Visual Interpretation of Vegetation Classes from Airborne Videography: An Evaluation of Observer Proficiency with Minimal Training

Sam Drake

Abstract

This study evaluated the ability of individual subjects and small groups to correctly identify Arizona plant communities from color airborne video footage, explored the relationship of five background variables to subjects' success, and determined which community types were easiest and most difficult for subjects to identify. Forty-six volunteers from the University of Arizona School of Renewable Natural Resources participated in a multiple-choice pretest-posttest experiment using 30 different plant communities depicted in 1-minute segments of videotape. Three hours of training increased the mean individual score from seven correct (pretest) to 21 correct (posttest), and mean group score from 11 to 24. All respondents significantly improved their performance, regardless of background. Posttest results showed no significant difference in ability among individuals or between individuals and groups. The most difficult community to identify was creosote-tarbush desertscrub; the easiest was paloverde-saguaro desertscrub. Findings support the feasibility of video interpretation by minimally trained personnel.

Introduction

Airborne videography is an emerging technology used by natural resource managers to identify and map various Earth surface features, including vegetation. Georeferenced videography can function as an independent data source or as an adjunct to satellite imagery for both map creation and map verification. In particular, it is an inexpensive proxy for ground truth. Initial equipment costs are relatively low, and image acquisition costs amount to little more than aircraft rental. Interpretation of video footage, however, can be time consuming and, thus, costly if done by upper-level professionals.

Though methods do exist for video framegrabbing and computer image processing, a human observer is essential for synthesizing cues from plant size, shape, color, texture, and configuration, and for rendering a judgment as to what community is depicted in a frame or segment of tape. The overall cost-effectiveness of airborne videography as a tool for resource managers could be maintained if some or all of the necessary human interpretation could be done by minimally trained staff.

Information is lacking on the adequacy of interpretation by such staff and on the feasibility of widespread adoption of video technology for decision support without reliance on specialists. This study is an initial effort to determine the accuracy of video interpretation by naive observers, the kind and degree of variation among observers, and the efficacy of a minimal training program for improving observers' performance.

Videography can be used for many tasks once performed with aerial still photography. The spatial resolution and image clarity of video cannot match those of still photography, but the range of spectral sensitivity possible with video cameras is much wider than that of film cameras, and the turnaround time for video acquisition is negligible compared with that of traditional aerial photography (Graham, 1993). Gain settings and other adjustments on the video camera can be used to compensate for poor flying conditions, especially low light. Image motion and the opportunity to view a spot on the ground from different perspectives in successive frames often aids interpretation. Also, features are seen at different scales in successive frames as the camera zooms in or out. The markedly lower cost of video acquisition makes repeated resource monitoring surveys more feasible than they would be with film photography.

Videography cannot match the synoptic views afforded by satellite imagery, but can serve as an accessory source of ground-truth information for mapping resources over large areas. For example, georeferenced video sampling (identification) of plant communities can be used as input for a supervised classification of satellite imagery, or to confirm and label polygons resulting from an unsupervised classification (Kalliola and Syrjanen, 1991). Usually, some actual ground sampling will be necessary to create a key for interpretation of the videography, but this can be minimal. Graham (1993) explains how to tag video frames with latitude-longitude coordinates received from the Global Positioning System (GPS) constellation of navigation satellites to facilitate location of plots for ground truthing. For some resource analysis projects the application of airborne videography will obviate ground sampling altogether. The advantages of airborne videography over ground survey methods are evident: tremendous savings of time and money, with continuous pictorial coverage of an area of interest. Flown at 120 mph with a 1/3-mile swath width, an airborne video camera can sample 40 square miles per hour. The cost of aircraft rental and all other mission expenses should be less than $300 per hour (Charles C. Curtis, pers. comm.).

While financial incentives exist to use airborne video, and technological factors permit the acquisition of quality footage, its ultimate utility still depends on human factors which bear on interpretation. Users must be able to recognize what they are looking at on the video monitor in order to distinguish features with desired resolution and accuracy. Some features or patterns may be consistently easier to inter-

Arizona Remote Sensing Center, 1955 E. 6th Street, Tucson, AZ 85719-5224.

Photogrammetric Engineering & Remote Sensing, Vol. 62, No. 8, August 1996, pp. 969–978.

0099-1112/96/6208–969$3.00/0
© 1996 American Society for Photogrammetry and Remote Sensing

pret than others, or interpretability may vary with aspects of a person's background, education, and aptitude. Different people looking at the same complex pattern of vegetation may "see" the same community type, or different types. Collaborating small groups may produce more accurate interpretation than single observers, or may not. These issues affect the reliability of the interpretations on which management decisions must be based.

As more people become involved and larger inventory or monitoring projects are conceived, the issue of training video interpreters inevitably arises. Two phases of this training can be recognized. In the initial phase, an identification or interpretation key is created, with the aid of independent ancillary information, which covers the range of patterns likely to be encountered in project video footage. This key may be informal: indeed, it may be simply the accumulated knowledge of one person, derived from field work and reading. In the second phase of training, this key can be used to transfer the ability to interpret video imagery to multiple observers, who can then process as many hours of footage as are required.

The present study focuses on this second phase of training and the above issues concerning human interpretation of videography. Objectives are (1) to evaluate the ability of individual subjects and small groups to identify correctly Arizona plant communities from color airvideo footage, both before and after a brief training program; (2) to explore the relationship of five background variables to subjects' success; (3) to judge the efficacy of a minimal training program for improving subjects' performance; and (4) to determine which communities are easiest and which are most difficult for subjects to identify. The five background variables examined are formal education in plant identification or community ecology, time of residence in Arizona, frequency of outdoor recreation, self-assessment of knowledge of Arizona vegetation communities, and previous exposure to airborne videography or the view from light aircraft.

