Spatial Resolution Requirements for Automated Cartographic Road Extraction

Susan Benjamin and *Leonard Gaydos*
U. S. Geological Survey, 345 Middlefield Road, Menlo Park, CA 94025

ABSTRACT: Ground resolution requirements for detection and extraction of road locations in a digitized large-scale photographic database were investigated. A color infrared photograph of Sunnyvale, California was scanned, registered to a map grid, and spatially degraded to 1- to 5-metre resolution pixels. Road locations in each data set were extracted using a combination of image processing and CAD programs. These locations were compared to a photointerpretation of road locations to determine a preferred pixel size for the extraction method. Based on road pixel omission error computations, a 3-metre pixel resolution appears to be the best choice for this extraction method.

INTRODUCTION

LARGE-SCALE, STANDARD SERIES MAPS, such as the 1:24,000-scale, 7.5-minute series published by the U.S. Geological Survey, need to be updated frequently in order to reflect cultural and landscape changes. Map revision is required when features are constructed or removed, or when routes or sites are relocated. The revision process traditionally employs photointerpretation of aerial photographs of the region. Automation of many of the revision processes would shorten the time between large-scale map editions. Automated techniques for printed map revision could also be employed to revise the Digital Line Graph (DLG) data produced by the Survey of features on 1:24,000-scale quadrangles (U.S. Geological Survey, 1986).

This study was therefore designed to determine the best pixel spatial resolution to use for the identification and extraction of roads from digitized aerial photographs (Benjamin and Gaydos, 1984). The extraction technique has shown promise for use in revising roads on large-scale maps. Determining an optimal spatial resolution for any digital mapping procedure is essential in order to minimize data processing cost, volume, and time while maintaining mapping quality. Selection of a pixel size that reduces data processing, while still retaining sufficient map information, will lead to the best use of the extraction technique.

RESEARCH BACKGROUND

Digitized photographs are under investigation as a direct source for the feature information included on large-scale topographic quadrangles. The digital photo data set can be combined with scanned map feature layers in a GIS-like digital photorevision process (Usery and Welch, 1989) or serve as a data source for an automated feature extraction process. Two approaches have been cited for extraction of linear features in digitized photographs. The first employs knowledge of the region or type of features in the photograph to extract an information layer. Fischler *et al.* (1981) and Fischler and Wolf (1983) scanned high contrast black-and-white photos, then used a variety of line following algorithms to trace the structure of the road network. Algorithm selection was based on *a priori* knowledge of the road width, direction, approximate location, terrain type, scene elevations, and degree to which the road network was obscured. Pries and Schowengerdt (1987) developed a partially automated system for the detection of roads and trails on digitized black-and-white photographs. An operator selects control nodes along a linear feature, the paths between nodes are examined for contrast variations, and automated decisions are made as to the presence of a feature. Nagao and Matsuyama (1980) developed a knowledge-based system for the identification of all land cover features in data sets formed from high resolution digitized color infrared photos. Knowledge of the physical properties of linear features, such as roads, rivers, and railroads (including their linearity), was used to separate them from the other features (buildings, cropland, water, etc.) in the frame.

The second major approach to extraction of linear features from scanned aerial photographs is to use multiband digital data sets developed from color infrared photographs, reduce the multispectral information through digital image classification to a single information band, then refine the band further to complete the extraction process. Scarpace and Quirk (1980) defined land cover types that included linear features listed above by classifying and interpreting digitized color infrared photos of wetland areas. Townsend (1981) scanned large-scale color infrared photograph stereopairs to produce stereo-image data sets with fine spatial resolution (0.75-metre pixels). The features within the images were identified through classification, logical smoothing, and region clustering to extract the boundaries of scene objects. Photogrammetric methods were then used to map positions of the extracted boundaries accurately from the digital image stereopairs. Benjamin and Gaydos (1984) classified scanned color infrared photos of urban and suburban areas to isolate land cover and road networks for updating large-scale topographic guadrangle maps. The road network information contained in the classifications was refined with automated editing techniques to isolate and trace networks within the frame

Techniques for assessing optimal spatial resolutions for applications of remotely sensed data have been established. Woodcock and Strahler (1987) developed a method to evaluate spatial resolutions used in all types of digital imagery. A measure of local variance at different spatial resolutions was computed and plotted, and the location and height of the peak of the curve was used to select an optimal spatial resolution for the scene. An optimal resolution of 15 to 20 metres was determined for a scanned color infrared photograph of suburban southern California. Markham and Townshend (1981) evaluated the effect of decreasing spatial resolution on the ability to classify land cover types on a per pixel basis. Airborne multispectral scanner images were degraded from 5-metre resolution pixels to 40-metre resolution pixels, each data set was classified, and each classification was evaluated within the classifier training pixel areas. In an urban-suburban test site, land cover classification accuracy (percent of area at each resolution that was properly classified) increased as spatial resolution was degraded.

STUDY OBJECTIVES

The main objective of this study was to identify the spatial resolution required to effectively extract road locations with the

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 56, No. 1, January 1990, pp. 93–100.

0099-1112/90/5601–93$02.25/0

linear feature extraction technique developed by Benjamin and Gaydos (1984). This procedure involves scanning a color infrared aerial photograph, clustering and classifying the multispectral data set, interpreting the classification into feature and material classes, refining the linear features (roads) with raster CAD (computer-aided design) programs to enhance road network continuity, and finally thinning the refined network to its center path and converting that to a vector representation. By applying this extraction technique to a scan-digitized aerial photograph at several spatial resolutions, the best spatial resolution was identified.

