A Class of Algorithms for Enumerating Rare Objects Using Spectral and Spatial Data in Real Time

David R. Anderson
Colorado Cooperative Fish & Wildlife Research Unit, Colorado State University, Fort Collins, CO 80523
Clair L. Wyatt
Space Dynamics Laboratory, Utah State University, Logan, UT 84322
Mohan M. Trivedi
Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803
*Roger K. Harshbarger**
Space Dynamics Laboratory, Utah State University, Logan, UT 84322

ABSTRACT: Airborne multispectral linear array systems for the remote detection and enumeration of rare objects must often employ specialized classification algorithms that allow fast data processing and accurate discrimination. In some cases, the spectral information from several bands is inadequate to classify accurately the object of interest from its background. This limitation leads to objects of interest being undetected and the occurrence of false counts (errors of commission). If the *a priori* probability of occurrence of the object of interest is low, the probability of a false count must be extremely small or the estimated number of objects will be greatly overestimated. Pixel mixtures frequently result in significant errors in some classification methods.

A class of multistage classification methods based on capture-recapture sampling theory is presented to cope with the problems and requirements outlined above. Ratios of the radiant flux in two spectral bands are the primary basis for classification. Each of these ratios is compared to a threshold value to discriminate the object of interest from its background. The threshold values are set conservatively to avoid false counts. However, a substantial fraction of the objects of interest then remain undetected by each of the ratio classifiers. Capture-recapture theory is used to estimate the population size of the objects of interest from these incomplete counts. The classification procedure is illustrated with data from the results of a Monte Carlo study and with data from a mule deer detection program. These studies indicated that a class of jackknife estimators performs well, especially when heterogeneity among the pixels and dependence among the classifiers is significant and detection probabilities are above about 0.3 or 0.4. Relative bias was less than one percent when three classifiers were simulated and achieved confidence interval coverage was at the nominal level. It is essential to avoid false counts when using this classification system.

Selection of spectral bands is addressed and a simple algorithm based on spatial information is given to further reduce the number of false counts. The entire algorithm can be implemented with currently available electronic components and can be operated in real time, as demonstrated with prototype system of field tests.

INTRODUCTION

SOME REMOTE SENSING APPLICATIONS are concerned with correctly classifying relatively small objects of interest and counting them by means of instrumentation aboard an aircraft. Examples might include aerial surveys for biological populations like caribou, deer, kangaroos, or polar bears as well as many nonbiological inventories. These applications involve several common problems that must be overcome before success can be achieved.

A high degree of accuracy is not essential in some survey programs. In biological surveys an accuracy of ± 15 percent is often satisfactory. These objectives are difficult to obtain because the *a priori* probability of the object of interest may be on the order of 10^{-5}. Therefore, unless false counts are essentially eliminated, say a rate on the order of 10^{-6}, the count will be far too large. Error types and terminology are illustrated by the following table:

* Presently with Space Data Corporation, 1333 West 21st Street, Tempe, AZ 85282.

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 52, No. 8, August 1986, pp. 1159–1170.
0099-1112/86/5208-1159$02.25/0

		True State	
		Object of interest	All other objects
Classification Outcome	Object of interest	Correct classification	Error of commission (False count)
	All other objects	Error of omission	Correct classification

The classification algorithm must be very fast because real time data processing is often required due to the high data rates and small pixel size. The large amount of data prohibits storage in the aircraft until the data can be analyzed at a later time on the ground.

Various classification methods can be unsatisfactory due to a common and unavoidable measurement problem that is herein referred to as "pixel mixtures." For example, a given 10-cm by 10-cm pixel contains a mixture of snow and sagebrush and might be incorrectly classified as "deer." The system is designed to correctly classify snow as "nondeer" and sagebrush as "nondeer." It is certain mixtures of the snow and sagebrush scene where classification errors are made. In addition, a pixel might be only partially "deer," giving rise to classification problems.

Often, data in three or four spectral bands are not sufficient for accurate classification and enumeration. This is true for certain scenes, even without pixel mixtures. If a system could use information in 8 to 12 spectral bands and limit high atmospheric absorption spectral regions, perhaps the spectral data would be sufficient for accurate classification. However, current technology often prevents the development of such systems and, therefore, spectral information can be used with simple spatial information to enhance correct classification. The spatial data available are the classification outcomes for adjacent pixels.

The purpose of this paper is to explain and illustrate a series of classification algorithms based on capture-recapture theory (Otis *et al.*, 1978) that are very fast, are based on a simple method of selecting spectral bands that minimizes the problem of misclassifying pixel mixtures, are accurate in the face of imperfect spectral discrimination, incorporate spatial data, and are appropriate for surveying populations of small objects that are relatively very rare.

