Comparison between Digital and Manual Interpretation of High Altitude Aerial Photographs

Paul W. Snook
USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO 80521
Kenneth C. Winterberger
USDA Forest Service, Pacific Northwest Forest and Range Experiment Station, Anchorage, AK 99508
Norman E. Merritt and *Raymond L. Czaplewski*
USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO 80521

ABSTRACT: Second generation color infrared transparencies of the Tanana River Basin in Alaska were digitized using a scanning microdensitometer. Unsupervised clustering was performed independently on each of 20 digital images. Area estimates from 20 8-hectare sample plots were obtained by manual and computer aided interpretations. Manual interpretation of large-scale (1:3,000) photographs and ground truth served as reference. Computer-aided interpretation was consistently more accurate than the manual interpretation when compared to reference. However, the time and cost for digital processing was much higher than manual interpretation if information for only a small portion (e.g., an 8-ha plot) of the digital image is required.

INTRODUCTION

THE USDA FOREST SERVICE (FS) Forest Inventory and Analysis Project (FIA) in Alaska has implemented a four-phase sampling design for a large-scale, production inventory of land cover (Schreuder, 1981). The sampling design consists of digital analysis and interpretation of Landsat imagery, photointerpretation of high altitude (1:60,000-scale) aerial photographs from ASCS, photointerpretation of low altitude aerial photographs (1:3,000-scale) taken by the FS, and ground sampling. At each phase, a systematic 25 percent subsample of permanent 8-hectare (ha) sample plots are examined, and more detailed data are collected.

Manual interpretation of aerial photographs has been widely used to inventory land cover. However, this can be a labor intensive process with inconsistencies possible among photointerpreters. Digital interpretation has several advantages over manual interpretation. It permits mathematical transformations and enhancement of the imagery. It can also reduce bias and inconsistency among interpreters due to differences in skill levels, fatigue, and other factors. However, human interaction with the digital analysis is still required. By digitizing the photographs, the advantages of computer-aided interpretation can be combined with the fine resolution of aerial photography.

The purpose of this study was to determine if computer-aided estimates of land cover were comparable in accuracy and cost effectiveness to estimates obtained from manual interpretation in the USDA FS FIA production inventory of Alaska. FIA photinterpretation procedures (e.g., classification system, plot size) were observed to effectively compare the two types of interpretation. Second generation, small-scale photography from ASCS was evaluated because this type of photograph is available for large areas of Alaska and other parts of the United States.

REVIEW OF LITERATURE

Estimation of land cover using digitized photographs assumes that the desired land-cover features can be separated based on film density. A relationship is assumed between the incident energy on the film and the resulting film density values. However, this relationship is nonlinear for multi-emulsion film, such as color and color infrared (CIR) photographs. Therefore, radiometric calibration procedures have been developed to correct for this nonlinearity (Scarpace, 1978; Scarpace and Friedericks, 1978). Also, Kalman and Scarpace (1979) recommended correcting for decreased intensity of incident light towards the outer edges of a photograph (i.e., vignetting or lens fall-off). This causes different film densities for the same cover type.

Without the aid of ancillary information, such as digital terrain data, the computer-aided interpretation of digital images is limited to measurements of spectral reflectance in classifying cover type. Differences in sun angle will cause differential illumination (e.g., shadows) of target scenes (Kimes, 1980). This problem is most severe in extreme northern latitudes such as Alaska. Photointerpreters, however, can integrate other subtle clues such as tone, texture, color, and spatial features in an image. These allow the photointerpreter to more accurately classify an area influenced by topography and shadowing.

Digitized photographs have been used with varying degrees of success. Accuracy depends on the detail of information desired and the use of the proper correction procedures. Hoffer *et al.* (1972) reported accuracies of 94 percent and 95 percent for digitized 1:120,000-scale color infrared (CIR) and black-and-white multiband photography, respectively. Quirk and Scarpace (1982), working with a digitized 1:24,000-scale CIR photograph from the Green Bay Wisconsin area, concluded that analysis of a digitized photograph could analyze large areas at a finer resolution faster than manual interpretation.