A secondary objective was to determine the effect of spectral cluster signature extension at the different resolutions. The technique relies on spectral signatures of land cover to define features that are extracted. During this study, additional areas beyond that used to develop the feature signatures were analyzed, and effectiveness of the technique was evaluated within and outside of the region.

TEST SITE

The test site is in Sunnyvale, California, (Figure 1) and is entirely contained on the Cupertino, California, 1:24,000-scale topographic quadrangle (U.S. Geological Survey, 1961). The site was selected for its wide variety of road types, paving materials, states of weathering, intersection patterns, hierarchies, and edge delineators. Within the area are many other features with reflectance characteristics similar to the roads, such as rooftops, parking lots, sidewalks, and dark vegetation, that needed to be distinguished from the road network. Sunnyvale's structural morphology is typical of suburban neighborhoods in California with frequent new road construction and changes. The test site, therefore, retains the characteristics of an area that might require numerous map revisions. Dominating vegetative cover in suburban neighborhoods may actually improve the separability of linear road features providing increased contrast with the paved surfaces.

During this study, the spatial extent of the original test site was expanded in all directions to add samples of features poorly

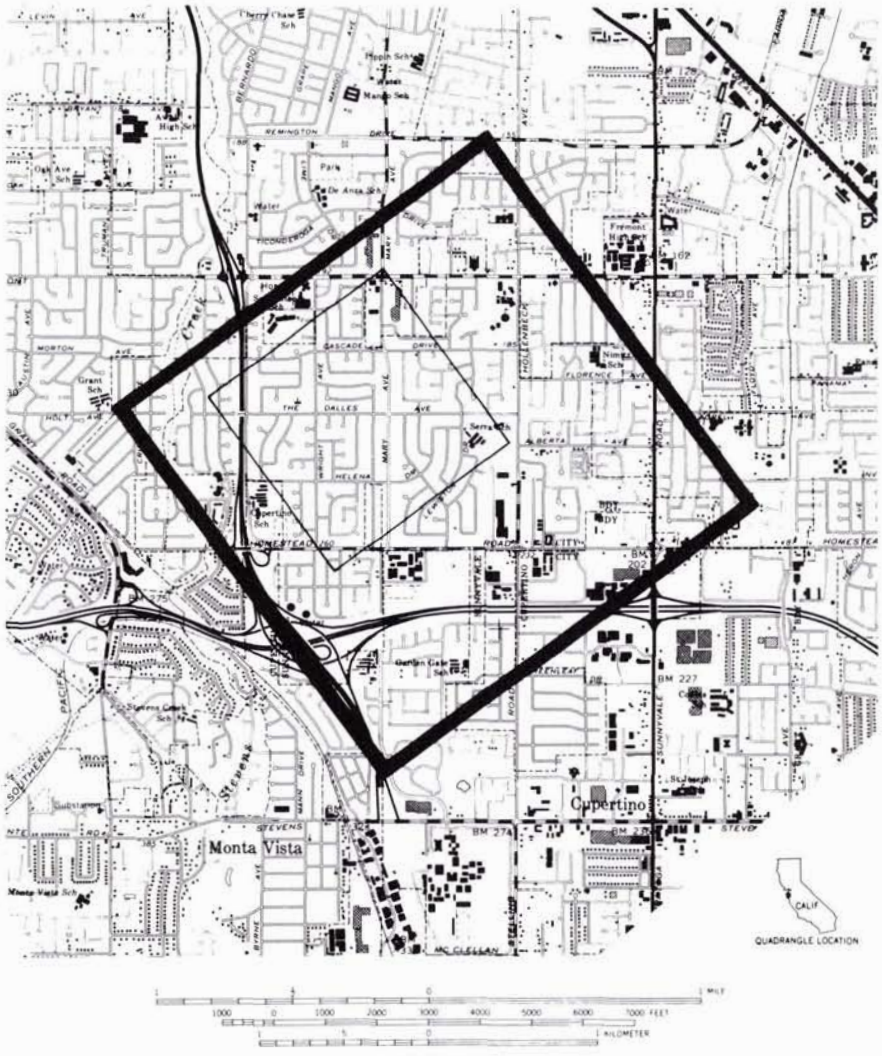

FIG. 1. Roads in the Sunnyvale study site as mapped on the Cupertino 1:24,000-scale quadrangle. Inner boundary is the original test site; outer boundary is the test site used in this study.

represented in the original test site. These features included freeway area, large institutions (hospital and fire station), and riparian vegetation. This expansion had the added effect of permitting the evaluation of extending the spectral signatures into areas not included when the signatures were computed. Figure 1 shows both the original and expanded test site boundaries.

Within the expanded test site, land use type is primarily suburban single-family residential housing with small commercial centers, as well as institutions such as schools, churches, hospitals, and parks. Buildings are located in a matrix of vegetation consisting of houses, lawns, gardens, trees, and bushes. The road network is well defined with small residential streets feeding into larger arterials. Road connections are not always at right angles; a significant number of the residential streets are curved or arced. Sections of two multilane freeways that serve the region were included as well. The complete test site covers an area approximately 2.5 kilometres on a side; the region used to develop cluster statistics is a 1.3 by 1.3-kilometre section near the center of the site (Figure 1).

DATA ACQUISITION

Aerial photography was acquired on 28 May 1982 by a U2 aircraft flying at 50,000 feet using an HR-732 camera system with a 24-inch focal length lens (NASA, 1982). The film was standard color-infrared type SO-131 with emulsions sensitive to green, red, and infrared radiation, and a yellow filter was used to absorb blue band radiation. The center of one photograph was scanned with an Optronics microdensitometer, and reflectances were recorded in each emulsion at 50-micrometre intervals to form a three-band data set with 1.25 metre spatial resolution.