Interpretation of particular features in airborne videography is mentioned perfunctorily in the literature, usually as a topic ancillary to discussion of hardware technology. No published research has addressed the more specific subject of training video interpreters, though the training of aerial photointerpreters is analogous and a body of literature does exist on this subject. Quinn (1947) reviews aspects of the training of military photointerpreters during and just after World War II. At that time Colwell (1946, 1948) developed methods for interpreting vegetation and associated ground conditions from aerial photos, and for teaching the skill to students in the armed forces. In rare experimental research, Klingberg *et al.* (1963) investigated the ability of experienced and inexperienced photointepreters to detect change in pairs of comparison photographs with side-by-side, overlay or apparent-motion display. Other work has been done on more general aspects of photointerpretation, such as the production of interpretation keys (Loelkes, 1983). Basic principles of photointerpretation are presented in standard texts and manuals (Avery, 1985; Estes *et al.*, 1983). Pertinent to the present study, Julesz (1975) illustrates the difference between perceptual and cognitive recognition in a useful discussion of the visual perception of texture. Much recent research has focused on pattern recognition and image interpretation by computer; an excellent theoretical review is given by Argialas and Harlow (1990).

Materials and Methods

Participant Pool

This study was conducted with the aid of 46 volunteer subjects recruited from the School of Renewable Natural Resources at the University of Arizona. Twenty-four of these were graduate students and 11 were undergraduate students taking at least one class in the School. Five participants were faculty members, four were technical staff, and two were alumni and currently senior staff at The Nature Conservancy, Tucson, Arizona. All participants were clearly well-educated, and similar in their claims of an academic major related to biology or natural resources. Even so, the subject pool represented a range of familiarity with Arizona vegetation types, and some subjects had limited previous exposure to airborne videography, as revealed in an initial questionnaire completed by each participant (Appendix A). Subjects were not compensated materially for participating, but undergraduates were offered class credit by their instructor.

Vegetation Classes

For this study, 30 Arizona plant communities were selected from the list established by Brown, Lowe, and Pase (Brown, 1982). Communities were selected to represent some of the most important or extensive vegetation types found in Arizona, encompassing three distinctly different life forms (grassland, shrubland, forest), with sets of similar examples within each life form to measure subjects' ability to make fine distinctions. Appendix B lists the 30 communities used, with numerical codes and names adapted from Brown (1982).

Videography

Airborne videography footage for this study was extracted from some 90 hours of footage acquired for the Arizona Gap Analysis Project (GAP) by researchers in the University of Arizona's Advanced Resource Technology Program. The natural color Super-VHS footage was recorded from about 600 m above ground level using a Panasonic AG-7400 portable videocassette recorder and a Panasonic 300CLE professional grade video camera aimed vertically downward through the belly port of a Cessna 180 aircraft. The camera was fitted with a Canon 12X motorized zoom lens; at approximately 10-second intervals, the camera operator in the airplane would fully extend the lens, hold the zoomed view for about 3 seconds, then pull back to the wide-angle view (Graham, 1993). This footage was edited into segments lasting 45 to 60 seconds, and containing at least two zooms, for each community. The edited, composite S-VHS tapes were replayed for test subjects on a JVC HR-S4700U commercial grade VCR with a Sony Trinitron 13-inch color monitor. The VCR was capable of crystal freeze-frame and variable speed reviewing, both forward and backward. All of the footage had been GPS-referenced when acquired (Graham, 1993), but location information was deleted during editing and subjects were not told where in Arizona the segments were recorded. GAP footage was recorded over a one-year period, from June 1991 to June 1992, and showed seasonal variation in the appearance of some vegetation communities. Segments selected for this study were representative of each community as it appears most of the year in Arizona; leafless wintertime scenes were avoided.

Experimental Design and Procedure

The experimental design was the one-group pretest-posttest design described by Campbell and Stanley (1963). Each participant was asked to take part in an individual pretest, a three-person group pretest, two training sessions, an individual posttest, and a three-person group posttest. Each training session lasted 1 1/2 hours and each test 1 hour. The "individual" tests were sometimes given to as many as five subjects simultaneously, but each worked independently and interaction was prohibited. To maintain uniformity, the experimenter operated the VCR during all of the individual tests, pausing and going slowly through the zooms in particular, spending about 2 minutes on each of the 30 segments. Each participant was given a community list (Appendix B) and was told to choose from it the one name which best applied to each segment of videotape. Subjects were instructed that they could use a name more than once but that they should *not* change any answers to previous segments once the tape had advanced to a different community. Each group was instructed to give its best collective response, also in about 2 minutes' time, and to not return to any segment. To encourage free discussion among group members and to avoid biasing their answers, the experimenter did not monitor group tests.

Pretest groups were constituted by circumstance and scheduling convenience, not for any particular affinity among members, though in most cases group members knew and liked each other. Not all groups could be exactly reconstituted for the posttest. Of 13 pretest groups, two were not reconstituted at all due to attrition, two changed one member, and one group of four lost two members. Four groups had only two members; most had three. Not all individuals could be assigned to groups. Of the 46 individuals pretested, 38 completed the posttest.

For the posttests, a second composite tape was made, similar to the pretest tape. The segments appeared in a different order and were edited from different source footage, so that no specific patterns or features could be recognized from the pretest tape or any of the training tape used.

The 3 hours of training were conducted in two 90-minute sessions, repeating each one six times over a 3-week period to accommodate all participants. In the first session, everyone was given a copy of his/her own pretest answer sheet, group pretest score (#correct/30), and an ordered key to the segments on the pretest tape. Then the pretest tape was reviewed in detail, focusing on cues that allow identification of plant species and differentiation of the communities of interest. In the second training session, about 45 minutes of videotape were shown and described, giving new examples of the 30 communities; then participants examined a book of (labeled) still video frame prints, and their questions about video interpretation were answered. The video prints, showing both zoomed and wide angle views of each community, were produced directly from the pretest and training tapes using a Hitachi VY-200A color video printer. Trainees were allowed to take notes and to use those notes on the posttests, but were reminded that time would be limited to 1 hour.

Participants were asked to take the posttests as soon as possible after completing both training sessions. Many did so within a few days, most within one week, but several took longer than one week to schedule both an individual and a group posttest. Everyone was given an opportunity to review the book of video prints for 10 minutes immediately before taking a posttest. The experiment lasted a total of 54 days (18 February to 12 April 1993), from the first individual pretest to the last group posttest.

