Practical Paper

Interactive Boundary Delineation of Agricultural Lands Using Graphics Workstations

Thomas D. Cheng
TGS Technology, Inc., MS 242-4, NASA Ames Research Center, Moffett Field, CA 94035-1000
Gary L. Angelici
Sterling Software, Inc., MS 242-4, NASA Ames Research Center, Moffett Field, CA 94035-1000
Robert E. Slye
Ecosystem Science and Technology Branch, MS 242-4, NASA Ames Research Center, Moffett Field, CA 94035-1000
Matt Ma
TGS Technology, Inc., MS 242-4, NASA Ames Research Center, Moffett Field, CA 94035-1000

ABSTRACT: The National Agricultural Statistics Service of the United States Department of Agriculture presently uses labor-intensive aerial photographic interpretation to divide large geographical areas into manageable-sized units for estimating domestic crop and livestock production. Prototype software to automate the boundary delineation procedure, called the Computer-Assisted Stratification and Sampling (CASS) procedure, was developed by the National Aeronautics and Space Administration. The procedure uses image processing software on a Hewlett-Packard graphics workstation for the development of the operational approach. Over a background display of a Landsat Thematic Mapper image and United States Geological Survey Digital Line Graph data, the operator uses a cursor to delineate agricultural areas, called sampling units, which are assigned to strata of land-use and land-cover types. The sampling units are used for subsequent sampling procedures. Three counties in Missouri were chosen for evaluation of the CASS procedures. Analysis indicates that CASS is up to six times faster than the manual techniques in creating sampling units.

INTRODUCTION

THE NATIONAL AGRICULTURAL STATISTICS SERVICE (NASS) of the United States Department of Agriculture (USDA) is responsible for providing domestic crop and livestock production statistics. These statistics are derived from data collected though a variety of sampling techniques and surveys. An automated boundary delineation procedure, called the Computer-Assisted Stratification (CAS) procedure, was initially developed at the National Aeronautics and Space Administration (NASA) Ames Research Center (ARC) to delineate the boundaries of sample units for survey procedures, a task that is currently performed manually (Cheng *et al.*, 1989). To reflect a subsequent addition of the sampling capability of CAS, the acronym was changed to CASS, which stands for computer-assisted stratification and sampling (NASS, 1990). The CASS stratifies the sampling units by land-use and land-cover type, using image-processing hardware and software. The procedure generates coverage areas and the boundaries of stratified sampling units which are used for subsequent sampling procedures from which agricultural statistics are developed.

BACKGROUND

An *area sampling frame* is a construct that serves as a sampling strategy for large-area surveys. An area to be surveyed is divided into contiguous parcels, each of which can be easily found on the ground. The area sampling frame for a state or county consists of a collection of all of the parcels of land in an area of interest framed by either political or natural boundaries. The Area Frame Section (AFS) of NASS in Fairfax, Virginia, is responsible for the development of area sampling frames.

The current manual procedure employed by NASS to develop and edit sampling frames was designed to ensure operational efficiency and to ensure a statistically valid sample. The key steps in the area sampling frame procedure include (from Cotter and Nealon, 1987)

- Stratification—divide the land in an area, according to land-use and land-cover strata, into primary sampling units (PSUs) by means of photo-interpretation processes;
- Digitization—digitize PSUs electronically for area measurements that are used to allocate the samples proportionally; and
- Multi-step sampling—sample PSUs randomly for further breakdown into segments. A segment is randomly chosen and is visited by a field worker.

The manual procedure uses Landsat satellite false-color composite prints, aerial photographs, USGS topographic maps, and other hard-copy materials. The boundaries are delineated, transferred among different scales and media, and digitized using these paper-based materials on an *x-y* digitizer to compute the acreages of PSUs.

The existing NASS manual procedure for creating and editing area sampling frames is slow, labor-intensive, and sometimes inaccurate. The "automated" procedure described here performs area frame functions with display hardware and software on a graphics workstation. The goals of the current work on the automated procedure are to complete the software and implement an operational procedure by the end of 1992.

