign code: a five-digit code corresponding to a drawing of the sign;

- Sign size: width and height in inches; and
- Sign condition: a subjective rating of good, fair, or poor.

Once the attributes were identified for data collection, a data collection program was developed for use with the GPS equipment. This program allows the user to enter the sign attribute information directly into a hand-held microcomputer that composes part of the GPS unit. This is done through the computer's keypad or a bar code reader (or both). The program was developed using software provided with the GPS unit and was downloaded into the GPS units used to conduct the inventory.

The last task conducted involved identifying satellite conditions (number of satellites and satellite geometry) in the area of the inventory before actual data collection. This process is referred to as "mission planning." To conduct this process, a current almanac file (a file containing the orbital information of all available GPS satellites) was collected using a GPS unit. This almanac file was used to anticipate satellite paths and determine the time periods that good positional data could be obtained in the study areas. This was done using the software provided with the GPS unit.

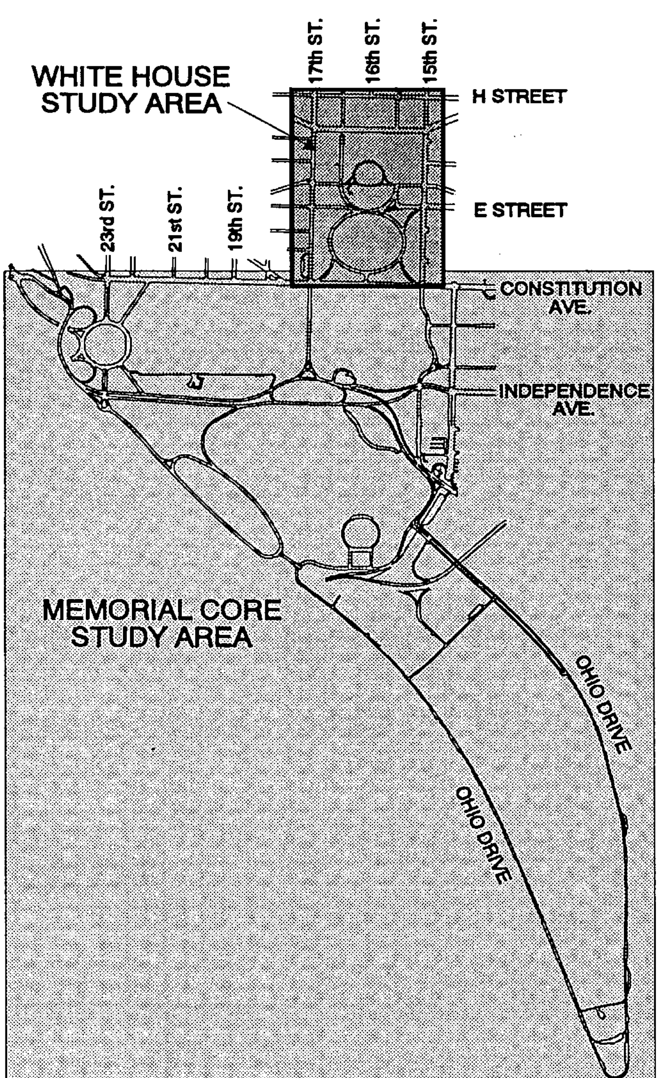

FIGURE 1 Study area.

FIELD DATA COLLECTION

The sign inventory data collection was conducted during four data collection trips to Washington between October 1992 and February 1993. At the onset of the project it was anticipated that the GPS units would be used to conduct the entire inventory. However, on the basis of problems encountered during data collection efforts in October and December 1992, it was decided that the GPS equipment be abandoned for a more traditional method of data collection in parts of the study area. These problems encountered included leaf coverage from trees and tall buildings, which blocked out portions of the skyline and often resulted in insufficient satellite coverage, and poor satellite geometry and possible multipathing as a result of radio equipment in the White House compound.

When the data collection efforts began in October 1992, the trees within the study areas were heavily covered with leaves. During the inventory of signs adjacent to the trees, locks on one or more of the satellites being tracked were frequently "lost" (i.e., the satellite signal was not strong enough to detect or use), often resulting in an insufficient number of satellites being tracked or unacceptable satellite geometry. This problem was fairly easily avoided, however, by using two methods. First, data collection was postponed until December, when the leaves had fallen off most of the trees. Second, a taller range pole (12-ft) was used for the GPS antennae. During the October data collection efforts, an 8-ft antenna was used.