Statistical Analysis

For a simple test of significant difference or improvement in raw scores (number correct), a paired t-test was used (Ott, 1988), with significance assumed at the 95 percent confidence level. Pretest scores were compared to posttest scores for the 38 individuals and 11 groups which completed both. Individual pretest scores were compared to group pretest scores and individual posttest scores to group posttest scores with the paired t-test. In fact, a more illuminating result emerged from simple inspection of the scores.

To examine subjects' responses in a more comprehensive and detailed way, they were entered into error matrices (e.g., Appendix C) and analyzed with a basic discrete multivariate technique described by Congalton *et al.* (1983). The summary statistic KHAT, and its associated variance, were calculated for each of 30 composite error matrices compiled. KHAT uses all cell values in a matrix; it is a measure of the overall accuracy of a matrix and can be used to test for differences between matrices. This statistic can test for significant differences between photointerpreters, and for the consistency of the same interpreter over time (Congalton and Mead, 1983). KHAT varies from 0 to 1, with 0 indicating all responses in the matrix maximally incorrect, and 1 indicating a perfectly correct matrix with all responses on the major diagonal.

Each individual and each group test response sheet yielded one "primary" error matrix; in all, 108 were produced. These primary matrices could have cell values of only 1 or 0, with 1 denoting a response in that cell (whether correct or not) and 0 no response there. Each of these matrices could contain up to 30 1's; fewer if a participant or group left blanks on their answer sheet or mistakenly recorded a plant community code number not on the list provided. Matrices were constructed as spreadsheets for manipulation with Quattro Pro software (Borland International Inc., Scotts Valley, California).

Secondary or composite error matrices were compiled from sets of primary matrices by summing corresponding cell values from each primary matrix in the set. Some sets were of obvious interest: the set of all individual pretests, that of all individual posttests, those of all group pretests, and all group posttests. Also, the 38 individual and 11 group pretests that had counterpart posttests were combined. Other sets were composed based on participants' answers to questions on the initial questionnaire: those who had taken 2+ pertinent classes vs. those who had taken <2 classes; those who had lived in Arizona 5+ years vs. <5 years; those who go hiking often or quite frequently vs. rarely or sometimes; those who rated their own knowledge of Arizona vegetation communities good or excellent vs. poor or fair; and those who fly frequently in light aircraft or have had previous exposure to airborne videography vs. those who had neither experience. Lastly composed was a set of all participants having high "expected" familiarity with Arizona vegetation because they displayed at least two of the following three qualifications: had taken 2+ pertinent classes, had lived 5+ years in Arizona, or went hiking often or quite frequently. For the latter sets, only individual (not group) test matrices were combined and compared.

Using KHAT and its variance, 15 pairwise comparisons of composite matrices were made, testing for significant difference (Table 3). Following the procedure of Congalton and Mead (1983), significance was assessed with a Z-statistic, where

$$Z = (\text{KHAT1} - \text{KHAT2}) / \sqrt{\text{var1} + \text{var2}}.$$

At the 95 percent confidence level, a difference is significant when Z is greater than 1.96.

Inspection of cell values in composite matrices allowed an interpretation of which plant communities were easiest for participants to identify and which were most difficult. The communities defined to be easiest were those most frequently identified correctly on *both* the pretest and posttest. Likewise, the most difficult communities were those missed most often on both pretest and posttest. The three or four communities at each extreme were determined for the set of all individual participants and for the sets of all three-person groups; individuals with previous airborne video exposure

Table 1. Raw Scores, Reported as Number of Vegetation Communities Identified Correctly out of 30 Possible. Individual Participants Are Identified with an Arbitrary Sequential Number, and Group Composition Can Be Inferred from Group Scores

	Individual		Group	
subject	pretest	posttest	pretest	posttest
1	16	—	15	—
2	11	—	15	—
3	15	29	15	—
4	12	30	18	29
5	11	27	18	29
6	13	28	10	25
7	6	23	10	25
8	6	20	10	25
9	10	25	9	24
10	6	23	9	24
11	10	21	11	22
12	5	19	11	22
13	3	20	11	22
14	10	26	15	27
15	4	22	15	27
16	3	24	15	27
17	9	26	10	20
18	9	23	10	20
19	5	15	10	20
20	7	22	12	26
21	7	15	12	26
22	4	23	12	26
23	8	20	9	20
24	3	7	9	20
25	7	—	9	—
26	4	23	9	—
27	8	20	6	25
28	6	—	6	—
29	5	22	6	25
30	4	17	6	—
31	10	24	14	27
32	9	24	14	27
33	6	26	14	—
34	7	20	—	27
35	7	20	8	16
36	7	15	8	16
37	2	—	8	—
38	1	20	—	16
39	9	20	—	—
40	7	—	—	—
41	5	24	—	—
42	5	—	—	—
43	4	—	—	—
44	4	18	—	—
45	4	14	—	—
46	2	21	—	—

(video "experts"); and individuals giving a "good" or "excellent" self-assessment of their (ground-based) knowledge of Arizona vegetation (plant "experts"). Differences were sought between individuals and groups, and between airborne video "experts" and ground-based plant "experts."

Results

Participants were able to learn enough in the course of this experiment to improve significantly the accuracy of their video interpretation. Improvement was universal and dramatic. The mean individual score increased from seven correct (23.3 percent ± 11.3 percent) on the pretest to 21 correct (70 percent ± 15.1 percent) on the posttest. The mean group score increased from 11 (36.7 percent ± 11.3 percent) to 24 (80 percent ± 12.8 percent). All individuals and all groups improved their performance (Table 1). Values of KHAT reiterate the marked improvement, increasing from 0.206 for individual pretests to 0.709 for individual posttests, and from 0.358 for group pretests to 0.786 for group posttests (Table 2).