SYSTEM DESCRIPTION

The CASS is essentially a photographic interpretation procedure using a color display and on-screen interactive delineation of stratified sampling units, that is, PSUs. The objective is to display Landsat Thematic Mapper (TM) imagery and Digital Line Graph (DLG) data using an image processing system as an aid in the on-screen stratification process. To emulate the current manual stratification and sampling-unit delineation procedures as performed by NASS, the CASS procedure uses interactive on-screen digitizing. The work under the CASS procedure is performed on a graphics workstation using digital imagery instead of on false-color composite prints, aerial photographs, and topographic maps.

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 58, No. 10, October 1992, pp. 1439–1443.

0099-1112/92/5810–1439$03.00/0

Hardware System

Both NASS-AFS and ARC acquired similar hardware systems to enable parallel software development and testing; the software was developed at ARC and tested on the same basic system at NASS-AFS. The Unix-based Hewlett-Packard 9000 Model 360 TurboSRX color graphics workstations were selected for the operational development of the software and were procured in late 1989 at both NASS and ARC (Cheng *et al.*, 1989; 1991).

The TurboSRX graphics subsystem has 24-bit, 1024- by 1280-pixel color display memory and four overlay planes (one reserved, three usable); and it comes with an HP Starbase Graphics Library, upon which the HP version of the CASS display is implemented. The system has a mouse and a 32-key box which are essential for a friendly interactive user interface software design.

Software Environment

The prototype CASS software was written as a part of the PEDITOR system for ease of development and testing (Cheng *et al.*, 1989; 1991). The CASS programs have been placed in the PEDITOR menu structure under the menu selection of "Area Frame Development." There are 20 Pascal-based functional programs specifically written as a prototype for the CASS (Cheng *et al.*, 1989). The programs can be categorized into three groups: (1) displaying or manipulating TM and DLG data, (2) creating and editing polygons, and (3) system or file management. Because data files generated by CASS are in PEDITOR format, they can be read and processed by other PEDITOR functions designed to produce crop estimates. In addition, a file, called "_statefile," serves as a CASS program-to-program communication file to retain pertinent CASS processing information, including image and screen coordinates of the display, type of file, latitude and longitude of the scene, and the pixel size of the image.

CASS PROCEDURES

CASS is accessed through the "Area Frame Development" menu selection of the PEDITOR system. The general work flow of boundary delineation involves displaying a TM image, overlaying it with DLG data, delineating and sampling PSUs and segments, editing, and then saving the results. The following is a description of the basic steps involved in the CASS procedures.

Boundary Delineation

Display TM Image. The "Load and Display an Image" or *rtdisp* program is invoked, reads three of seven TM channel windows from a disk file, and displays the superimposed red, green, and blue (RGB) images onto the system color display screen in a user-selected order. The display window is 1024 by 1280 pixels, with an 8-bit color depth for each of the RGB images. The program supports TM data in three different formats: band sequential (BSQ), band interleaved by line (BIL), or pixel interleaved by line (PIL). The program "Mapping Function" or *mapima* adjusts the range of brightness for each of the RGB images in order to enhance certain land-use features. A subfunction of *rtdisp* redisplays a new window of the image with the same RGB parameters that were set in the previous display, and enters new window center coordinates into the communication file.

In operation, the operator displays a 1024- by 1280-pixel window of the TM image for a county, starting with the northeastern corner of the county. He/she then adjusts the color to enhance vegetation areas, agricultural or cultivated areas, and nonagricultural areas. Each enhancement can be saved in a color table file for future use. After completing one 1024- by 1280-pixel displayed area, the operator places the cursor on the edge of the display window and proceeds to display an adjacent area. The new cursor location is marked as the center of the new display window, and the process continues until the entire county is digitized.

Register DLG Data. The DLG data are used as a reference for the subsequent polygon digitization process in which the polygons (or PSUs) are delineated using known features such as roads or rivers. A set of programs prepares the DLG data for display of the polygons. The "Scan DLG Tape" scans each file on a USGS DLG tape and stores the tape's attribute codes into a file. The "Digital Line Graph Tape Read" program creates disk files, in PEDITOR format, of those DLG files on the USGS tape that are desired by the user. The "Add DLG Files" program reads selected PEDITOR-format DLG files and combines them to create a single larger file. The "DLG File Display" program then reads the DLG file and displays the DLG lines for the requested attributes, such as roads or water features, onto the overlay planes of the display. The "Register DLG to Image" allows users to more precisely register DLG data to any target image using a set of control points. The least-squares transformation generated is subsequently used to adjust and overlay the registered DLG lines onto the image.