Mace and Bonnickson (1982) compared digitized 1:120,000-scale CIR photography with Landsat multispectral scanner (MSS) data for classifying forest-cover types on the Apostle Islands in Wisconsin. Film calibration procedures were followed except that the vignetting correction was omitted because the small study area was close to the image principal point. A fourth band, consisting of texture transformation data, was added to the three spectral reflectance bands. Ground resolution was set at 36 m. Overall classification accuracies for the MSS were less than for the digital CIR data (47.9 percent and 68.3 percent, respectively). Accuracies increased as the detail of land-cover classes was decreased by aggregating forest type classes.

Turner and Thompson (1982) digitized eleven 1:21,000-scale CIR photographs to classify a North Carolina barrier island into water, sand dune, mesic shrub, xeric shrub, shrub thicket, pine

* The use of trade and company names is for the benefit of the reader. Such use does not constitute an official endorsement or approval of any service or product by the U.S. Department of Agriculture to the exclusion of others that may be suitable.

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 53, No. 5, May 1987, pp. 531–534.
0099-1112/87/5305–531$02.25/0

woodland, and high and low salt marsh. No film calibration procedures were applied to the digital data. The overall classification accuracy was 62 percent. The classification accuracy for vegetation types was 48 percent. The authors acknowledged that film calibration and a classification system suited for remote sensing might have increased accuracy.

Jensen and Estes (1978) indicated that vignetting can severely reduce the accuracy for digital analysis of multitemporal high altitude (1:120,000-scale) CIR photographs. Efforts were made to reduce the effects of vignetting by using antivignetting filters on the camera system and preprocessing the digitized density values by an antivignetting filtering algorithm. However, signature extension of crop types across the digitized photography was still hampered by vignetting effects.

As indicated above, there has been varied success in estimating land cover by use of digitized photographs. Various scales, classification levels, and geographic locations have been investigated. This study was designed to evaluate the use of digitized photographs for the FIA production inventory in Alaska, incorporating the existing photographs, classification scheme, and plot size used by this FIA unit.

STUDY AREA

Plots analyzed in this study were from a portion of the Tanana River basin in Alaska between 147 and 152 degrees west longitude, and 63 and 66 degrees north latitude. This area consists of three major physiographic regions (Wahrhaftig, 1965).

The Tanana-Kuskokwim lowland consists mainly of a broad glacial outwash plain, overlain in some areas by aeolian deposits. It is characterized by large, relatively low-productivity forest stands with a few large lakes. Lower elevations consists of flood plains of major rivers and many small lakes and oxbows interspersed with bogs and relatively high productivity forest stands.

The Yukon-Tanana upland is an area of rolling hills ranging in elevation from 335 to 1615 metres. Lower elevations are dominated by aspen (*Populus* tremuloides), birch (*Betula* papyrifera), and mixed birch/spruce (*Picea* spp.) forests. Much of the southern half is characterized by deep, well drained soils with high productivity upland forests.

The northern foothills of the Alaskan range have shallow soils and sparse tree cover. The major vegetation includes sedges (*Carex* spp.) and low to dwarf shrubs, including *Betula* spp. and *Salix* spp. Trees are found in small areas of deeper soil along drainage ways.

METHODS

Twenty second-generation CIR 9- by 9-inch (22.9- by 22.9-cm) aerial photo transparencies of 1:60,000 scale were used. Reference data (large scale photointerpretation and ground truth) were available for one plot on each photograph. The photos were acquired by NASA during August 1978, 1979, 1980, and 1981, and were photographed with a combination of Zeiss and Wild RC-10 cameras with 12-inch focal length lenses and Wratten 12 filters. The film used was Aerochrome CIR SO-103.