Results of pairwise comparisons of composite matrices (Table 3) allow some clear statements to be made about participants' performance. All subgroups significantly improved their identification of communities from pretest to posttest. In the pretest, groups scored higher than individuals, individuals having some previous exposure to airborne videography scored higher than those without, and those individuals rating themselves as having a good or excellent knowledge of Arizona plant communities scored higher than those rating their knowledge fair or poor. None of these differences persisted into the posttest; all subgroups were of statistically similar ability at the posttest. There was no difference between participants whose "expected" knowledge was rated high versus those rated low based on initial questionnaire Items 2, 3, and 4. The attrition of subjects during the experiment did not appear to affect results.

Contrary to the aggregate analysis using KHAT, the more sensitive paired t-tests indicated a generally significant difference between individual and three-person group scores on both the pretest and the posttest, but this result is somewhat misleading. Inspection of the raw scores shows that groups almost always scored about the same as their highest-scoring member did individually. Usually the group score was slightly higher than the best individual score within it but, in the case of some very high scoring individuals, it was slightly lower. Either way, for one-third or more of the participants there was no real difference between their individual and group scores. This fact is masked by the summary t-test. Moreover, the paired t-test assumes that the population of differences between scores is normally distributed and

Table 2. KHAT Values and Associated Variances Computed from Composite Error Matrices Representing Various Sets of Participants. N is Number of Respondents in Set. PRE in Matrix Code Indicates Pretest Results; POS Indicates Posttest Results

Set of Participants	Matrix	KHAT	Variance	N
had 2+ pertinent classes	YQ2PRE	0.21663	0.00037183	29
had <2 pertinent classes	NQ2PRE	0.18728	0.00054460	17
lived 5+ years in Arizona	YQ3PRE	0.18887	0.00037130	25
lived <5 years in Arizona	NQ3PRE	0.22626	0.00054256	21
outdoors often-frequently	YQ4PRE	0.23440	0.00051476	23
outdoors rarely-sometimes	NQ4PRE	0.17729	0.00037843	23
self-assessment high	YQ5PRE	0.33128	0.0020757	09
self-assessment low	NQ5PRE	0.17583	0.00023216	37
"expected" knowledge high	YEXKPRE	0.21139	0.00040440	26
"expected" knowledge low	NEXKPRE	0.19869	0.00049165	20
airborne video exposure	YVIDPRE	0.25579	0.00069061	19
no previous exposure	NVIDPRE	0.17063	0.00031042	27
groups which finished	GPFINPRE	0.35388	0.0018500	11
all groups started	GROUPPRE	0.35784	0.0015875	13
subjects who finished	INFINPRE	0.20194	0.00026177	38
all subjects started	INDIVPRE	0.20589	0.00022218	46
	YQ2POS	0.74282	0.0042210	24
	NQ2POS	0.65188	0.0047046	14
	YQ3POS	0.69329	0.0039741	20
	NQ3POS	0.72695	0.0051830	18
	YQ4POS	0.74319	0.0053313	19
	NQ4POS	0.67524	0.0038536	19
	YQ5POS	0.82565	0.0027714	06
	NQ5POS	0.68760	0.0024150	32
	YEXKPOS	0.72560	0.0042223	22
	NEXKPOS	0.68685	0.0048081	16
	YVIDPOS	0.77155	0.0073500	16
	NVIDPOS	0.66359	0.0031674	22
	GPFINPOS	0.78619	0.016660	11
	INFINPOS	0.70925	0.0022535	38

Table 3. Results of Pairwise Comparisons among Participant Subgroups on Pretest, on Posttest, and between Pretest and Posttest. See Table 2 for Group Descriptions. A Difference Is Significant (p<.05) when Z Is Greater than 1.96. Group with Higher Score Is Underlined

Comparison	Z-statistic	Significance
INDIVPRE × INFINPOS	10.12	S
GROUPPRE × GROUPPOS	3.17	S
INDIVPRE × GROUPPRE	3.57	S
INFINPOS × GROUPPOS	0.56	NS
YQ2PRE × NQ2PRE	0.97	NS
YQ3PRE × NQ3PRE	1.24	NS
YQ4PRE × NQ4PRE	1.91	NS
YQ5PRE × NQ5PRE	3.24	S
YEXKPRE × NEXKPRE	0.42	NS
YVIDPRE × NVIDPRE	2.69	S
YQ2POS × NQ2POS	0.96	NS
YQ3POS × NQ3POS	0.35	NS
YQ4POS × NQ4POS	0.71	NS
YQ5POS × NQ5POS	0.80	NS
YEXKPOS × NEXKPOS	0.41	NS
YVIDPOS × NVIDPOS	1.05	NS
YQ2PRE × YQ2POS	7.76	S
NQ2PRE × NQ2POS	6.41	S
YQ3PRE × YQ3POS	7.65	S
NQ3PRE × NQ3POS	6.62	S
YQ4PRE × YQ4POS	6.65	S
NQ4PRE × NQ4POS	7.65	S
YQ5PRE × YQ5POS	2.86	S
NQ5PRE × NQ5POS	9.95	S
YEXKPRE × YEXKPOS	7.56	S
NEXKPRE × NEXKPOS	6.71	S
YVIDPRE × YVIDPOS	5.75	S
NVIDPRE × NVIDPOS	8.36	S

free of outliers. This is certainly not true when comparing individual posttest scores to group posttest scores, so the paired t-test is not strictly valid here.

Some qualitative statements can be made about the ease or difficulty of identification of plant communities. By far the easiest community to identify was the Sonoran paloverde-saguaro mixed desertscrub; nearly all respondents named it correctly on both pretest (93 percent) and posttest (97 percent). Following this community in overall ease of identification were the ponderosa pine forest, cottonwood-mesquite riparian forest, and ponderosa pine-quaking aspen forest. The four most difficult communities to identify were Chihuahuan creosote-tarbush desertscrub, Great Basin greasewood shrubland, semidesert burroweed-mesquite disclimax, and mixed evergreen sclerophyll chaparral (without scrub oak). There was much more similarity than difference between individuals and groups and between airvideo "experts" and ground-based plant "experts" in which communities they found easy or difficult to identify. Most of the same eight communities appeared exceptional for all subgroups of subjects (Table 4). Evidently, the easiest communities were dominated by one or two distinctive tree species, and the most difficult communities by relatively indistinct shrub species.