From the digitized raster array, a large rectangular "window" was selected for the analysis. This window contained the area of the test site used to develop the extraction technique and extended beyond it on all sides. The original test site window contained 1,024 by 1,024 pixels while the test site window for this study contained 2,048 by 2,048, or four times as many pixels.

FEATURE EXTRACTION PROCESS

Before applying the extraction technique, the scanned data were registered to a map grid to provide the opportunity for accuracy evaluation through direct digital comparison with the topographic quadrangle. The digitized data were planimetrically registered and corrected to a Universal Transverse Mercator (UTM) map grid with a 1-metre spatial resolution. The process used to register the data set was developed for the geometric correction and registration of satellite imagery to map orientation and projections (Ford and Zanelli, 1985). Twelve points were selected throughout the test site to provide a good distribution of registration control. The corners of large buildings that could be accurately identified on both image and map were used. Large buildings were the only features located on the quandrangle and the photo which retained their correct outlines. (Road intersections mapped as right angle intersections are curved on the photo.) The quadrangle itself was the largest-scale map available of the area constructed with sufficient horizontal and vertical control to select ground control points. This limited the number of points that could be precisely located. These ground control points were then used to establish a uniform 1-metre UTM grid for the image fit. After least-squares analysis and point deletion, seven points remained to establish the second-degree transformation polynomial, with an RMSE of 0.91 pixels horizontal and 0.99 pixels vertical. The three bands of digitized photo data were then fit to the UTM grid according to the transformation polynomial.

The registered photo data set (1-metre resolution) was then degraded to coarser spatial resolutions: 2-, 3-, 4-, and 5-metre pixels. The data degradation process was modeled after that used by Wrigley *et al.*, (1984) to simulate the spatial characteristics of Landsat MSS and TM images using airborne Thematic Mapper Simulator images and is similar to that used by Woodcock and Strahler (1987) to examine the optimal spatial resolutions for digital analysis of different types of landscapes. The 1-metre data set was first filtered with a square (boxcar) filter of the same size as the desired output spatial resolution (for example, a 4- by 4-pixel filter was used when generating the 4-metre data set). This process averaged the pixel values within the filter and assigned the average to the center (or nearest center) pixel. After filtering, the data were reduced in size by the same factor (for example, four times smaller in the case of the 4-metre data set) using a nearest neighbor reduction program.

The spectral clusters used to classify the data sets were developed in the original test site (Benjamin and Gaydos, 1984) using 1.25-metre resolution pixels. The data values in the original test site were examined and iteratively divided into 48 groups using an algorithm that maximizes the separation between the groups (clusters). Statistics were computed for each cluster (mean, variance, and covariance matrix) to describe it. These same cluster statistics were used to classify all spatial resolution data sets in this study (rather than developing new clusters for each data set) so that extraction results at the different spatial resolutions would be comparable.

The classification process, using a maximum likelihood decision rule (Jensen, 1986), reduced each multispectral image to a single band (for each spatial resolution) with data values ranging from 0 to 48. Each of the five classified images was then interpreted for land cover type designation. Following land cover interpretation, the numeric classes were grouped into five feature and material classes (Table 1). Land cover classes remained constant for all spatial resolutions, although spectral classes differed slightly due to the pixel averaging that occurred during the data registration processes. Finally, the classifications were smoothed with a 3 by 3 nearest neighbor filter to remove single occurrence pixels and improve class connectivity (Jensen, 1986).

During the data refinement phase, the road network was en-

TABLE 1. FEATURE AND MATERIAL CLASSES

Class label	Constituent elements
Roads	Asphalt roads, primarily those in the single-family residential areas but also including arterials and some portions of freeways; also included are dark, weathered concrete roads and patches of pixels located on dark rooftops and in parking lots.
Dark surface	Manmade surfaces, generally lighter in tone than those assigned to the roads class; included are wood shake roofs, weathered concrete and asphalt, and definable building and object shadows.
Light surface	Manmade and natural surfaces, generally lighter in tone than those assigned to the dark surface class; included are concrete (in sidewalks, driveways, and freeways), very light-toned weathered asphalt, composite construction roofs, gravel and dirt roads and patches, and disturbed soil areas.
Vegetation	Any section of vegetation with a significant infrared reflectance; included are trees, shrubs, lawns, park grass, and definable vegetation shadows.
Background	Areas outside of the data subwindow.

hanced through editing procedures and thinned to its center path. The refinement processes were carried out on a USGS CAD raster/vector processing system used for map production. These methods were derived from procedures and programs used by the Survey to edit digitized transportation and hydrographic linework from 1:100,000-scale quadrangles in order to create DLGs (Callahan and Broome, 1984). The data set for each spatial resolution was analyzed with the same set of procedures to produce final products that could be compared directly.

The first step in the data refinement process was to remove islands of pixels of one class that were entirely surrounded by pixels of other classes and reassign them to a new class of edited pixels. This process resulted in a more continuous road network with better edge definition. Manual examination and editing of the images was then undertaken to remove any remaining road class pixel connections between the road network and rooftop areas. The connections between the two feature types (road and rooftop) were "painted" out by substituting pixels of the edited class for road class pixels. Editing decisions were based on object (or class) shape, location, and size.