These problems and solutions are illustrated by an example involving mule deer (*Odocoileus hemionus*) on winter ranges in the western United States and by means of a Monte Carlo study. Based on these results, the classification algorithm described appears promising for a fairly broad class of remote sensing applications where enumeration is the objective.

BACKGROUND

The classification procedure described is illustrated by ongoing efforts to develop a system to enumerate mule deer. Trivedi *et al.* (1982) presented an optimal single-stage Bayes classifier and the results of controlled field tests relating to a multispectral deer detection system. Classification error rates were estimated and several different *a priori* probabilities were considered. Trivedi *et al.* (1984) presented a sequential or multistage classification algorithm as a more practical and equally effective alternative. The results of Trivedi *et al.* (1982, 1984) support the use of ratios of the radiant flux in two spectral bands as the basis of classification. Such ratios are easy to implement electronically, and classification of them can be based on a comparison with a threshold to discriminate the classes "deer" versus "nondeer."

In view of these results, Voorheis (1982) developed a prototype electro-optical system to be used in a wide variety of controlled tests from the ground and from aircraft. The prototype instrument employs four 1728-element detector arrays totaling 6912 detectors. Four spectral bands are used (0.725, 0.764, 0.863, and 0.981 μm, all $\pm$ 0.01). The system resolves a 10-cm by 10-cm pixel at about 445 m above the ground (additional details can be found in Voorheis (1982)).

Experiments with this instrument and the single-stage and multi-stage classification algorithms were encouraging, but revealed several problems. Spectral data in three or four bands were not sufficient for accurate discrimination into the two classes "deer" and "nondeer." False counts could not be totally eliminated, and because of the overwhelming majority of the "nondeer" class, gave greatly exaggerated counts. Deer were sometimes undetected by the instrument (an error of omission). Finally, the pixel mixture problem was found to be serious and resulted in many false counts.

TABLE 1. RATIOS OF FLUX MEASUREMENTS OF THE SIX SPECTRAL BANDS SATISFYING THE BOUNDARY CONDITION. INVERSES ARE NOT SHOWN AS THEY CONTAIN SIMILAR INFORMATION (DIFFERENCES IN THE COEFFICIENT OF VARIATION <3%). FROM TRIVEDI *et al.* (1984).

		Ratio of radiant flux measurements in two bands					
Class		[0.672]/[0.603]	[0.725]/[0.603]	[.764]/[0.603]	[0.863]/[0.603]	[0.981]/[0.603]	[0.725]/[0.672]
Deer	Mean	1.73	2.36	2.63	2.80	1.49	1.36
	Coeff. Var. %	9.86	14.22	16.18	19.62	24.75	8.09
Snow	Mean	1.41	1.54	1.52	1.26	0.48	1.09
	Coeff. Var. %	4.18	5.52	5.71	6.51	18.31	2.73
Juniper	Mean	1.21	6.71	7.89	7.39	3.04	5.57
	Coeff. Var. %	9.07	10.96	11.15	12.57	20.20	14.45
Sagebrush	Mean	1.59	2.65	2.92	2.82	1.32	1.67
	Coeff. Var. %	5.86	10.17	12.00	14.40	21.66	7.27
Rabbit-brush	Mean	1.62	2.30	2.50	2.37	1.06	1.42
	Coeff. Var. %	7.41	9.11	10.39	13.14	20.53	6.33

		Ratio of radiant flux measurements in two bands								
Class		[0.764]/[0.672]	[0.863]/[0.672]	[0.981]/[0.672]	[0.764]/[0.725]	[0.863]/[0.725]	[0.981]/[0.725]	[0.863]/[0.764]	[0.981]/[0.764]	[0.981]/[0.863]
Deer	Mean	1.51	1.61	0.86	1.11	1.18	0.63	1.06	0.56	0.53
	Coeff. Var. %	10.00	12.89	19.13	3.27	7.72	15.87	5.61	14.40	10.66
Snow	Mean	1.08	0.89	0.34	1.00	0.82	0.31	0.83	0.32	0.38
	Coeff. Var. %	2.97	4.82	18.64	1.50	3.64	18.27	3.36	18.10	15.61
Juniper	Mean	6.56	6.13	2.53	1.18	1.10	0.45	0.94	0.38	0.41
	Coeff. Var. %	14.63	15.39	22.15	2.37	4.68	13.89	3.82	13.23	10.49
Sagebrush	Mean	1.84	1.77	0.83	1.10	1.06	0.45	0.96	0.49	0.46
	Coeff. Var. %	8.87	11.18	19.28	2.74	5.35	13.00	3.63	14.01	10.85
Rabbit-brush	Mean	1.55	1.46	0.65	1.09	1.03	0.46	0.95	0.42	0.44
	Coeff. Var. %	7.36	8.88	16.70	2.63	5.72	15.06	4.20	14.25	11.06