Discussion

Four basic assumptions about participants in this study should be stated explicitly and evaluated. First, it was assumed that participants had normal visual acuity and color vision. One subject (15, Table 1) acknowledged that he was partially color blind, and reported that this did impair his ability to interpret the videotape. Second, it was assumed that participants could understand the verbal content of the training sessions, delivered in English. English was not the first language of seven respondents. While no complaints or requests for repetition were given, it is likely that at least one subject (24) was lacking in comprehension and so did not benefit fully from the training sessions. As a group, however, speakers of English as a second language scored slightly above average on the posttest. The third assumption was that participants were motivated to do their best on both pretest and posttest, not "holding back" on the pretest in order to display more improvement on the posttest. All indications were that this assumption was valid and participants were conscientious throughout the experiment. Fourth, it was assumed that participants were similar in their interpretation of terms on the initial questionnaire ("rarely," "sometimes," etc.), and were truthful in their responses. Some differences in usage undoubtedly existed, but probably not enough to alter the coarse dichotomous divisions made when creating subgroups for analysis.

Table 4. The Easiest and Most Difficult Arizona Vegetation Communities to Identify, for Four Sets of Participants. Percentage of Respondents Correctly Identifying the Community is Indicated (Pretest% - Posttest%)

	all individuals	3-person groups	video "experts"	plant "experts"
EASY	154.1211 paloverde-saguaro (93 - 97)	154.1211 paloverde-saguaro (100 - 100)	154.1211 paloverde-saguaro (100 - 100)	154.1211 paloverde-saguaro (100 - 100)
	122.622 ponderosa pine (41 - 92)	122.622 ponderosa pine (69 - 100)	122.622 ponderosa pine (42 - 100)	
	224.534 cottonwood-mesquite (33 - 95)		224.534 cottonwood-mesquite (42 - 100)	224.534 cottonwood-mesquite (56 - 100)
		122.628 ponderosa-aspen (62 - 100)		122.628 ponderosa-aspen (56 - 100)
		234.721 tamarisk (54 - 100)		
DIFFICULT	153.213 creosote-tarbush (4 - 26)	153.213 creosote-tarbush (15 - 18)	153.213 creosote-tarbush (0 - 31)	153.213 creosote-tarbush (0 - 50)
	152.171 greasewood (9 - 37)	152.171 greasewood (15 - 45)	152.171 greasewood (16 - 50)	152.171 greasewood (11 - 67)
	143.163 burroweed-mesquite (9 - 42)		143.163 burroweed-mesquite (16 - 31)	143.163 burroweed-mesquite (22 - 50)
	133.362 chaparral (9 - 34)	133.362 chaparral (8 - 45)	133.362 chaparral (5 - 56)	
		152.121 shadscale (15 - 55)		152.121 shadscale (11 - 50)

At least within the scope of this study, background variables seem to have little value as predictors of subjects' success interpreting vegetation communities from airborne videography. Because posttest results were essentially homogeneous, it seems that if a cadre of video interpreters are to be given brief training before being set to work, then their backgrounds are irrelevant. However, these results show that differences in ability do exist before training, and it is conceivable that significant differences could reappear if subjects were given more advanced training. Other variables related to education or aptitude could be more important than those examined.

Self-assessment of knowledge of Arizona vegetation communities was the best predictor of pretest success at interpreting communities from video. However, only nine subjects rated their knowledge "good" or "excellent," creating a very select group that did have superior ground-based knowledge and apparently were able to use it to identify vegetation communities from airborne videography footage. The experimenter did not define criteria on which to base a "good" or "excellent" assessment; most of those so rating themselves were graduate students or faculty having some professional interest in Arizona's native vegetation.

Previous exposure to airborne videography was the second-best predictor of success on the pretest. The 15 participants having only casual video exposure did only slightly better than others on the pretest and not significantly better on the posttest. The three participants having substantial experience with airborne videography (4, 9, and 11, Table 1) were, not surprisingly, more accurate in identifying vegetation communities during this experiment than were other participants. However, for people with a track record of airborne video interpretation, the idea of predicting success based on training loses its meaning.

The question of whether collaboration in small groups aids interpretation does not have a simple answer. For minimally trained subjects, individual performance was just as good as group performance. For untrained subjects, group scores were slightly better, on the whole, than individual scores. A majority of participants could claim better group pretest scores than individual pretest scores and these people were, in a sense, "helped" by group activity. For a substantial minority of higher-scoring individuals, however, group activity was no help at all and their group scores were the same as, or slightly lower than, their individual scores. Collaboration seemed to have a noticeable synergistic effect on four groups, but for four others, group scores were actually lower than the best individual member's score, indicating a detrimental effect of collaboration. Only the vague catchall of "group dynamics" can be offered as explanation for this difference among groups. It is entirely possible that these relationships could change with more highly trained subjects.

The determination of which vegetation communities are easiest and which are most difficult to identify is subject to some ambiguity due to the confounded effects of communities' familiarity to observers versus their distinctiveness. In order to correctly interpret a community from the videotape, subjects had to recognize a distinctive pattern or quality of some kind (perceptual recognition), then apply the appropriate vegetation community name to it (cognitive recognition). Several communities, notably tamarisk riparian scrub, tobosa grassland and Great Basin bunchgrass, were distinctive and recognizable as patterns but were not familiar to participants, so they could not be named correctly. Understandably, these were frequently missed on the pretest but easily interpreted correctly on the posttest, once the names had been attached to the images during training. Some communities, such as the semidesert burroweed-mesquite disclimax, were undoubtedly quite familiar to Tucson residents but were often misidentified, even on the posttest, because their patterns were either nondescript or similar to those of one or more other communities. Given this inherent complexity of the recognition process, communities were designated, as explained earlier, "difficult" if they were very frequently missed on both pretest and posttest, and "easy" if they were very frequently identified correctly on both.