The final refinement step created DLG-like raster representations of the road networks. The fully edited road pixels in each data set were thinned to their center paths. These thinned line segments of raster pixels were converted to vector representations and then reconverted to a raster file with fewer spikes and smoother contours. These final lines were judged to be the roads extracted at each spatial resolution (Figure 2).

Problems were noted in editing processes with the coarser resolution data sets (4 and 5 metres). In the eastern portion (right) of the images, the majority of pixels were classified as roads and separations were less evident. Automated techniques that were effective in other portions of the same image were not useful in the eastern sections. In addition, manual editing in these areas was virtually impossible because of difficulties in visually separating the correctly classified road pixels from the misclassified pixels assigned to the same class.

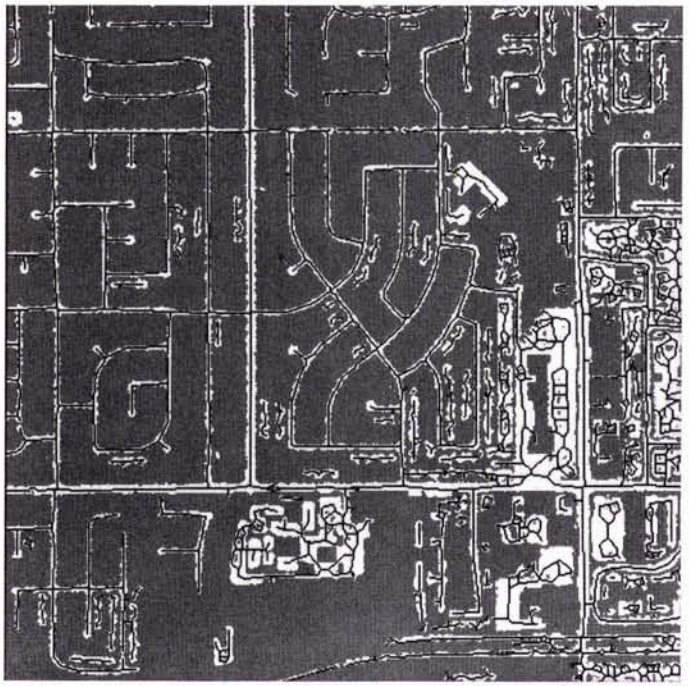

FIG. 2. Subsection of the 3-metre road center path data set, a good example of the final extraction product (extracted road center path pixels overlaid on edited road class pixels).

COMPUTATION OF OMISSION ERRORS

Errors of omission were computed to determine the effectiveness of the extraction technique with differing sizes of pixels. These errors indicate the length of road in place when the photo was acquired that was not included in the extraction process. Omission errors are critical because extra roads could be erased, but missing roads would have to be plotted manually if the extraction product were used to make a map or a DLG.

An attempt was made to compute extraction accuracy through direct digital comparison to roads in the quadrangle. However, universal alignment of a scanned map separate and the scanned photo data was not possible due to internal distortions in the photograph caused by terrain variation and aircraft tilt. A ground-truth data set used to assess the accuracy of the extracted road networks had to take into account the slight positional distortions contained in the photograph that was scanned. Therefore, the data set created in the registration process was used to generate ground-truth information through photointerpretation of an enlarged print of the registered 1-metre, color-infrared data set. Road locations on the print were interpreted, and edges of the roadways were identified, transferred to a mylar overlay, and compared to the quadrangle map. Features used to register the scanned photo data to the 1-metre UTM grid were also located and marked on the overlay.

This overlay was then scanned to create a raster digital image of the photointerpreted road locations in the test site. The image was registered to the 1-meter UTM grid with the same control points used to register the multispectral photo data set, then converted to a binary raster image containing only the outlines of the roads and a background class. The pixels within the outlines were changed to a different value, forming a registered ground truth image with correct road widths that would accurately align with the digitally extracted road networks.

Because the effectiveness of extraction was to be compared over the range of spatial resolutions, a ground-truth image for each spatial resolution was needed. The 1-metre resolution ground-truth image was spatially degraded to 2-, 3-, 4-, and 5-metre data sets with the filtering method used to degrade the 1-metre scanned photo data. The final step in the ground-truth data set preparation was to thin the filled casings down to their center paths to create DLG-like raster representations of the road networks.

To determine the omission errors, the ground truth was compared to both the unedited interpreted classifications and to the classified data sets after manual and automated editing had been performed. When the ground-truth pixels were overlaid on the extracted roads, the pixel values in the resulting data sets indicated whether the feature had been extracted (ground-truth center path with road class) or had been omitted (ground-truth center path with a nonroad class). The pixel values were tallied for each omission error data set and converted to a percentage of the total ground-truth road center path pixels at each spatial resolution to compare between resolutions. As a final step, the spectral cluster generation region (original test site, Figure 1) was isolated within each omission error image and omission error percentages were computed for just that region at each spatial resolution.