Tall buildings in the area provided a larger problem. Attempts to inventory signs in the downtown CBD portions of the study area were unsuccessful. In these areas where buildings blocked much of the skyline, it was difficult to impossible to obtain signals from enough satellites. In addition, when enough satellites were available, the geometry of the satellites was often unusable. In these areas the GPS equipment was set aside for more manual methods of data collection. The process used for both methods of data collection are discussed in the following.

GPS Data Collection Process

The GPS unit used to collect the inventory data consisted of a receiver, an antennae, a hand-held microcomputer, and an optional bar code reader. The microcomputer controls the receiver and is used to input and store the attribute data. The equipment used provides location accuracies of 2 to 5 m when used with differential correction. Data collection was conducted with the GPS equipment operating in the manual three-dimensional mode. This mode requires a constellation of four satellites with adequate geometry for positioning. This mode also provides the greatest accuracy for data collection and reduces postprocessing time.

To collect the inventory data, the GPS software was started, a file was created to store the field-collected data, and the data dictionary for the sign inventory was selected. The GPS antennae was placed by a sign and the "feature on" key was pressed. This causes the receiver to begin collecting position data, which were collected at a rate of one position per second.

As the position data were being recorded, the data collector was prompted to enter the sign attribute information. When this had all been entered, the data collector pressed the "feature off" key, which causes the receiver to stop collecting position data. The data collector then moved on to the next sign. When a sign was encountered that had not been assigned a sign code, it was assigned a code, sketched, and photographed. A CAD drawing was later made of the sign.

To obtain the best possible accuracy, a minimum of 180 positions were obtained at each signpost location as recommended in the GPS documentation. At locations with more than one sign per post, enough positions were collected for each individual sign so that the sum of all positions collected at the signpost location exceeded the 180 threshold. For example, if the post contained three signs, the data collector might obtain 60 positions for each of the three signs for a total of 180 positions.

At the end of each day's data collection, the data collection file was closed and the data downloaded to a personal computer for postprocessing. Data were collected for 2,772 signs (77 percent) using the GPS equipment.

Manual Data Collection Process

During the manual data collection activities, the sign location was determined using a measuring wheel, and the sign attributes and locational information were noted on hard copy (paper) or audiotape. This information was later postprocessed and entered into electronic form for entry in the GIS package. When a sign was encountered that had not been assigned a code, it was assigned a code, sketched, and photographed, and a CAD drawing was made of it later.

To determine the location of each sign, two measurements were required: the offset from the face of curb and the distance from the face of curb of an intersecting street. Using this information and a digitized map of the roadway network, the coordinates of the sign were determined. Data were collected for 830 signs (23 percent) using the manual method.

Postprocessing GPS Data

After data collection, the data were postprocessed to manipulate them into a form compatible with the GIS packages being used. The first step in postprocessing the GPS data was to differentially correct the rover files (field data collection files). During the data collection process, the National Park Service set up a base station to collect base station data. Using the GPS software, the base station files were referenced and used to differentially correct each day's data. This process eliminates many of the errors inherent in GPS, primarily selective availability, which accounts for the largest portion of all errors.

After differential correction, the attribute information was exported into an ASCII file format compatible with the GIS package used by the National Park Service and the GIS package used by Lee Engineering. This also was done using software provided with the GPS unit.

Processing the position data proved to be a more complicated task than expected, as most of the information collected was for multiple signs on one post. The GPS software, by default, determines the location of a feature by averaging all of the positions collected for the feature, or in this case each sign. However, the location of each signpost was needed. Fortunately, the locations could be found by forcing the software to average the data over time breaks, allowing the position data collected for all signs on a common post to be grouped together and then averaged to obtain the signpost location.

These signpost locations were then exported to an ASCII file format compatible with the GIS packages. Because the number of locations (one for each signpost) was now less than the number of signs, a utility program was developed that assessed the attribute and location data and assigned signpost coordinates to each sign. The new location data were then exported in a format compatible with the GIS packages and the sign inventory was imported into the GIS.

Postprocessing Manually Collected Data

The manually collected data also required postprocessing. The first step in postprocessing the manually collected data was to load a street coverage file that had been previously digitized from aerial maps in the NAD27-CONUS geodetic datum into a CAD program. A street coverage file had been provided by the National Park Service in a .DXF format.

Sign locations were then determined using an offset command and the curb faces digitized from aerial photographs. To enter the attribute information, a "block" was created that prompted the user to enter the attribute information collected during the data collection activities. Once all of the signs had been located and the sign attribute information entered into blocks, the data were exported to an ASCII file containing the x- and y-coordinates of the sign as well as all attribute information. This file was then split into two files, a point file (containing coordinate information) and an attribute file, compatible with the GIS packages. The data were then imported into the GIS.