Sun angle is one characteristic of the videography that influenced the apparent ease or difficulty of interpretation of several communities. A high sun angle (near noon) eliminates shadows and tends to "flatten" the look of trees and shrubs. A low sun angle (morning or afternoon) produces shadows on the ground that greatly aid interpretation; such species as saguaro (*Carnegiea gigantea*) and ocotillo (*Fouquieria splendens*) are readily identifiable by their shadows but virtually invisible without them. As an example, in the pretest many subjects missed the Douglas fir-mixed conifer forest because it was recorded at a high sun angle and people interpreted the tall trees as low shrubs. This forest appeared the same way in the posttest but was not often missed because people had learned in training how to gauge and compensate for sun angle effects. Similarly, scale and size of features in the videography were problematical for people during the pretest, but much less so on the posttest.

Without extensive photographic illustration, it is impossible to detail the cues used by participants to identify plant communities. General strategy involved identification of dominant species, evaluation of the complexity of species composition, and mnemonic devices to differentiate a community from one or two others of similar appearance. Analogies were used as succinct descriptions of species or communities (e.g., Douglas fir looks like green dreadlocks, Great Basin bunchgrass-scrub looks like bacterial culture in a petri dish, creosote-bursage sometimes has a pattern like magnified muscle tissue). Ponderosa pines were identifiable by their branchless lower trunks and dull, dark-green foliage. Patches of quaking aspen in ponderosa pine forest were readily apparent from their different leaf morphology, contrasting bright-green color, and sometimes-visible white trunks. Pinyon-juniper woodland was identified largely by the nearly spherical appearance of the junipers. Scrub oak (*Quercus turbinella*) was seen as irregular low patches of a distinct gray-green color. Tobosa grassland had a unique whitish or silvery color. Yucca, mesquite, and tamarisk all had a strong radial, starlike pattern, but differed in other details. Sagebrush and blackbrush shrublands showed a very uniform, overdispersed pattern and very uniform individuals, but were easy to differentiate by color and size. Sagebrush plants were larger and bluish-green, while blackbrush appeared as small dark puffs. Joshua trees were distinctive with their multiple gangly arms. Cottonwood crowns looked like textured green cumulus clouds or bubbling green water. In great contrast to cottonwoods, paloverde crowns appeared very indistinct and wispy. Sonoran interior strand was marked by expanses of bright, reflective sand surface dotted with a variety of small plants. Other identifications were made with myriad other, more subtle cues.

Interestingly, although this study was primarily an exercise in pattern recognition, and participants were urged to concentrate on retaining images in their mind's eye, all but one person relied heavily on notes when taking the posttest. Some barely glanced at the video monitor, spending virtually the entire hour perusing copious notes. At least during the early stages of learning video interpretation explored in this experiment, words seem to be more important than, or necessary for evoking, mental images.

The duration of images or concepts held in memory is, of course, limited. By the end of this experiment, subjects who had let more than 1 week elapse between training and posttest were reporting considerable difficulty remembering what they had learned.

The major limitations of this study are that it was uncontrolled, it examined only one of many possible training regimes, and it focused on a somewhat specialized target population. Further research could better define the training process, evaluate longer or different training regimes, and encompass a more extensive target population so that results could be generalized.

As an alternative to the one-group pretest-posttest design, some experimental subjects could have been assigned to a control group which received the pretest and posttest but not the "treatment" of the training sessions. This would have provided a formal internal validity to the experiment, albeit at the cost of a reduction in sample size. Internal validity is the certainty that experimental results are due solely to treatment effects and not to other factors. While a great improvement in subjects' performance surely occurred during the course of this experiment, it cannot be said with certainty whether the improvement was due to the training sessions, testing effects, or some other extraneous variable. This theoretical lack of certainty does not seem to pose any practical problem, however, in designing a training regime. The pretests could in fact be considered part of the training process and not as a separable confounding factor.

Instrumentation could have accounted for some improvement in scores if the posttest tape had been considerably easier for subjects to interpret than the pretest tape. If anything, the opposite was true because pains were taken to use the best, most clearly representative segments available for the pretest tape, and the second-best examples for the posttest tape.

Experimental mortality could have very slightly affected group test results because group composition did not remain constant from pretest to posttest. Any effect from this must be vanishingly small, however. Attrition did not affect individual results at all, as shown by the identity of analyses based on all 46 subjects and on the 38 finishers.

Selection of subjects in a completely nonsystematic way from a limited source population eliminates the possibility of generalizing these results. No claims are made for the external validity of this study; although results were clear-cut, they cannot be construed to apply to a dissimilar population of subjects. To predict the performance of different groups, further studies are needed in which experimental subjects are selected randomly from well-defined source populations.

Conclusions

This study has shown that a minimal training program can effectively increase the accuracy of observers' interpretation of vegetation communities from airborne videography. Three hours of training significantly improved the performance of all individuals tested, regardless of background, and of all small groups. Before training, groups performed measurably better, on the whole, than individuals; participants with some previous airborne video exposure performed better than those without; and subjects rating their own knowledge of Arizona vegetation "good" or "excellent" did better than those rating their knowledge "fair" or "poor." Three other background variables had no significant effect. After training, these few differences had disappeared and all respondents were of statistically similar ability. Thus, within the scope of this experiment, personal background and the opportunity to collaborate in small groups do not seem to be important factors in selecting and training video interpreters.

The most difficult communities for subjects to identify were Chihuahuan creosote-tarbush desertscrub, Great Basin greasewood shrubland, semidesert burroweed-mesquite disclimax, and mixed evergreen sclerophyll chaparral (without scrub oak). The communities most easily identified were Sonoran paloverde-saguaro mixed desertscrub, ponderosa pine forest, cottonwood-mesquite riparian forest, and ponderosa pine-quaking aspen forest.

The findings of this study should be encouraging to any group planning an investment in airborne videography as a resource management tool, and the assignment of personnel to interpret vegetation features from video footage. Given a suitable key (teacher), competent operators can be trained in a very short time. Once trained, multiple interpreters can be assigned to work independently on the same project with the knowledge that their contributions will be essentially homogeneous, at least for a short term. It is not necessary to have a team of interpreters working on the same footage, and specialists are not required for the bulk of video interpretation. For all of these reasons, airborne videography is likely to be the least-cost option for conducting many resource inventory and monitoring projects.