OMISSION ERROR RESULTS

Omission error analysis produced four sets of omission error statistics: (1) for the entire test site, before editing was performed; (2) for the spectral cluster generation region only, before editing was performed; (3) for the entire test site, after editing was completed; and (4) for the spectral cluster generation region only, after editing was completed (Table 2). Error Set 3 gives the most comprehensive assessment of the effectiveness of the extraction technique because it is based upon

TABLE 2. ERRORS OF OMISSION (PERCENTAGE OF PHOTOINTERPRETED ROAD CENTER PATH PIXELS)

Error set	Spatial resolution (metres)	Ground-truth road center pixels	Omission error pixels	Omission error (percentage)
(1)	1	80,002	24,879	31.098
Full test	2	39,803	9,763	24.528
site	3	26,423	5,509	20.849
before	4	19,828	5,079	25.615
editing	5	15,641	4,206	26.891
(2)	1	20,711	6,835	33.002
Original	2	10,246	2,199	21.462
test site	3	6,770	1,164	17.193
before	4	5,073	1,341	26.434
editing	5	4,023	1,253	31.146
(3)	1	78,984	18,692	23.666
Full test	2	39,655	7,152	18.036
site	3	26,315	4,619	17.553
after	4	19,747	4,366	22.109
editing	5	15,603	3,546	22.726
(4)	1	20,703	4,072	19.669
Original	2	10,246	984	9.604
test site	3	6,770	682	10.074
after	4	5,073	903	17.800
editing	5	4,023	892	22.713

TABLE 3. PIXELS REQUIRED FOR FEATURE DEFINITION

Data set spatial resolution (metres)	Road Widths		
	Freeway (40 metres) (pixels)	Arterial (20 metres) (pixels)	Residential (7 to 10 metres) (pixels)
1	40	20	7 - 10
2	20	10	3.5 - 5
3	13.3	6.6	2.3 - 3.3
4	10	5	1.8 - 2.5
5	8	4	1.4 - 2

the entire test site, employing the full extraction process. Comparing Sets 1 to 2 and 3 to 4 demonstrates the effect of cluster signature extension on the extraction process because the spectral clusters were "extended" beyond the boundaries of the region used to define them in order to classify additional areas. Comparing Sets 1 to 3 and 2 to 4 demonstrates the effect of the editing processes on the extraction technique.

OMISSION ERROR ANALYSIS

The 3-metre data set exhibited the lowest omission errors; the only variation was in error Set 4 where errors for the 2-metre data set were slightly lower. The highest errors were found in the data sets with the largest and smallest pixels, with an absolute error maximum at about 33 percent for 1-metre pixels in the original test site before editing.

Three-metre resolution pixels result in a more accurate road extraction product for many reasons. One factor to consider is the size of the extraction objects. The accuracy evaluation method considers roads linear objects with width being the critical dimension. Road width ranges from 40 metres on freeways, to 20 metres on arterial roads (two lanes in each direction), to between 7 and 10 metres for smaller residential streets. Table 3 lists the number of pixels required to define these road widths at each spatial resolution.

Secondly, many of the roads in the test site are residential streets and three 3-metre pixels would span their width. The ground-truth road center was therefore located with a single pixel allowance for error on each side. In contrast, at most one or two pixels define residential road widths in the 4- and 5-metre data sets (and may be below the minimum sampling width to be resolved). At these coarser resolutions, if a road pixel was averaged to a more reflective digital value in the degradation process, then the pixel would be assigned to a nonroad class and be recorded as an error pixel when compared to the ground-truth data. This may be related to the Nyquist frequency in sampling theory (Bloomfield, 1976) that requires at least two samples to identify an object; in this case, the 3-metre data set is more likely to use at least two pixels to define a road than are the 4- or 5-metre data sets. It is also possible that, in the coarser resolution ground truth data set, the center path is slightly offset due to the road being contained in one or two pixels.

Another factor in favor of 3-metre resolution pixels is the elimination of extraneous elements (cars and trucks, vegetation in center dividers, vegetation and shadow overhangs at the side of the roads, crosswalks, etc.) within the roads. These elements would be detected with smaller pixels because the degradation process would begin to eliminate their effects in the coarser resolution pixels. Cars and trucks are "confusing" elements on residential streets and are generally smaller than 3 metres. Thus, in the 3-metre data set, the pixel covering a vehicle would also include a substantial portion of roadway, retaining low digital values in all bands, and therefore be assigned to the road class. In contrast, in the 1- or 2-metre data sets, the vehicle might be contained in a separate pixel that would be assigned to a nonroad class and later assessed as an error pixel.

Two other trends are noteworthy. First, omission error decrease overall as a result of the editing processes. The decrease can be attributed to the elimination of extraneous elements in the roads (as noted above). A major goal of the automated editing was to improve road network continuity by removing islands of nonroad class pixels contained within the defined roadway. These islands included the objects that were entirely surrounded by road class pixels, as well as the vegetation and shadow overhangs that were surrounded on three sides by road class pixels. Second, omission errors rise when examining the entire test site rather than just the original test site. This trend indicates that spectral cluster signature extension does have an effect on extraction accuracy. Errors may be expected to increase when moving beyond the region where the spectral clusters were defined because the features of interest (roads) may have different spectral signatures in those locations and the classification would not be as homogeneous.

COMPONENTS OF ERRORS

The components of the omission errors are the class assignments at pixel locations that are defined as road center paths in the ground-truth data set but not as road pixels in the photointerpreted data sets. These components have been computed for each error set, then converted to a percent of the total error pixels in the data set (Table 4 and Figure 3). No definitive component of the error is present in all error sets, although comparisons between the error sets indicate certain trends.

If error set 3 (full test site after editing) is regarded as the most comprehensive assessment of errors of the technique, then its values must be examined closely. The largest component of the omission error statistics are the pixels of the light surface class consisting mostly of light-toned concrete and rooftops. The concrete freeways are assigned to this class rather than to the darker toned road class. This component is the source of most of these errors. The second largest component of the error comes from the vegetation class — this results primarily from trees that hang over the road surface and interrupt the continuity of the network.