APPROACH

AVOIDING PIXEL MIXTURE ERRORS

A method of avoiding large classification errors due to pixel mixtures is the proper selection of spectral bands. Bands to be selected must be those where reflectance ratios for deer are higher (or lower) than reflectance values for all nondeer classes (i.e., a boundary constraint is enforced). Mixing of the nondeer classes does not cause errors if we consider the ratios of the spectral bands that satisfy this boundary condition. The most common class expected to be mixed with other nondeer classes is snow in wintertime surveys. Such mixtures often improve the discrimination between deer and nondeer classes. Average scans of the data collected by Trivedi (1979) were examined for spectral bands that met this boundary constraint. Six spectral bands were identified: 0.603 μm, 0.672 μm, 0.725 μm, 0.764 μm, 0.863 μm, and 0.981 μm. These six bands allow calculation of $\binom{6}{2} = 15$ ratios (Table 1); however, only six of the 15 ratios satisfy the boundary constraint: [0.672]/[0.603], [0.981]/[0.863], [0.863]/[0.725], [0.981]/[0.725], [0.863]/[0.764], and [0.91]/[0.764] (in addition to inverses of these six). The notation above expresses the value of a spectral measurement corresponding to wavelength *i* as [*i*]; thus, [*i*]/[*j*] would represent the ratio of measurements made in the spectral bands corresponding to wavelengths *i* and *j*, respectively.

Mixed Spectral Data. Approximately 2,000 spectral signatures involving 105 classes (e.g., snow; bare ground; grass, shrub, and tree species; rocks) were collected near Logan, Utah during the winter of 1981 (Trivedi *et al.*, 1984). A large data base was generated by purposely mixing these spectral reflectance data to gain information on the problem of pixel mixtures. In general, the various classes were mixed with snow (the typical case on a winter range for mule deer). Specifically, the data base of mixtures was generated for each spectral band using Equation 1 (see Conners *et al.*, (1984) 288–289).

$$\alpha Y_i + (1-\alpha)\, Y_j \qquad [1]$$

where α = mixing coefficient $0 < \alpha < 1$ values in the increments of 0.1 were considered.
In Equation 1,

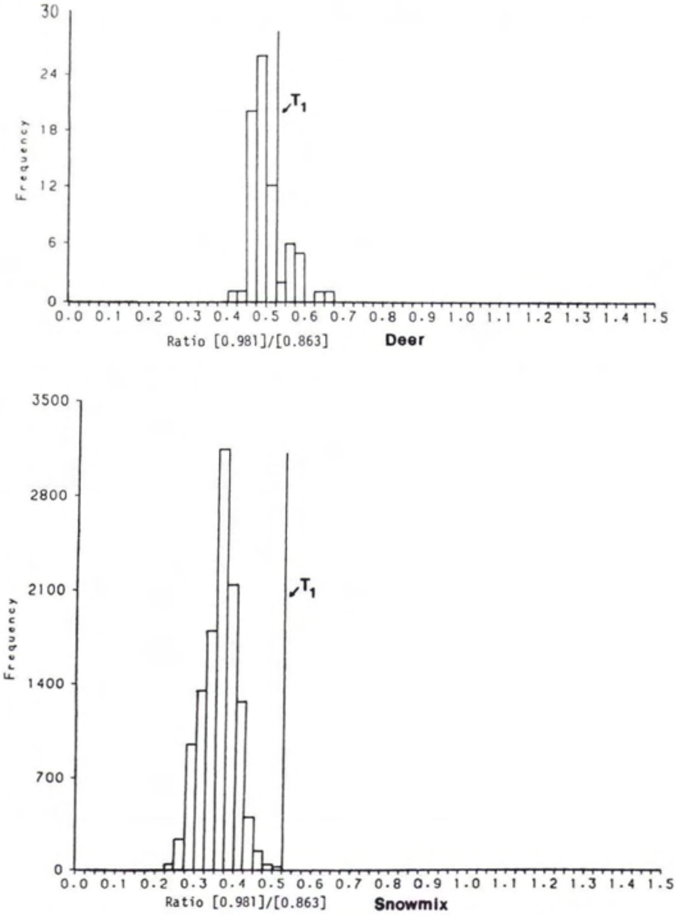

FIG 1. Histograms of the ratio of spectral reflectance in two wavelength bands [0.981]/[0.863] for mule deer (top) and 105 nondeer classes mixed with varying proportions of snow (bottom). A threshold T_1 is shown and if the ratio $> T_1$ the scene is classified as deer. This rule avoids false counts but fails to correctly classify most of the deer.