Acknowledgments

The author wishes to thank all study participants for their generous effort, and gratefully acknowledges the fundamental contributions of Drs. Lee A. Graham and Guy R. McPherson. Facilities and assistance were provided by the Advanced Resource Technology Program at the University of Arizona and the Arizona Remote Sensing Center.

References

Argialas, D.P., and C.A. Harlow, 1990. Computational image interpretation models: An overview and a perspective, *Photogrammetric Engineering & Remote Sensing*, 56(6):871–886.

Avery, T.E., 1985. *Interpretation of Aerial Photographs, Fourth Edition*, Burgess Publishing Co., Minneapolis, 324 p.

Brown, D.E. (editor), 1982. Biotic communities of the American southwest-United States and Mexico, *Desert Plants* (special issue), 4(1–4):1–342.

Campbell, D.T., and J.C. Stanley, 1963. *Experimental and Quasi-Experimental Designs for Research*, Houghton-Mifflin Co., Boston, 84 p.

Colwell, R.N., 1946. The estimation of ground conditions from aerial photographic interpretation of vegetation types, *Photogrammetric Engineering*, 13(2):151–161.

———, 1948. Aerial photographic interpretation of vegetation for military purposes, *Photogrammetric Engineering*, 14(4):472–481.

Congalton, R.G., R.G. Oderwald, and R.A. Mead, 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques, *Photogrammetric Engineering & Remote Sensing*, 49(12):1671–1678 (N.B. Erratum, 50(10):1477).

Congalton, R.G., and R.A. Mead, 1983. A quantitative method to test for consistency and correctness in photointerpretation, *Photogrammetric Engineering & Remote Sensing*, 49(1):69–74.

Estes, J., E. Hajic, and L. Tinney, 1983. Manual and digital analysis in the visible and infrared regions, *Manual of Remote Sensing, Second Edition*, American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia, 1:987–1123.

Graham, L.A., 1993. Airborne video for near real-time vegetation mapping, *Journal of Forestry*, 91(8):28–32.

Julesz, B., 1975. Experiments in the visual perception of texture, *Scientific American*, 232(4):2–11.

Kalliola, R., and K. Syrjanen, 1991. To what extent are vegetation types visible in satellite imagery? *Ann. Bot. Fennici*, 28:45–57.

Klingberg, C.L., C.L. Kraft, and C.L. Elworth, 1963. *A Study of Photointerpreter Performance in Change Discrimination*, Rome Air Development Center, Air Research and Development Command, U.S.A.F., Griffiss Air Force Base, New York, Task #624402.

Loelkes, G.L., 1983. *Land Use/Land Cover and Environmental Photointerpretation Keys*, U.S. Geological Survey Bulletin No. 1600, 142 p.

Ott, L., 1988. *An Introduction to Statistical Methods and Data Analysis, Third Edition*, PWS-Kent Publishing Co., Boston, 945 p.

Quinn, A.O. (editor), 1947. Symposium of ideas relative to education in photogrammetry, *Photogrammetric Engineering*, 13(3):331–394.

(Received 4 March 1994; revised and accepted 11 January 1995; revised 13 March 1995)

Appendix A

INITIAL QUESTIONNAIRE

1) What is your major? ALL RELATED TO BIOLOGY OR NATURAL RESOURCES

2) How many classes have you taken in plant identification or community ecology? ______

CLASSES:	0-1	2	3	4	5+
# PARTICIPANTS:	17	10	8	6	5

3) How long have you lived in Arizona? ______ years ______ months

YEARS:	0-2	2-3	3-5	5-10	10+
# PARTICIPANTS:	10	7	4	10	15

4) Do you go hiking or do other outdoor recreation:

 4 rarely, 19 sometimes, 14 often, or 9 frequently?

5) Would you say your knowledge of Arizona vegetation communities is:

 11 poor, 26 fair, 7 good, or 2 excellent?

6) Are you a private pilot or frequent passenger in light aircraft? 4 yes 42 no

7) Have you had any previous exposure to airborne videography?

 18 yes 28 no If yes, describe briefly.

ONLY 3 SUBJECTS HAD SUBSTANTIAL EXPERIENCE (>50 HRS) VIEWING AIRBORNE VIDEOGRAPHY FOR INTERPRETATION. THE OTHER 15 ANSWERING 'yes' HAD CASUAL EXPOSURE OF A FEW HOURS AT MOST.

Appendix B

VEGETATION COMMUNITY LIST

code	community name
122.414	Great Basin Pinyon-Juniper Woodland
122.612	Madrean Douglas Fir-Mixed Conifer Forest
122.622	Madrean Ponderosa Pine Forest
122.628	Madrean Ponderosa Pine-Quaking Aspen Forest
123.311	Madrean Encinal Oak Woodland
133.314	Interior Scrub Oak-Mixed Sclerophyll Chaparral
133.362	Interior Mixed Evergreen Sclerophyll Chaparral
142.214	Great Basin Bunchgrass-Scrub
142.222	Great Basin Grass-Scrub
143.121	Semidesert Tobosa Grassland
143.151	Semidesert Mixed Grass-Yucca
143.152	Semidesert Mixed Grass-Mesquite
143.163	Semidesert Burroweed-Mesquite Disclimax
152.111	Great Basin Sagebrush Shrubland (*Artemisia tridentata*)
152.121	Great Basin Shadscale Shrubland (*Atriplex confertifolia*)
152.131	Great Basin Blackbrush Shrubland (*Coleogyne ramosissima*)
152.171	Great Basin Greasewood Shrubland (*Sarcobatus vermiculatus*)
153.1221	Mojave Blackbrush-Joshua Tree Desertscrub
153.213	Chihuahuan Creosote-Tarbush Desertscrub
154.1112	Sonoran Creosote-Bursage Desertscrub
154.1116	Sonoran Creosote-Bursage/Paloverde-Mixed Desertscrub
154.118	Sonoran Brittlebush-Creosote Desertscrub
154.121	Sonoran Paloverde-Mixed Desertscrub
154.1211	Sonoran Paloverde-Saguaro Mixed Desertscrub
154.1215	Crucifixion Thorn-Juniper-Paloverde Desertscrub
224.521	Sonoran Riparian Mesquite Forest ("Bosque")
224.531	Sonoran Riparian Cottonwood-Willow Forest
224.534	Sonoran Riparian Cottonwood-Mesquite Forest
234.721	Sonoran Tamarisk Riparian Scrub
254.711	Sonoran Interior Strand, Mixed Scrub Species