PSU Delineation. The "Polygon Functions" or poly program allows the user to digitize or create an arbitrarily shaped polygon, and to modify or edit existing polygons. The function currently uses a two-button cursor on a digitizing tablet. The operator performs the delineation by typing in an identifying stratum and PSU number, delineating the area of interest, and closing the polygon. As the polygon is closed, the program annotates the longest edge of the polygon with a stratum-PSU label. The area in acres (converted to square kilometres for this report) and square miles is automatically computed based on the scale factor listed in the active communication file. The boundaries and other associated information are written into a polygon file for future processing by subsequent CASS or PEDITOR software programs.

On the image display, the operator delineates PSUs, or polygons of uniform land-use and land-cover type bounded by permanent features, according to predetermined strata definitions. He/she uses the polygon function on a second overlay plane with the cursor and the digitizing tablet while the DLG roads or rivers are displayed on the first overly plane as an aid. The periodic magnification of a selected area by a factor of 2, 4, 8, or 16, using a button-controlled zoom function, aids the operator in the definition of the PSU boundaries. In addition, the DLG data on the overlay plane can be turned on and off independently to allow the operator to see permanent features, such as roads or rivers, which are sometimes obscured by the DLG lines. The statistical data on crop types and acreages of the county in question also aids the operator in identifying a PSU and in digitizing the boundaries properly. Each polygon, or PSU, is labeled with its strata code as the prefix, and with a sequential number unique within the county as the suffix. For example, 11-23 means that this PSU is categorized as Stratum 11, "75% or more cultivated land," (Table 1) and is the 23rd PSU delineated in the county. The label and a calculated coverage area in acres and square miles (converted to square kilometres for this report) are displayed on the operator's text monitor and written into a computer file, called an area file, for future use.

Checking and Editing. A set of subcommands within the *poly* function allows the operator to identify any ill-defined polygons, that is, those that overlap other polygons or those with gaps between themselves and adjacent polygons. The subcommands also allow the operator to correct polygon problems by splitting polygons, moving or adding node points on the boundaries, and deleting entire polygons.

After completing the polygon delineation for a county, the operator redisplays the PSU boundaries on the screen and uses the Polygon Functions to edit those boundaries. Polygons that overlap or have gaps between themselves and other polygons are corrected. The output file of PSU boundaries is retained as a permanent record of the *"area sampling frame."* Editing of previously defined polygons is necessary because of changes in land-use patterns. With CASS, the update can be done at any time using the polygon delineation software.

Sample Allocation and Selection

A sample selection procedure is used by NASS-AFS to allocate and select PSUs and segments (Cotter and Nealon, 1987). For most of the agriculturally oriented states, an equal-probability selection and a land-use and land-cover types based method is used. First, one or two PSUs are randomly selected within each stratum or land-use and land-cover type. The selected PSU is divided into a predetermined number of segments based on its area coverage. Then the sample segments are randomly selected, with equal probability, from this PSU (Cotter and Nealon, 1987). The CASS function does not include the PSU selection; however, an attempt was made to simulate the delineation and random selection of segments.

Segment Delineation. A CASS prototype function called "Segment Delineation," or *sgpoly,* was recently implemented to delineate segments from each selected PSU. The function uses the capability of "Polygon Functions" that includes delineating segments within a selected PSU. Segment delineation is also based on the land-use and land-cover type, and the size of the coverage area, e.g., 1 square mile. As in the polygon function, the coverage area of a segment is immediately computed upon closing the polygon boundaries. The operator may choose to modify the boundaries based on the area reported. A 5-digit identification number is then assigned to each of the segments selected for the field observations.

CASS Products

The CASS products for each county consist of a polygon file containing the PSU boundaries and a list of the labels and areas of the PSUs.

Polygon File. The polygon file is then used to further divide selected PSUs into segments as briefly described in the previous section, but was not demonstrated in the example. More detailed information will be collected during field visits for the segments in the selected PSUs.