Y_i = snow reflectance data for a particular spectral band, and
Y_j = some other class that would "mix" with snow to form a pixel mixture.

Because nearly all other classes could be expected to be measured with snow as a pixel mixture, the new data base was relatively large (11,466 records) and was used to investigate classifier performance with pixel mixtures.

Formulating the Decision Rules. Histograms of "deer" versus "nondeer" for each of the six ratios of spectral reflectance were prepared (the best three are shown in Figures 1 to 3). While these seem to represent the best ratios of radiant flux, there is considerable overlap between the histograms of deer and nondeer. Considering the very high *a priori* probability of nondeer (on the order of $1\text{-}10^{-6}$), it is clear that the probability of a false count must be essentially zero or very serious overestimation will occur. From Figure 1, if threshold T_1 is established at 0.525, the nondeer class is always $< T_1$. A given pixel is then classified as "deer" if the value of the ratio for spectral bands $[0.981]/[0.863] > T_1$. When this occurs, a counter c_1 is incremented by one. While this classifier avoids false counts, it also fails to count correctly about 4/5 of the deer (large errors of omission) (see Figure 1) and is hardly satisfactory. A similar argument can be made for ratios [0.764]/[0.981] and [0.981]/[0.725] (Figures 2 and 3), although the classification rule is reversed for [0.764]/[0.981].

In summary, the problem of pixel mixtures can be avoided if only ratios of spectral reflectance are chosen with the boundary constraint. In this way, mixing certain classes actually allows further separation of the distributions of deer and nondeer. Ratios of reflectance were chosen to allow the greatest separation of the distributions (all possible histograms and linear discriminate functions were examined to select the

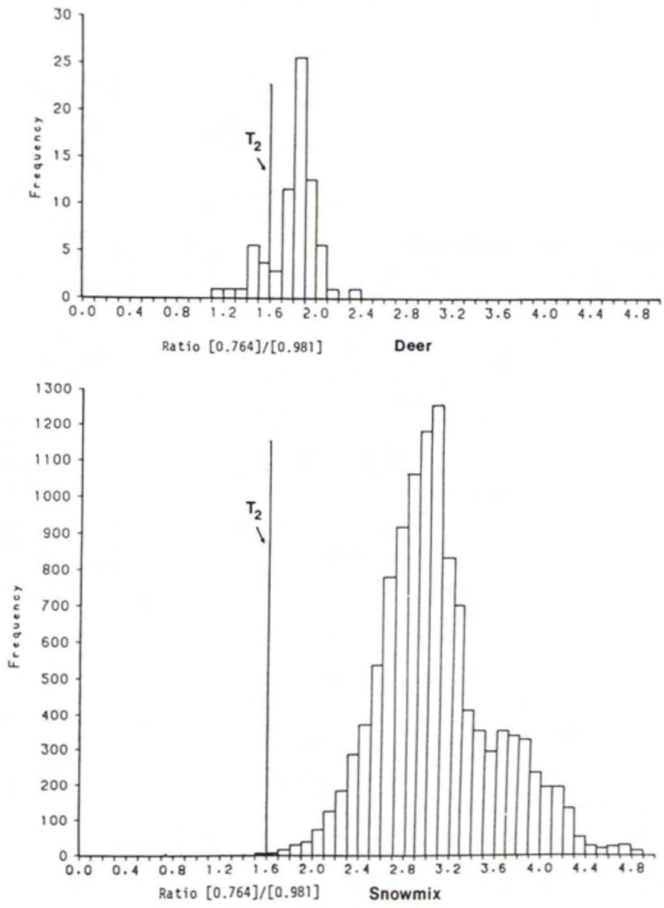

FIG. 2. Histograms of the ratio of spectral reflectance in two wavelength bands [0.764]/[0.981] for mule deer (top) and 105 nondeer classes mixed with varying proportions of snow (bottom). A threshold T_2 is shown and if the ratio $< T_2$ the scene is classified as deer. This rule avoids most false counts but fails to correctly classify most of the deer. A smaller T_2 would be required to lessen the number of false counts.

best ratios). These considerations and the information in Figures 1 to 3 lead to the following simple decision rules for a given pixel:

If ratio [0.981]/[0.863] > 0.525, classify "deer" and increment the counter c_1 (0.525 is a threshold, T_1);
If ratio [0.764]/[0.981] < 1.60, classify "deer" and increment the counter c_2 (1.60 is a threshold, T_2); and
If ratio [0.981]/[0.725] > 0.70, classify "deer" and increment the counter c_3 (0.70 is a threshold, T_3).