One 8-ha circular sample plot was registered onto each of the 20 photo transparencies by FIA inventory personnel in Alaska. A 2.5 cm^2 area of the photo, which included each 8-ha plot (0.5 cm on photo), was digitized on the University of Wisconsin's Environmental Remote Sensing Center's Optronics P-1700* scanning drum microdensitometer, which had been modified for color scanning. An area surrounding each sample plot was digitized to provide an adequate number of pixels for developing statistics needed for the classification algorithm. The aperture setting of 50 μm resulted in a nominal pixel resolution of 3 m on the ground. The analytical densities were approximated by separating the integral density with the use of narrow band lens filters, centered at wavelengths of 0.45 μm, 0.55 μm, and 0.65 μm. These wavelength regions correspond to the blue, green, and red film dye layers. In CIR photography, these film layers represent green, red, and IR reflectance, respectively. Filters allowed only the transmitted light from the desired dye layer to be detected by the microdensitometer.

The ratio of the IR and the red bands was subsequently added as a fourth variable. The IR/RED ratio can be a good indicator of green biomass (Tucker, 1979) and can help discriminate between vegetation cover types. This transformation can also help standardize spectral response across the scene. Vignetting and bidirectional reflectance causes spectral response (film density) of a cover type to vary. The ratio should be less affected because both bands would be similarly, although not identically, affected.

Although film calibration procedures are known to be important, it was not possible to apply the calibration procedures in this study. Film wedges for the second generation photographs, necessary for the radiometric calibration, were not available. Similarly, the lens and filter combinations for each camera system, which were necessary for antivignetting corrections, were unavailable. This situation may not be ususual for production inventories when second or third generation photographs must be used.

Digital data from the 2.5 cm^2 area were classified using CLUSTER, a minimum distance to mean unsupervised clustering algorithm. Preliminary classification results using the 3-m pixel resolution showed a high level of spectral variability because of this high spatial resolution. Shadows or gaps within the vegetation cover were detected in addition to spectral responses from the vegetation. Spatial resolution of the data was reduced by averaging density values over pixel blocks to reduce or mask the spectral variability. A series of pixel blocks, up to 5 by 5, were tested. A 3 by 3 block was subjectively judged to provide the best reduction of spectral variability while retaining adequate spatial resolution. This reduced the ground resolution to a nominal 9 m.

Unsupervised clustering was performed independently for each of the 20 image files using all pixels from all four bands within the 2.5 cm^2 digitized area. Each image was classified independently because color balance, vegetation phenology and species mixture, sun angle, and atmospheric conditions differed among photographs. Therefore, a cover type on one photograph did not exhibit the same color and tone as a similar feature on another photograph.

Upon completion of the clustering, the classification was displayed by assigning a color coded print pattern to each cluster. Where more than one cluster appeared to represent a single cover class, those clusters were aggregated into the same print pattern. If separation between cover types, as discerned by the analyst from the area surrounding the sample plot, were judged inadequate, then the entire scene was reclustered using a different combination of input parameters to the clustering algorithm.

The sample plot center was visually registered from the high altitude photograph onto the final classified image. Each print pattern on the classified image was independently labeled by a trained photointerpreter who was familiar with the study area as to the cover type each print pattern represented on the ground (e.g., a red pattern may represent a deciduous forest, a green pattern a coniferous forest, and a black pattern water). To avoid biasing identification of cover types within the sample plot, the labeling process was conducted on areas within the digitized section of the photo but outside of the sample plot. Within each 8-ha sample plot, the percentage of each classified cover class (Table 1) was calculated.