PSU List. The tabular list of PSUs displays the acreage of each delineated PSU. A subcommand of *poly* automatically computes the area in square miles and in acres when the polygon is closed and labeled.

AN EXAMPLE

The CASS procedures, except for segment delineation, were applied to each of three counties in Missouri as an evaluation. PSU polygonal data file products, including a PEDITOR-format file of PSU boundaries and a listing of area statistics, were generated for Livingston, Linn, and Macon counties (Figure 1). These three counties fall into the nominal coverage area of three TM scenes: Path 25 Row 32 (P25R32), P26R32, and P26R33. Two quadrangles (quarter scenes, or quads) of 1987 TM imagery were used for this evaluation, and two DLG data categories, transportation and hydrography, for the corresponding quads were acquired (Figure 1).

There are 85, 88, and 107 PSUs delineated using CASS for Livingston, Linn, and Macon Counties (Table 2, Plate 1) with areas of 1,371, 1,583, and 1,981 km², respectively. The proportions of agricultural areas in these counties were 96.9 percent, 99.0 percent, and 84.8 percent for Livingston, Linn, and Macon, respectively. Based on NASS' PSU and strata definition (Table 1), there are eight strata categories for all three counties identified using CASS. There are five, five, and eight strata for Livingston, Linn, and Macon counties, respectively. For Livingston and Linn, the "75% or more cultivated," or stratum 11, was the predominant land-use/land-cover type. The "15-49% cultivated," or stratum 20, was the predominant type for Macon county. Area files for all counties are used to randomly select PSUs for further breakdown into segments.

DISCUSSION

Experience has demonstrated that the time required to construct area frames using these procedures was reduced substantially compared to the manual technique currently used. For example, the construction time for an operator to complete area frames for Macon County was 2.5 days using the CASS procedure compared to 15 days using the conventional manual method. The PSU delineation, performed manually by transferring boundaries between maps and image media, can be done electronically. The CASS procedure also eliminates the time and labor of digitizing for area estimates.

Because the PSU boundaries can be easily edited electronically

TABLE 1. USDA PRIMARY SAMPLING UNIT (PSU) AND STRATA DEFINITION

Stratum	Definition	Area, km² Minimum	Desired	Maximum
11	75% or more cultivated	2.6	15.6 to 20.8	31.1
12	50-74% cultivated	2.6	15.6 to 20.8	31.1
20	15-49% cultivated	2.6	15.6 to 20.8	31.1
31	Ag-urban, less than 15% cultivated	0.6	2.6 to 5.2	7.8
32	Residential-commercial	0.3	1.3 to 2.6	2.6
40	Range and pasture, less than 15%	5.2	31.1 to 41.4	62.2
50	Nonagricultural	2.6	None	None
62	Water	2.6	None	None