In each case, the counts c_1, c_2, and c_3 are very incomplete (negatively biased) because they are well below the actual number of deer N.

ESTIMATING THE NUMBER OF DEER FROM INCOMPLETE COUNTS

Consider a system employing three ratios of four spectral bands and three simple threshold classification rules as presented. We are left with three counts c_1, c_2, and c_3, all of which are less than the true total N. We can say that the counts c_i are incomplete. Magnusson *et al.* (1978) discussed the use of Petersen's (1896) method to estimate the total number (N) in a population from two incomplete counts. The example given dealt with crocodile nests being counted and mapped by two airborne observers. They define c_1, c_2, and n_{12} (our notation) as the number of nests seen by observer 1, observer 2, and both, respectively. An estimate of the total N is (from Chapman 1951)

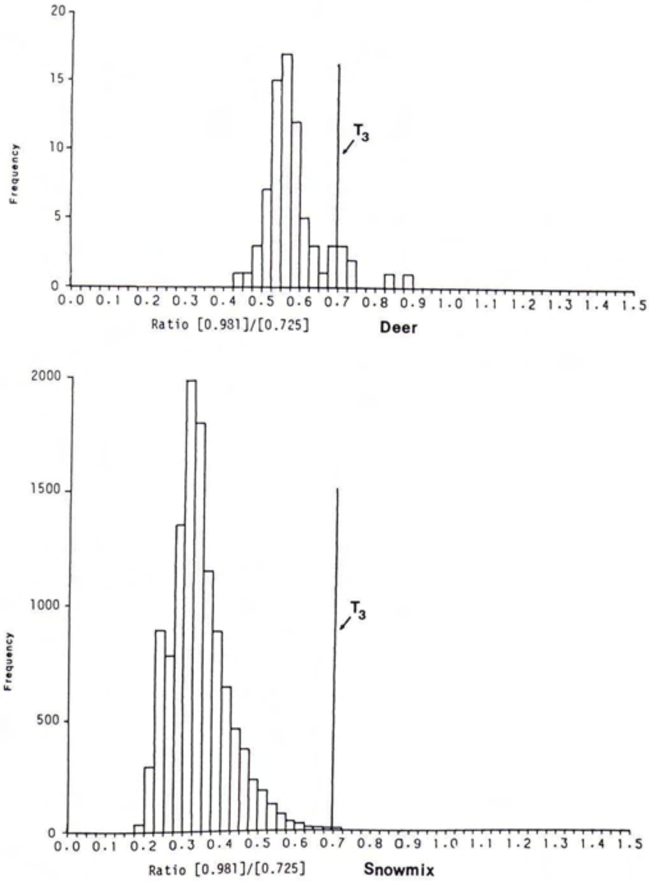

FIG. 3. Histograms of the ratio of spectral reflectance in two wavelength bands [0.981]/[0.725] for mule deer (top) and 105 nondeer classes mixed with varying proportions of snow (bottom). A threshold T_3 is shown and if the ratio $> T_3$ the scene is classified as deer. T_3 must be larger to avoid all false counts, thereby giving an incomplete count of deer.

$$\hat{N} = \frac{(c\ + 1)(c\ + 1)}{(n_{12} + 1)} - 1.$$

More recently, Maxim *et al.* (1981) dealt with a more extended example involving two or three observers inspecting photographic images. They, like Magnusson *et al.* (1978), realized that the assumption that detection among observers must be independent, and they further emphasized that there must be no false counts. They made an empirical assessment of the assumption of independence and attempted to deal with a model allowing some dependence. Both papers fully recognized that the general problem falls under what is called capture-recapture sampling theory.

The problem of dependence between the two observers carries over to dependence among electro-optical sensors and mathematical classification algorithms. If the first observer fails to see a crocodile nest because it is somewhat hidden or partially shaded, then the second observer will also tend to fail to see the same nest, for the same reasons. Conversely, a large nest on a slight rise in open terrain will have a marked tendency to be seen by both observers. The principle of dependence is to be expected due to the *heterogeneity* of the objects of interest and their spectral information. Heterogeneity is defined here as the collection of differences in the spectral information for the objects of interest. Conceptually, there exists a probability density function (pdf) of this variability (see Burnham and Overton (1978) for the analogy of heterogeneity in capture probabilities in animal trapping studies). Objects of interest exhibit heterogeneity, and this gives rise to dependence among the classifiers.