Appendix C

Included are two composite error matrices, representing individuals' pretest and posttest responses. For each matrix, reference data lie along the horizontal axis and respondents' interpretations lie along the vertical axis; correct responses are found on the major diagonal (shaded). Plant community codes are the same as in Appendix B. Cell values are number of respondents giving each interpretation. Blank cells indicate no responses. Row and column sums are given for reference, but note that the value in the (shaded) lower right corner is the sum of the diagonals, not the sum of the marginal totals.

INFINPRE - pretests of 38 finishers; KHAT = 0.20194

reference data →

interpretations ↓

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	aa	ab	ac	ad	sum
A	122.414	10		2	2	3	5							1	2	1	1	1		1			3			2			4			38
B	122.612		2	6	6		2																					1	2		1	20
C	122.622		1	18			2																				2					23
D	122.628		2	5	6		1																				3					17
E	123.311	9	7	3	4	10	4	2			1			1							1		2			1	1		1			47
F	133.314	1	10	1	5	8	9	4	1				1	3	4		1	3		1		1	3						1			57
G	133.362	1	2	1	8	2	7	4							3			5			1			2		1					1	38
H	142.214								5	6	3				1	3	1		4	3	1	1										28
I	142.222							3	4	6	4	6	4		1	2			1	3		1	2							1		38
J	143.121								5	13	9	3				1		1					1									33
K	143.151									3	6	4				1	1		1		1									2		19
L	143.152			1		3				2	1	6	5	3	1	1	1		4	3	2	4	1	1								39
M	143.163				1	1		1				2	4	3	1		2	1	5	3	3		1			2				2	1	33
N	152.111	1						1	2	1	4	2	1	1	8	5	2	4		3	1	1									2	39
O	152.121			1					4	1	3	3			2	2	4	2		2			2							1		27
P	152.131	2							3		1		3		3	4	8	3		3	2									1		33
Q	152.171	1	2				1		1	2	1				4	1	1	3		2	1	1	2							2	3	28
R	153.1221							1	1			1		1		3	2	1	1			1	2	1		9				5		29
S	153.213	2	2				1	2	3			5	2	2		5	3	3		2	4	6	4			1						47
T	154.1112	4				1	1	2	4		1	1	1	4	2	5	4	4	2	3	6	3	1			1						50
U	154.1116							2		1			2	3			1		6	2		11	4	1	1	12					1	47
V	154.118	2	1					2	3		3	2	4	2	1	1	3		2	1	4	1				1				4	1	38
W	154.121	3			1	3	1	5	2	2		1	6	4	3			1	10		5	5	3	14	1	2				1	2	75
X	154.1211				1		1						1		1										36						1	41
Y	154.1215	1	2				2	4			1		2	1	1			1	1			2	3	4		5				2		32
Z	224.521		1			2		2						4							1			1			17	3	11	3	2	47
aa	224.531		1			1										1		1									8	21	3	1		37
ab	224.534		1		4	3							1	3			1			1	1		1	2			1	10	13	1		43
ac	234.721	1	2										1	1				1		3				7			5	3	3	6	8	41
ad	254.711		1			1		2		1		1				2	2	3	1	1	4		2	4		1	1			4	14	45
	sum	38	37	38	38	38	37	37	38	38	38	37	38	37	38	38	38	38	38	37	38	38	37	37	38	38	38	38	38	36	37	258

Appendix C, continued

INFINPOS - posttests of 38 finishers; KHAT = 0.70925

reference data →

interpretations ↓		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	aa	ab	ac	ad	sum
A	122.414	30				6	1																									37
B	122.612		35	1	2	3		1																			3					45
C	122.622	1		35		5																										41
D	122.628			2	34	1																										37
E	123.311	2	2		1	18	3					1											1	1			3					32
F	133.314	2	1		1	3	24	1					2														1					35
G	133.362					2	5	13						1	1	1	2	1						2			1					29
H	142.214								37							1	1															39
I	142.222									30			2			6	4	2					1									45
J	143.121									3	37					2				1												43
K	143.151							2				28	1			1			1	1				1								35
L	143.152												25	3		1	2	1		1	1			1								35
M	143.163						1	1					3	16			2	1		1		4		5								34
N	152.111	1					1	1					1		24	3	1	7		3												42
O	152.121									1						22	1	3		4	1		2									34
P	152.131														5		19	7		8		2									1	42
Q	152.171						2	2		2		1		3	3		3	14		8	1						2					41
R	153.1221											1							33							1						35
S	153.213							3	1				2		4	1	2	1		10		3	1	3								31
T	154.1112							2						2							33	3		1							1	42
U	154.1116							2						2					2	1		21		7	1			1				37
V	154.118							3		2					1		1					1	28	1								37
W	154.121							6					2	9					1		1	3	5	12		1						40
X	154.1211													1										1	37							39
Y	154.1215	1						1				5							1					2		36				1		47
Z	224.521						1																				25		1		1	28
aa	224.531																							1				33			1	35
ab	224.534																										1	3	36			40
ac	234.721											2		1														1		37		41
ad	254.711										1											1							1		34	37
	sum	37	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	37	38	38	37	38	38	38	38	38	36	38	38	38	38	1116

FREE Subscription to ImageNet

New members to ASPRS receive
a Free 1-year subscription to
ImageNet, Core Software's
on-line geographics data search and retrieval system.

You can join in 5 minutes over the phone. Call 301-493-0290.

Or contact:
ASPRS Membership
5410 Grosvenor Lane, Suite 210
Bethesda, MD 20814-2160
fax 301-493-0208 members@asprs.org www.asprs.org/asprs

A $500 Value!