Second generation CIR photo transparencies were used in the manual interpretation of the high altitude photography. The transparencies remained in roll form and were viewed in stereo on a light table. Plot centers were located and marked on the photographs corresponding to specific UTM coordinate locations as depicted on 1:63,360-scale USGS map sheets. A circle, representing the limits of an 8-ha plot at the scale of the photography, was drawn on mylar (average circle size was 0.5-cm diameter). Photointerpreters were familiar with the vegetation

TABLE 1. LEVEL I AND LEVEL II LAND-COVER CLASSIFICATION, BASED ON VIERECK *et al.* (1982)

LEVEL I	LEVEL II
Forest	Conifer Deciduous Mixed
Shrub	Tall Shrub Low Shrub
Grassland	Grassland*
Non-Vegetated	Barren Water

* Note that the Level I grassland class was not subdivided at Level II because only one grassland type was present among all plots.

and physiography of the lower Tanana River Valley. Proportional estimates of land cover for each 8-ha sample plot were made ocularly with the aid of a 19-point mylar grid. The vegetation classification scheme (Table 1) was based on a modification of the Preliminary Vegetation Classification System for Alaska developed by Viereck *et al.* (1982) Manual interpretation of 1:3,000-scale CIR photography served as a reference to the two high altitude interpretations. Delineated areas on the large-scale photographs were planimetered and field checked.

The sample plot was visually reregistered twice in this study; from the manually interpreted high altitude photographs to the photos used for digitizing and to the final classified image. Misregistration could increase errors in estimating land cover. Misregistration might also exist between the high altitude sample plots and the large scale reference data. Every effort was made to obtain accurate registration of sample plots. Errors due to misregistration are assumed to be negligible in further analyses.

Comparisons were made of total area estimates from all plots for each cover class between the two high altitude interpretations and the large scale interpretation. The Kendall tau statistic (Gibbons, 1976) was used to assess reliability. This nonparametric analysis was chosen to test the association between interpretations because of the continuous and non-normal nature of the data. Land cover was estimated as a percentage within a sample plot. Because there were no categorical point comparisons made, the use of error matrices was not possible. Parametric tests were not used because estimation errors were not expected to be normally distributed. Rosenfield (1982) provides a description of various nonparametric tests for remote sensing. The Kendall tau statistic values range from -1 to $+1$, where -1 indicates an inverse association, 0 no association, and $+1$ a direct association. This test was applied to both Level I and Level II cover classes.

RESULTS

COMPARISON OF ACCURACY

Estimated proportion of total area for all 20 sample plots at the Level I and Level II classes for manual interpretation (MAN), digital interpretation (DIG), and large-scale photography (LSP) are given in Table 2. At Level I classification, the DIG estimates were as accurate or more accurate than MAN. For Level II, accuracy was similar except for mixed forest (DIG underestimated), low shrub (MAN underestimated), and grassland (MAN overestimated).

Using an aggregated estimate for all plots may mask errors for individual plots by compensating over- and underestimates. The Kendall tau statistic was used to test reliability of individual area estimates for each cover class at the plot level (Table 3). At Level I, the DIG was significantly associated with reference data for all classes. For MAN, only forest and non-vegetated classes were significantly associated. At Level II, MAN interpretations of conifer, deciduous, and barren classes are significantly associated with reference data, while these plus the grassland and water classes were significant for DIG.

TABLE 2. A COMPARISON OF ESTIMATED AREA (IN PRECENT) FOR MANUAL (MAN), DIGITAL (DIG), AND LARGE SCALE PHOTO (LSP) INTERPRETATION BASED ON TWENTY 8-HA PLOTS

LEVEL I			
	MAN	DIG	LSP
Forest	64.4	61.8	62.6
Shrub	12.9	19.3	18.0
Grassland	11.1	7.5	7.8
Non-Veg.	11.6	11.4	11.6
Total	100.0	100.0	100.0
LEVEL II			
	MAN	DIG	LSP
Conifer	46.3	46.5	46.7
Deciduous	14.2	14.2	11.0
Mixed Forest	3.9	1.1	4.9
Tall Shrub	4.5	5.7	5.0
Low Shrub	8.4	13.6	13.0
Grassland	11.1	7.5	7.8
Barren	9.7	9.3	9.7
Water	1.9	2.1	1.9
Total	100.0	100.0	100.0

TABLE 3. VALUES OF KENDALL TAU STATISTIC COMPARING INTERPRETATIONS AT THE PLOT LEVEL (Ho: NO SIGNIFICANT DIFFERENCE BETWEEN THE TWO INTERPRETATIONS). TAU SIGNIFICANTLY DIFFERENT FROM 0 ($P \leq 0.05$) INDICATED BY "*".