Capture-recapture theory now includes a rigorous estimation theory for a wide class of models (Otis *et* *al.*, 1978; Seber, 1982; White *et al.*, 1982). This theory is reviewed briefly and cast in a remote sensing context. The application covers two or more observers or some sort of detector or electro-optical sensor system with two or more outputs. The term classifier will be used be it a person, a detector plus a classifier, or a result of a mathematical algorithm. Consider the number of classifiers $j = 1, 2, \ldots, t$ (where $t \geq 2$). Conceptually, the pixels in the class of interest can be numbered from 1, 2, . . . , to the last pixel N_p; that is, $i = 1, 2, \ldots, N_p$. The detection history of each deer pixel by each classifier can be expressed conveniently in a matrix **X**.

Let

$$[X_{ij}] = \begin{vmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1t} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2t} \\ X_{31} & X_{32} & X_{33} & \cdots & X_{3t} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ X_{n.1} & X_{n.2} & X_{n.3} & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ X_{N_p} & X_{N_p} & X_{N_p} & \cdots & X_{N_p} \end{vmatrix} \qquad [2]$$

where $X_{ij} = \begin{cases} 1 \text{ if the } i^{\text{th}} \text{ deer pixel is classified as ``deer'' by the } j^{\text{th}} \text{ classifier, or} \\ 0 \text{ otherwise.} \end{cases}$

The **X** matrix contains only zeros and ones, indicating whether a particular deer pixel is "not classified as deer" and "classified as deer," respectively. Each column details the history of classification as deer for the j^{th} classifier ($j = 1, 2, \ldots, t$). Therefore, a row in this matrix has at least one l (i.e., deer) in it, and may have as many as t ls in it (i.e., all t classifiers classified the particular pixel as a deer). The first $n.$ rows of the matrix relate to the classification history of each deer pixel classified as deer by at least one classifier. The remainder of the **X** matrix contains all zeros, because these deer pixels were never classified deer. Of course, one does not know how many remaining rows there are (because N_p is unknown). Note that, because deer are larger than the 10-cm by 10-cm pixel, we must first estimate N_p, the total number of pixels that are "deer" and later estimate the number of deer N.

Detection Statistics. Several statistics are computed from the **X** matrix for use in defining estimators of $\mathrm{N_p}$.

c_j = the total number of deer pixels classified as deer by the j^{th} classifier, $j = 1, 2, \ldots, t$ (where $t \geq 2$). c_j is merely the sum of the j^{th} column of the matrix **X**.

n_w = the number of deer pixels classified as deer for a particular classifier pattern, w. Examples will make this clear. n_2 is the number of deer pixels classified as deer by *only* the second classifier. n_{13} is the number of deer pixels simultaneously classed as deer by only the first and third classifiers. In a four classifier system n_{1234} is the number of deer pixels simultaneously classed as deer by all four classifiers. n_w is a set of counters.

f_j = The classification frequencies: the number of deer pixels classified as deer exactly j times by the t classifiers, $j = 0, \ldots, t$. For example, f_2 = number of deer pixels that were classified as deer by two classifiers. The statistic f_o is not observable for if we know f_o we would know N_p. Note,

$$\begin{aligned} f_1 &= n_1 + n_2 + \ldots + n_t; \\ f_2 &= n_{12} + n_{13} + n_{23} + \ldots + n_{t-1,t}; \\ f_3 &= n_{123} + n_{124} + \ldots + n_{t-2,t-1,t}; \\ &\ \ \vdots \\ f_t &= n_{12\ldots t}. \end{aligned}$$

$n.$ = the number of different deer pixels classified as deer by at least one of the t classifiers during the survey. This statistic is merely the number of nonzero rows in the **X** matrix. Of course, if N_p - $n.$ were known, the problem would be solved.

Estimators for N_p, $\hat{N}_p$. A variety of estimators of $\mathrm{N_p}$, denoted as $\hat{\mathrm{N}}_\mathrm{p}$, exists depending on the model assumed. In general, if it is assumed that classifiers are independent, the Chapman (1951) and Darroch (1958) estimators are appropriate:

For $t = 2$

$$\hat{N}_p = \frac{(c_1 + 1)(c_2 + 1)}{n_{12}} - 1$$

$$\hat{\mathrm{var}}\,(\hat{N}_p) = \frac{(c_1 + 1)(c_2 + 1)(c_1 - n_{12})(c_2 - n_{12})}{(n_{12} + 1)^2\,(n_{12} + 2)}$$

For $t > 2$ $\hat{N}_p$ is the solution of

$$1 - \left[\frac{M_{t+1}}{\hat{N}_p}\right] = \prod_{j=1}^{t}\left[1 - \frac{c_j}{\hat{N}_p}\right]$$

$$\mathrm{var}\,(\hat{N}_p) = \hat{N}_p\left[\frac{1}{\prod_{j=1}^{t}(1-\hat{p}_j)} + (t - 1) - \prod_{i=1}^{t}(1 - \hat{p}_j) - 1\right]$$

where $\hat{p}_j = \frac{c_j}{\hat{N}_p}$ is the detection ("capture") probability.