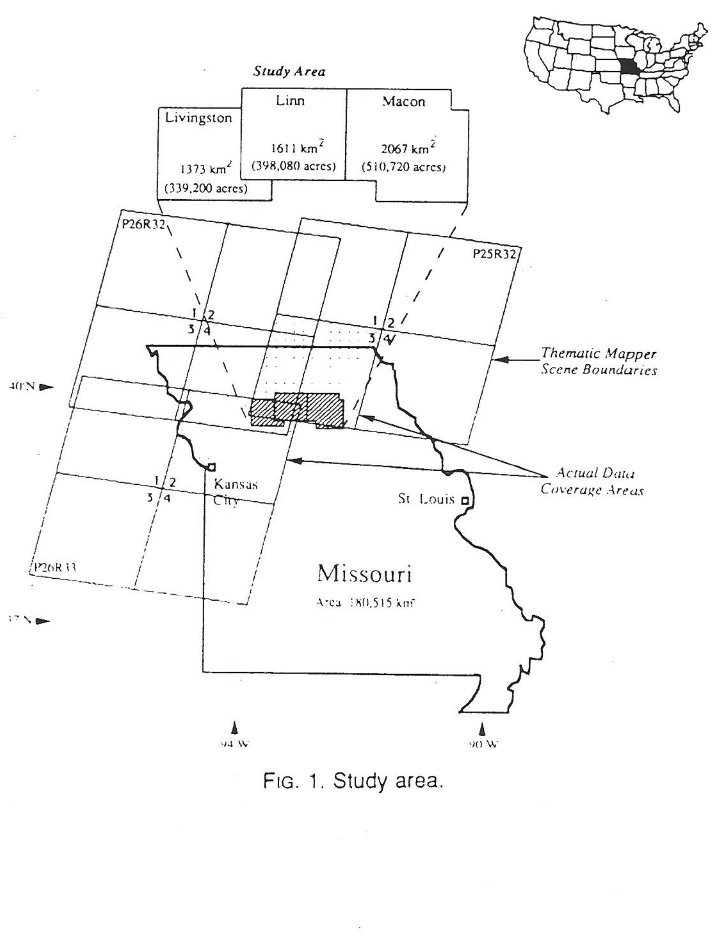

FIG. 1. Study area.

TABLE 2. LAND-USE PROFILE FOR TEST COUNTIES IN MISSOURI. THIS TABLE WAS COMPILED BY THE STRATA AND AREA COVERAGES (KM² IN PARENTHESES) GENERATED BY CASS, AND IS SHOWN HERE AS AN EXAMPLE FOR CASS' PRODUCT FROM WHICH AGRICULTURAL STATISTICS OF A STATE WILL BE DEVELOPED.

	Livingston		Linn		Macon		
Stratum	No.PSU[a]	Area, acres (km^2)	No.PSU[a]	Area, acres (km^2)	No.PSU[a]	Area, acres (km^2)	Total, acres (km^2)
11	39	150,860.0 (610.4)	43	210,304.3 (851.1)	19	86,154.8 (348.6)	447,319.1 (1,810.2)
12	23	81,715.7 (330.5)	34	148,575.4 (601.1)	16	58,168.7 (235.2)	288,459.8 (1,167.3)
20	18	96,051.6 (388.5)	8	28,457.2 (115.0)	52	270,883.1 (1,096.1)	395,391.9 (1,599.8)
31	1	2,031.8 (8.0)	2	1,844.8 (7.3)	2	838.1 (3.4)	4,714.7 (18.9)
32		–		–	1	199.8 (0.8)	199.8 (0.8)
40	4	8,198.8 (33.2)		–	15	70,008.8 (283.1)	78,207.6 (316.2)
50		–	1	1,919.8 (7.5)	1	1,141.3 (4.4)	3,061.1 (12.2)
62		–		–	1	2,085.5 (8.3)	2,085.5 (8.3)
Total PSUs	85		88		107		
Total acres		338,858.1		391,101.8		489,480.5	1,219,439.5
Total km^2		(1,371.1)		(1,582.5)		(1,980.8)	(4,934.7)
Agric. acres[b]		328,627.3		387,336.9		415,206.6	1,131,170.8
Agric. km^2		(1,329.7)		(1,567.3)		(1,680.0)	(4,577.0)
							Average
Agric./total acres ratio (%)		96.9		99.0		84.8	92.8
Avg PSU acres		4,458.6		4,444.3		4,574.5	4,499.8
Avg PSU km^2		17.9		17.9		18.4	18.1

[a] Number of PSUs (average PSU size, km^2).
[b] Agricultural or cultivated area.

on a graphics workstation, the area frames can be updated with greater frequency. The current average update cycle is about 5 years. This is an important benefit because of changes in land-use patterns. Once a state *"area sampling frame"* is established, periodic updates can be done for areas in which the land use or land cover has changed by acquiring the digital data for those areas and overlaying the existing boundaries of PSUs on a new image, with DLG lines also displayed. The boundaries can then be edited electronically, and the acreages can be recomputed. The CASS procedure is thus valuable not only in the construction phase, but also in the maintenance phase.

Accuracy of the assessments may be enhanced. The capability of CASS to view satellite imagery and map-attribute data at the same time may enable the operator to more accurately assign the PSUs to strata. Increased accuracy may also result from the avoidance of the step of transferring boundaries, better presentation of the ground cover electronically, and automatic coverage calculation.

A graphical user interface (GUI) can streamline CASS into an operational procedure, and increase productivity. This can be accomplished with a Macintosh-style screen display, which includes features such as pull-down menus, sliders, and mouse-driven functions. GUI designing tools were investigated for this objective, also.