Within the unedited classification error sets (1 and 2) the dark surface class and the vegetation class are of equal importance

TABLE 4. COMPONENTS OF OMISSION ERRORS (PERCENTAGE OF TOTAL OMISSION ERROR PIXELS)

Error set	Spatial resolution (metres)	Land Cover Class of Error Pixels: Dark surface (%)	Light surface (%)	Vegetation (%)	Editing	Background (%)
(1)	1	24.74	47.12	27.19	.....	0.40
	2	30.31	44.00	25.16	.....	0.54
	3	25.62	46.85	27.08	.....	0.45
	4	23.71	40.89	35.06	.....	0.33
	5	31.95	35.66	32.19	.....	0.19
(2)	1	38.89	34.94	26.07	.....	.....
	2	56.25	23.23	20.51	.....	.....
	3	52.75	25.48	22.77	.....	.....
	4	40.94	23.64	35.42	.....	.....
	5	55.95	15.56	28.49	.....	.....
(3)	1	0.98	57.65	34.60	5.97	0.18
	2	2.39	56.31	33.22	7.69	0.34
	3	2.40	55.23	31.63	10.19	0.54
	4	3.76	46.68	40.38	8.89	0.30
	5	9.33	41.91	37.90	10.63	0.23
(4)	1	0.10	53.29	43.52	3.12	.....
	2	3.35	44.00	44.21	8.43	.....
	3	3.26	41.79	38.86	16.13	.....
	4	4.54	35.11	52.60	7.72	.....
	5	21.75	21.86	40.02	16.37	.....

in computing the omission errors. This class contains weathered concretes and medium-toned roof materials and is frequently found adjacent to repaved asphalt streets. The automated editing process aided in removing pixels of this class and, therefore, its contribution to the overall error decreased considerably in the edited classification error sets (3 and 4).

Error attributable to light surface class pixels diminish in their contribution to the overall error within the original test site error sets (2 and 4). This decrease is due to the fact that the majority of the freeway area falls outside of the region used to generate cluster statistics. Vegetation component errors reach their maximum in error set 4 due to the effects of the editing steps. If editing removes most of the misclassified dark surface pixels without disturbing vegetation class pixels, then, proportionately, the vegetation class contribution to the total error should rise.

The only general trend that emerges from error component analysis is that, as pixel size increases, the contributions of the various classes become more equal. This trend is seen in Figure 3 where the columns of the contributing classes are closest to being equal in height in the 5-metre data sets, particularly in the full test site error sets (1 and 3). This result is likely caused by the averaging effect of the spatial degradation process.

LOCATIONS OF ERRORS

Omission error pixel locations for the 3-metre data set are shown in Figure 4. As previously noted, overall, the largest component of omission errors at each resolution are pixels assigned to the light surface class. Most of these errors are located in the freeways (routes CA-85 and I-280) running through the western and southern sections of the test site. Several other streets are also involved in this error category, including a small cluster of concrete and weathered asphalt roads adjacent to a public high school in the south-central section of the test site. The other major region of omission errors is in the northern tip of the test site with errors contributed mainly by the confusion between dark vegetation and asphalt paving. A similar region of omission errors is centered on the western tip of the test site. Figure 4 indicates that the original test site is relatively error free (in the 3-metre data set) and that similar sections (with respect to omission errors) exist throughout the test site. This result suggests that cluster signatures can be successfully extended in the extraction process but not necessarily to all regions.

PROBLEMS NOTED

The omission error analysis technique falters as an accuracy measure in the eastern portion of the test site. This section contains more commercial land-use parcels than does the remainder of the test site. More pixels in the eastern section were misclassified as roads than anywhere else in the test site because the roof reflectances of the commercial buildings are similar to road reflectances in the original test site. As a direct consequence, during the omission error analysis, few omission errors were detected in this section. Every pixel of ground-truth road path would probably coincide with a pixel classified as road because most pixels were given that assignment. This type of error should be noted because the resultant extraction product is very poor. When large regions are misclassified as roads, the editing process becomes very complicated because the boundaries between roads and adjacent land covers are obscured. The thinning algorithms applied to the edited data sets anticipate a linear input; thus, when given a blocky pixel mass rather than a roughly linear mass, the output is a set of lines that range throughout the mass linking corners and protuberances. This kind of extraction product can be seen in the southeastern portions of Figure 2. Extraction accuracy in this section of the test site requires a different method of evaluation.

Commission error analysis of the extraction products would give numeric and graphic indications of the pixels in each data set that are defined as road pixels at the end of the extraction but are not actually at road locations. Unfortunately, commission error statistics can't be accurately computed for these data sets because misclassified road pixels were retained in the data sets as long as they were not connected to the main road network.

CONCLUSIONS

The principle conclusion drawn from this study is that a pixel resolution element of 3 metres is preferential for road extraction from scanned aerial photography using the extraction method of Benjamin and Gaydos (1984). Data sets with this spatial resolution showed the lowest overall omission error terms. The 3-metre data set was also smaller in size than data sets with finer resolutions which reduced data processing requirements for critical phases in the analysis.

Extraction accuracy fell when cluster signatures were extended beyond the original section used for their development. This seems to be due to the inclusion of new types of landscape features not well represented in the original test site. To use this extraction technique most effectively for map or DLG revision, areas of change should be identified and unique cluster signatures should be developed using digital values from only those areas. These can be applied to the scanned photo data to extract the road locations.

A final conclusion is that the digital database must be orthorectified to eliminate geometric, elevation, and aircraft-related errors in the photograph for extracted road networks to align properly with existing mapped information if these techniques are employed for large-scale cartographic purposes. This rectification can be performed optically on the photo before scanning to produce a second-generation photograph that could then be scanned, or, after scanning the original photograph, using a digital orthorectification program (Gaydos *et al.*, 1986). If the photo used here had been rectified before scanning, the extracted products could have been compared directly to the map or DLG to determine accuracy.