If $p_j = 1$, all objects of interest are detected (i.e., correctly classified) by the j^{th} classifier. The assumption that classifiers are independent is not likely to be true. If it is false, then $E(\hat{N}_p) < N_p$, where E is the expectation operator. This estimator can be expected to be useful even if there is some dependence if the detection probabilities p_j are near 1 (say > 0.80 or 0.85). When dependence or heterogeneity of detectability is encountered, this procedure can give an estimated total less than the number observed ($\hat{N}_p < M_{t+1}$), an undesirable situation. This procedure critically assumes no false counts are made. This can be met by proper selection of each threshold, T_i (Figures 1 to 3) and section on Avoiding Pixel Mixture Errors).

A second model is more generally useful and allows heterogeneity in detectability. A series of estimators have been developed (Burnham and Overton, 1978, 1979) for this model based on the generalized jackknife. This is a nonparametric estimation theory and has been shown to be robust to violations to the underlying model (see Otis *et al.* (1978) 123–133).

Regardless of the pdf of heterogeneity, the minimal sufficient statistics are the classification frequencies f_j (Burnham and Overton, 1978, 1979), and the jackknife estimator is a linear combination of these statistics. If $t = 2$ (two classifiers),

$$\hat{N}_p = \left[1 + \left(\frac{t - 1}{t}\right)\right] f_1 + f_2$$

$$= (1 + 0.5)\, f_1 + f_2 \qquad \text{or } 1.5f + f_2$$

$$\mathrm{var}\,(\hat{N}_p) = \left[\left(1 + \frac{t - 1}{t}\right)^2 f_1 + f_2\right] - \hat{N}_p$$

$$= [2.25f_1 + f_2] - \hat{N}_p.$$

The procedure is often less than satisfactory in this case because little information is contained in just f_1 and f_2. In the case $t = 3$,

$$\hat{N}_p = \left(1 + \frac{3t - 6}{t}\right) f_1 + \left(1 - \frac{3t^2 - 15t + 19}{t(t - 1)}\right) f_2 + \left(1 + \frac{(t - 3)^2}{t(t - 1)(t - 2)}\right) f_3 = 2f_1 + \frac{5}{6} f_2 + f_3$$

$$\mathrm{var}\,(\hat{N}_p) = \left[\left(1 + \frac{3t - 6}{t}\right)^2 f_1 + \left(1 - \frac{3t^2 - 15t + 19}{t(t - 1)}\right)^2 f_2\right]\left(1 + \frac{(t - 3)^2}{t(t - 1)(t-2)}\right) f_3\right] - \hat{N}_p$$

$$= \left[4f_1 + \left(\frac{25}{36}\right) f_2 + f_3\right] - \hat{N}_p.$$

Higher order estimators exist for $t > 3$ and, while they become increasingly complex, they still exist in closed form (see Burnham and Overton, 1978).

The class of estimators is quite generally applicable. For example, consider a multispectral scanner employing four spectral bands. Six ratios could be defined resulting in 64 classification patterns n_w (i.e., $n_1, n_2, \ldots, n_6, n_{12}, \ldots, n_{56}, n_{123}, \ldots, n_{456}, \ldots, n_{123456}$). This represents a large amount of information, providing the firm basis for the estimation of N_p.

TABLE 2. RESULTS OF MONTE CARLO STUDIES OF THE PERFORMANCE OF THE DARROCH AND JACKKNIFE ESTIMATORS FOR TWO AND THREE CLASSIFIERS. ALL RESULTS ARE BASED ON 400 REPLICATIONS.

Estimator	Number of classifiers	True N	Scheme[1]	Average $\hat{N}$	Percent relative bias	Average C.I. width	C.I. coverage (95%)
Darroch	2	100	A	87	−13	35	56.8
		100	B	89	−11	191	72.5
		400	A	337	−16	163	58.2
		400	B	334	−16	323	70.0
		1000	A	839	−16	253	31.2
		1000	B	821	−18	483	58.5
	3	100	A	87	−13	91	74.0
		100	B	88	−12	98	70.8
		400	A	343	−14	68	15.5
		400	B	335	−16	—	—
		1000	A	862	−14	109	1.0
		1000	B	830	−17	265	30.8
Jackknife	2	100	A	61	−39	19	0
		100	B	40	−60	16	0
		400	A	248	−38	40	0
		400	B	164	−59	34	0
		1000	A	622	−38	64	0
		1000	B	414	−59	54	0
	3	100	A	99	−1	33	92.5
		100	B	69	−31	30	6.5
		400	A	397	0	68	95.5
		400	B	277	−31	61	0
		1000		998	0	108	96.0
		1000	B	694	−31	97	0

[1]The probability of correctly classifying a deer for each of the two (or three) classifiers was [0.4, 0.3, (0.5)] for scheme A, while for B it was [0.18, 0.25, (0.21)].