The CASS effort is a pilot design to automate the stratification and sampling-unit delineation portion of the NASS-AFS area sampling frame techniques. The basic steps of the process have not changed; they are merely replaced by computerized procedures. To take full advantage of the current computer technology, a more fundamental design is needed, which will greatly improve the productivity of the CASS process. The subsequent stage of the area frame techniques, the sampling procedures, will also be considered for a future automated procedure.

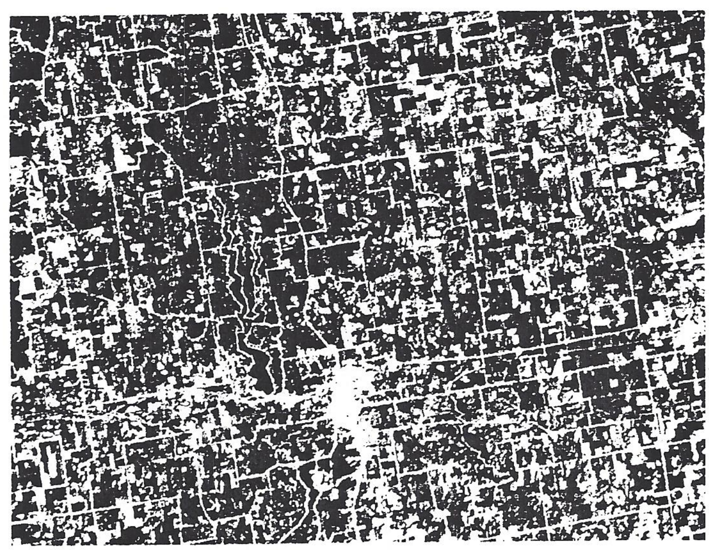

PLATE 1. DLG and PSU boundaries. A 1024- by 1280-pixel window of a TM image of Macon County was displayed using channels 2, 3, and 4 in blue, green, and red, respectively. Therefore, the highly vegetated areas appear to be dark red. Using the two graphic overlay planes, DLG transportation systems were overlaid in blue, and the boundaries of the PSUs were overlaid with strata labels in white.

A scenario for a future CASS procedure project begins with the classification of a satellite sensor image to produce a class map/image. This class image is processed through ground-truth verification procedures to combine spectral classes into land-use types or PSUs. The image is then processed using edge-finding, edge-tracing, vector-conversion, and other algorithms to produce boundaries for all the PSUs. A quality control procedure is implemented to verify the accuracy of the automated boundary-delineation process. This process potentially reduces the cost of labor-intensive photographic interpretation that is now a part of the stratification procedures. It can also potentially produce high-quality, vectorized boundaries for PSUs. These boundaries can later be used for PSU selection or as a basis of further sampling procedures.

ACKNOWLEDGMENTS

The CASS project was supported by NASS, and the work of NASS staff members J. Cotter, D. Atkinson, C. Mazur, Y. Dodson, and J. Gallegos on this project is greatly appreciated. This project was also funded through the Earth Observation Commercial Applications Program of NASA Code SE (NASA NRA-87-060, UPN 672-90-20-03).

REFERENCES

Cheng, T. D., G. L. Angelici, R. E. Slye, and M. Ma, 1989. *Computer-Aided Boundary Delineation of Agricultural Lands.* NASA TM–102243. NASA Ames Research Center, Moffett Field, California. 23 p.

———, , 1991. Compiling and Editing Agricultural Strata Boundaries with Remotely Sensed Imagery and Map Attribute Data Using Graphics Workstations. *Proceedings of IGARSS'91,* University of Technology, Espoo, Finland. pp. 2457-2460.

Cotter, J. J., and J. Nealon, 1987. *Area Frame Design for Agricultural Surveys.* National Agricultural Statistics Service, U. S. Department of Agriculture, Washington, D.C. 67 p.

NASS, 1990. Planned Technology Demonstrations: Computer-Assisted Stratification and Sampling. *Proceedings of the 1990 National Conference for the National Agricultural Statistics Service: Prospect of the Future.* 2-5 April, San Francisco, California, (Unpaginated)

(Received 11 February 1991; revised and accepted 15 April 1992)