LEVEL I		
	MAN/LSP	DIG/LSP
Forest	0.9252*	0.8645*
Shrub	0.4046	0.6228*
Grassland	0.4914	0.5389*
Non-Veg.	0.6866*	0.8986*
LEVEL II		
	MAN/LSP	DIG/LSP
Conifer	0.7639*	0.8001*
Deciduous	0.7539*	0.4933*
Mixed Forest	−0.0714	0.4454
Tall Shrub	0.3236	0.2509
Low Shrub	0.4097	0.3284
Grassland	0.4914	0.5389*
Barren	1.0000*	1.0000*
Water	0.4447	0.8281*

Several anomalies appear at the individual plot level which are not apparent in aggregated plot totals. The mixed forest classification by MAN, which had smaller differences in total area than the DIG, shows very low association at the plot level (Kendall tau of 0.07 compared to 0.44 for the DIG). Kendall tau for DIG indicates a better association for mixed forest (although it is not significant). The DIG interpretation of the low shrub class had small difference for the aggregated totals but had low accuracy at the plot level. However, the DIG interpretation was again more accurate than MAN.

COST EFFECTIVENESS

It required approximately 5 to 10 minutes per plot for manual interpretation of the high altitude photography, while an average of 1.5 hours were required to complete one cycle of the digital interpretation. The times given for computer processing were system dependent. For this study, a Data General Eclipse S230 minicomputer was used. The time for digital processing, as averaged over all plots, includes

- One-half hour for 12 iterations of clustering the digital image (233 by 233 pixels or approximately 228 ha). The entire image data file for one photograph was used to develop cluster statistics for classifying the image. The classifier was given freedom to produce up to 30 clusters to account for variability within classes.
- One-half hour for the analyst to print the cluster classes and determine if clustering parameters used provided sufficient

separation of the data. This time increased as the number of clusters produced increased.
- One-half hour to produce a final class map ready for labeling.

However, oftentimes the analyst determined that adequate separation of the cover classes was not achieved after clustering, and the clustering process was repeated using a different set of input parameters. For most scenes, at least two cluster runs were required. A set of input parameters which produced an adequate classification on one scene would not always work well on another. This was due to differences in color balance between photographs. The repeated clustering of the data was the single most costly procedure of the digital classification.

CONCLUSIONS

Computer-aided interpretation of high altitude color infrared photography produced results more similar to reference data than did manual interpretation of high altitude photography. At Level I of the classification scheme, computer-aided interpretation provided smaller relative errors of aggregated area estimates than the manual interpretation. Plot level estimates by computer-aided interpretation were significantly associated with reference data for all four Level I classes; only two of the four Level I classes were significant for manual interpretation. For Level II classification at plot level, five of eight classes were significantly associated for computer-aided interpretation, while only three of eight were significant for the manual interpretation.

Although the computer-aided classification was consistently more accurate than the manual interpretation, the time and cost involved for computer processing must be considered. If the information from a relatively small sample plot alone is required, as in this study, computer-aided interpretation is not cost effective compared to manual interpretation. The time for manual interpretation in this study was much less than for the digital. However, the 8-ha sample plot is only a small portion of the digitized photograph. The entire digital image (228 ha) was classified along with the 8-ha sample plot. If the sampling design of the inventory project could take advantage of the added information provided by the increased sample size of the full digital image, the computer analysis would be an effective alternative to manual interpretation. The potential of computer-aided interpretation increases with the size of the photographic image to be interpreted.

ACKNOWLEDGMENTS

The authors would like to achknowledge Dr. Lee F. Werth for the original study design and data acquisition. The authors would like to thank George Rosenfield, Frank Sadowski, and Ann Maclean for their reviews and comments.

REFERENCES

Gibbons, J. D., 1976. *Nonparametric Methods for Quantitative Analysis*. Holt, Rinehart and Winston, New York.