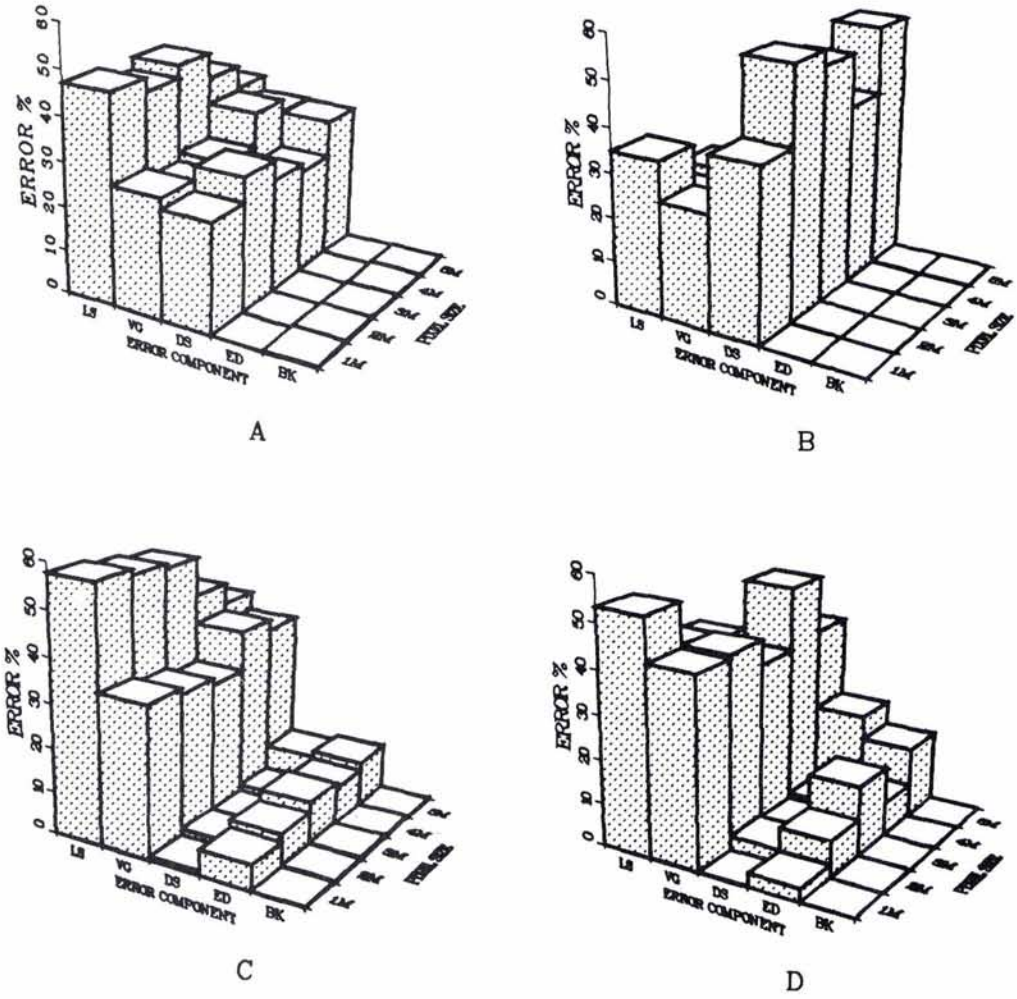

FIG. 3. Components of the omission errors: A, error set 1 (full test site before editing); B, error set 2 (original test site before editing); C, error set 3 (full test site after editing); D, error set 4 (original test site after editing). LS, light surface; VG, vegetation; DS, dark surface: ED, edited pixels; and BK, background. Note: order of columns selected to minimize hidden components.

Results achieved in this study can only be extended to other similar landscape environments. In other environments, the elements that make up the scene would have a different composition and would probably include elements of different sizes and materials. Presumably, in suburban areas with roads of similar widths, a similar composition of housing, commercial, and industrial buildings, with similar road construction materials, and with similar amounts, types, and locations of vegetation within the mosaic of elements, extraction products would be similar to those achieved in this study if 3-metre pixels were used. In other types of environments, the optimal spatial resolution might be larger or smaller.

To optimize this road extraction technique for map revision, other parameters aside from spatial resolution should be investigated. The spectral resolution of the data may play an equally large role in determining the effectiveness of the technique. Multiband imagery is imperative for this technique, but true color (rather than color infrared) photographs might be a better data source. Image quality aspects, both before and after scanning, should also be explored. There is clearly a trade-off between photographic scale and scanner aperture size when acquiring the photo data sets: is it better to achieve a defined pixel size with smaller scale photos and a small spot size, or to use many large-scale photos and a larger aperture? Questions concerning the best time of the year to acquire the photos to be scanned also need to be addressed. Most research in optimal dates of acquisition of remotely sensed data have centered on agricultural growth conditions. For this application, the best photograph date would be when the road network is least obscured and at the maximum contrast with the surrounding elements of the landscape.

REFERENCES

Benjamin, S., and L. Gaydos, 1984. Processing of scanned imagery for cartographic feature extraction. *Proceedings, 9th William T. Pecora Symposium*, Sioux Falls, South Dakota, pp. 222–230.

Bloomfield, P., 1976. *Fourier Analysis of Time Series: An Introduction*. New York: John Wiley and Sons, Inc., pp. 26–27.