COMPARISON OF DATA COLLECTION TECHNIQUES

After the data collection activities, an effort was made to assess the time required to collect and process the data for entry into the GIS packages for each of the collection methods. This was accomplished using time sheet information to determine the total number of man minutes required on a per-sign basis to accomplish each task. The results of the comparison are given in Table 1.

As indicated in Table 1, data collected with the GPS equipment were the most cost-effective, requiring approximately 4.6 man-min per sign for collection and postprocessing. The manually collected signs required approximately 5.5 man-min per sign. The signs collected using the GPS equipment require a longer data collection time (4.0 versus 2.6 man-min) but were postprocessed with much less time.

TABLE 1 Comparison of Data Collection Techniques (min/sign)

Task	GPS Collected (2722 Signs)	Manually Collected (830 Signs)
Field Inventory	4.0	2.6
Post-Processing	0.6	2.9
Total	4.6	5.5

The time difference in the data collection (field inventory task) is explained by the fact that 180 positions were collected at each sign location in order to obtain the best accuracy of the GPS equipment. Collecting positions at a rate of one position per second, this required a minimum of 3 min at each sign location. Locations with more than three signs on a post, typically required 3 or more minutes to enter all the attribute data for the signs. However, at locations with one or two signs per post, the attribute data were entered relatively quickly (40 to 60 sec per sign), but the data collector was required to remain at the location for a full 3 min to collect all of the position data. Using the manual method, the data collector moved on to the next sign as soon as the attributes were noted.

Postprocessing, however, was quite different. As the GPS collected data were already in electronic form, data reduction occurred quite rapidly. In part, this was because the data were reduced as a group (one file containing a day's data collection). However, the manually collected data had to be reduced one sign at a time, which resulted in a longer postprocessing time.

The resources and cost associated with each of the two methods were also compared. Many of the data collected with the GPS units were collected by one or two individuals, an engineer and a technician, when two GPS units were available for use. The data collected manually were collected by one or two engineers, although a technician-level person could have been used. In short, both methods used the same personnel resources. However, the equipment requirements were very different. The GPS units used for the data collection cost approximately $15,000 each, whereas the equipment used for the manual collection cost less than a couple of hundred dollars.

DEVELOPMENT OF GIS PROGRAM

The last step in the collection process was the development of a program to display the collected data in the GIS. The program was developed with several goals in mind, including

- Being easy to use (menu-driven),
- Providing the capability to display the locations of all signs or a select type of sign,
- Differentiating sign conditions using colors (i.e., green for good, blue for fair, and red for poor),
- Displaying the attribute information for any selected sign,
- Displaying an image of the sign, and
- Providing panning and zooming capabilities.

The result of the programming effort was the development of a program that met the goals listed. The program allows the user to show the locations of all signs on a roadway coverage of the total study area or any portion of the study area. In addition, the location of a select sign type (stop signs for instance) may be shown and the sign condition may be illustrated in color. To see sign specific information, a second-level menu was developed that allows the user to select a sign or group of signs and view the attribute information of each. The CAD drawing files for each sign were exported into a Raster format so that an image of each sign could be viewed within the GIS program.

CONCLUSION

Several conclusions can be drawn from this study. First, from a time standpoint the GPS data collection activities proved to be more cost-effective than the manual data collection activities, resulting in approximately 1 man-min less effort required per sign to inventory and postprocess the data. This is attributed to the fact that postprocessing the data is a much quicker task when the data are collected with the GPS equipment because the data are treated as a group and not processed separately for each individual sign. Additionally, both methods require similar staffing levels and requirements.

However, equipment costs are extremely different. The GPS units are very expensive pieces of equipment costing thousands of dollars, whereas the wheel and other equipment used for the manual data collection cost only a few hundred dollars.

ACKNOWLEDGMENTS

This paper was prepared based on the results of the sign inventory activities undertaken as part of the Memorial Core Area Transportation Study/White House Area Traffic Data and Analysis project. The inventory was conducted by Lee Engineering, Inc., Phoenix, under subcontract to BRW, Inc., Denver, for the National Park Service, U.S. Department of Interior.

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the National Park Service, U.S. Department of Interior. This report does not constitute a standard, specification or regulation. Trade or manufacturers' names which may appear herein are cited only because they are considered essential to the objectives of the report. The National Park Service does not endorse products or manufacturers.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.