Hoffer, R., P. E. Anuta, and T. L. Phillips, 1972. ADP of multiband and multi-emulsion digitized photos. *Photogrammetric Engineering and Remote Sensing* 38(10):989–1000.

Jensen, J. R., and J. E. Estes, 1978. High-altitude versus Landsat imagery for digital crop identification. *Photogrammetric Engineering and Remote Sensing* 44(6):723–733.

Kalman, L., and F. L. Scarpace, 1979. Determination of lens fall-off through digital analysis of photographic imagery. *Proceedings, American Society of Photogrammetry Convention*, W:`ington D.C. 18–24 March 1979. pp. 116–135.

Kimes, D. S., J. A. Smith, and K. J. Rnason, 1980. Vegetation reflectance measurements as a function of solar zenith angle. *Photogrammetric Engineering and Remote Sensing* 46(12):1563–1573.

Mace, T. H., and T. M. Bonnicksen, 1982. A comparison of digitized high-altitude photography and Landsat imagery for forest-cover mapping—Apostle Islands, Wisconsin. *J. of Applied Photographic Eng.* 8(3):111–116.

Quirk, B. K., and F. L. Scarpace, 1982. A comparison between aerial photography and Landsat for computer land-cover mapping. *Photogrammetric Engineering and Remote Sensing* 48(2):235–240.

Rosenfield, G. H., 1982. The analysis of aerial data in thematic mapping experiments. *Photogrammetric Engineering and Remote Sensing* 48(9):1455–1462.

Scarpace, F.L., 1978. Densitometry on multi-emulsion imagery. *Photogrammetric Engineering and Remote Sensing* 44(10):1279–1292.

Scarpace, F. L., and G. A. Freiedericks, 1978. A method of determining spectral analytical dye densities. *Photogrammetric Engineering and Remote Sensing* 44(10):193–1301.

Scarpace, F. L., B. K. Quirk, R. W. Kiefer, and S. L. Wynn, 1981. Wetland mapping from digitized aerial photography. *Photogrammetric Engineering and Remote Sensing* 47(6):829–838.

Schreuder, H. T., 1981. Sampling statistics for multi-resource inventories. *Bulletin of International Statistical Institute* 49(1):250–270.

Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment* 8(2):127–150.

Turner, B. J., and D. N. Thompson, 1982. Barrier island vegetation mapping using digitized aerial photography. *Photogrammetric Engineering and Remote Sensing* 48(8):1327–1335.

Viereck, L. A., C. T. Dynness, and A. R. Batter, 1982. *1982 Revision of the Preliminary Classification for Vegetation of Alaska*. Unpublished report on file at the Institute of Northern Forestry, Fairbanks, Alaska.

Wahrhaftig, C., 1965. *Physiographic Divisions of Alaska*. Geological Survey Prof. Paper 482. Washington, D.C. 52 p.

(Received 5 May 1986; revised and accepted 21 January 1987)

Forthcoming Articles

Albert K. Chong, A Robust Method for Multiple Outliers Detection in Multi-Parametric Models.
Max M. Ethridge, Combined Control Densification and Mapping.
Jess Grunblatt, An MTF Analysis of Landsat Classification Error at Field Boundaries.
Gordon Hoover and *Anne B. Kahle*, A Thermal Emission Spectrometer for Field Use.
Gail D. Kelly and *Greg J. E. Hill*, Updating Maps of Climax Vegetation Cover with Landsat MSS Data in Queensland, Australia.
Byung-Guk Kim, Surface Water Velocity Measurements on Delavan Lake Using Aerial Photographs.
Daniel C. Oimoen, Evaluation of a Tablet Digitizer for Analytical Photogrammetry.
Dan Rosenholm, Multi-Point Matching Using the Least-Squares Technique for Evaluation of Three-Dimensional Models.
S. A. Veress and *Huang Youcai*, A Method for Improving the Efficiency of the Sequential Estimation Procedure in Photogrammetry.