Callahan, G. M., and F. R. Broome, 1984. The joint development of a national 1:100,000-scale digital cartographic data base. *Technical Papers, ACSM 44th Annual Meeting*, Washington, D.C., pp. 246–256.

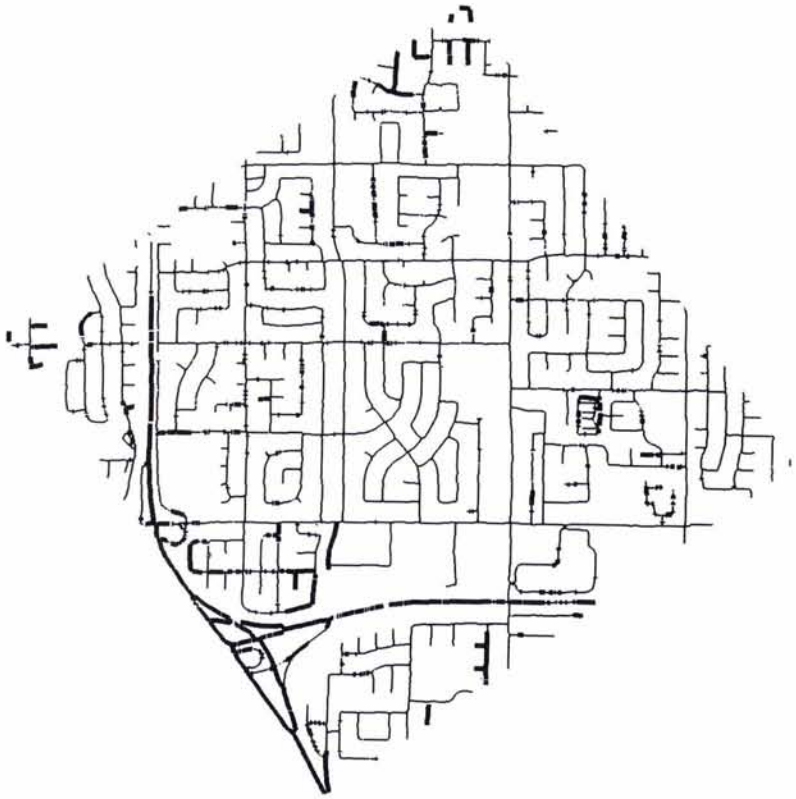

FIG. 4. Omission error pixel locations in the 3-meter data set; error pixels are shown in darker and thicker lines.

Fischler, M. A., and H. C. Wolf, 1983. *A General Approach to Machine Perception of Linear Structure in Imaged Data*. SRI International Technical Note 276. Menlo Park, California, 28 p.

Fischler, M. A., J. M. Tenenbaum, and H. C. Wolf, 1981. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Computer Graphics and Image Processing*, Vol. 15, pp. 201–223. New York: Academic Press, Inc.

Ford, G. E., and C. I. Zanelli, 1985. Analysis and quantification of errors in the geometric correction of satellite images. *Photogrammetric Engineering and Remote Sensing* 51(11): 1725–1734.

Gaydos, L., L. Ladner, R. Champion, and D. Hooper, 1986. Production of orthophotographs by digital image processing techniques. *Proceedings, ACSM-ASPRS Spring Convention*, Washington, D.C., March 1986.

Jensen, J. R., 1986. *Introductory Digital Image Processing*. Englewood Cliffs, New Jersey: Simon and Schuster, pp. 140, 212–214.

Markham, B. L., and J. R. G. Townshend, 1981. Land cover classification accuracy as a function of sensor spatial resolution, *Proceedings, 15th International Symposium on Remote Sensing of Environment*, Ann Arbor, Michigan, pp. 1075–1090.

Nagao, M., and T. Matsuyama, 1980. *A Structural Analysis of Complex Aerial Photographs*. New York: Plenum Press, 198 p.

NASA, 1982. Flight Summary Report 1603, Flight No. 82-087, 28 May 1982, Moffett Field, California: Ames Research Center, Airborne Missions and Applications Division.

Pries, R. A., and R. A. Schowengerdt, 1987. Extracting roads from digital images for cartographic products. *Proceedings, 5th Annual NW Conference on Surveying and Mapping*, American Society for Photogrammetry and Remote Sensing, Whistler Resort, British Columbia, 14-17 June 1987.

Scarpace, F. L., and B. K. Quirk, 1980. Land-cover classification using digital processing of aerial imagery, *Photogrammetric Engineering and Remote Sensing* 46(8): 1059–1065.

Townsend, F. C., 1981. *Feature Extraction and Planimetric Mapping by Computer Processing of Digital Stereopairs*. Ph.D. thesis, University of Wisconsin, Ann Arbor, Michigan, 482 p.

U.S. Geological Survey, 1961. Cupertino Quadrangle, California, 7.5-Minute Series (Topographic).

———, 1986. Digital line graphs from 1:24,000-scale maps. Data Users Guide 1.

Usery, E. L., and R. Welch, 1989. A raster approach to topographic map revision, *Photogrammetric Engineering and Remote Sensing* 55 (1):55–59.

Woodcock, C. E., and A. H. Strahler, 1987. The factor of scale in remote sensing. *Remote Sensing of Environment* 21:311–332.

Wrigley, R. C., W. Acevedo, D. A. Alexander, J. S. Buis and D. C. Card, 1984. *The Effect of Spatial, Spectral, and Radiometric Factors on Classification Accuracy Using Thematic Mapper Data*. NASA Technical Memorandum 855995.