The estimators of n_p were evaluated using Monte Carlo methods whereby 400 replications of the $\mathbf{X}_{ij}$ matrix were generated for each of 24 cases. For each case, N_p was known (e.g., 100, 400, or 1000) and for each Monte Carlo trial, $\hat{N}_p$ was computed. The evaluation was based on

$$\text{percent relative bias} = \left(\frac{\text{ave}(\hat{N}_p) - N_p)}{N_p}\right) \times 100$$

$$\text{where} \qquad \text{ave}(\hat{N}_p) = \frac{1}{400}\sum_{i=1}^{400} \hat{N}_\pi$$

and the percentage of the 400 replications where the 95 percent confidence interval covered the true parameter, N_p.

ANALYSIS RESULTS

MONTE CARLO RESULTS

A Monte Carlo study was performed to indicate the utility of capture-recapture theory in a remote sensing context. The purpose of this study was to examine the average performance of the estimators (e.g., bias and achieved confidence interval coverage). The method outlined by Otis *et al.* (1978:123) and White *et al.* (1982:218–221) was used to generate 2,400 data sets according to two different parameter sets, A and B. Scheme A was similar to the situations depicted in Figures 1 to 3 in that the detection probabilities p_j of deer ranged from 0.3 to 0.5. Scheme B was more pessimistic in that the detection probabilities ranged from 0.18 to 0.25. Both schemes reflect the fact that certain ratios of spectral reflectance are better than others at discriminating deer versus nondeer and that the ratios are partially dependent due to heterogeneity of the target and its background.

Moderate heterogeneity was incorporated in the underlying model to generate Monte Carlo data that would reflect the fact that some deer are more difficult to classify correctly than others because of the varying quality of the spectral information. This also reflects a moderate degree of dependence among the classifiers and the p_j. Three population sizes were used: 100, 400, and 1,000. Systems with two and three classifiers ($t = 2$ and 3) were simulated and the results appear in Table 2.

Chapman's (1951) and Darroch's (1958) estimators, the traditional capture-recapture approach, have

TABLE 3. RESULTS OF A FIELD TEST TO COUNT DEER USING TWO RATIOS OF THREE SPECTRAL BANDS.

	No. of deer detected	Expanded No. of False Counts		Deer detection probability	False count probability[3]
		All counts[1]	Only multiple counts[2]		
Classifier 1 [0.764]/[0.981]	41	20	0	0.82	0.000046
Classifier 2 [0.981]/[0.725]	43	40	5	0.86	0.000093
Classifiers 1 and 2	34	5	0	0.68	0.000012

[1]Without spatial constraints. The unexpended counts can be found by division by 5; i.e., 4, 8, and 1, respectively.
[2]With spatial constraints (i.e., classify as "deer" only if two or more adjacent pixels are "deer").
[3]Unexpended number of false counts/(1728 elements by 50 scans).

moderate negative bias (-11 to -18 percent) and poor confidence interval coverage (1 to 74 percent when the nominal value is 95 percent) for both schemes A and B for all population sizes. This is to be expected because these methods assume independence. The jackknife estimator (Burnham and Overton, 1978) does very poorly using data from only two classifiers. For three classifiers ($t = 3$), the performance of the estimator for data generated under scheme A is excellent (relative bias < 1 percent and 92.5 to 96 percent confidence interval coverage), but poor for scheme B (-31 percent relative bias and 0 to 6.5 percent confidence interval coverage). This is to be expected from what is known about the magnitude of the capture probabilities in capture-recapture theory. If these detection probabilities are low, the jackknife estimator performs poorly. While this simulation study is not exhaustive, three inferences can be drawn from it, and these can also be supported from what is known about this method in traditional capture-recapture studies. First, the jackknife estimator is preferable in some cases, especially when heterogeneity and dependence are significant and capture probabilities are about 0.3 or 0.4. Second, the jackknife can be expected to perform better if three or more classifiers are used. For example, recall in a four-color system, 64 n_w patterns exist and provide significant additional information about the population size. Third, improved performance is expected of the jackknife estimator as sample size (f_j) increases.

INCORPORATING SPATIAL INFORMATION

Our field experiments have shown that most false counts in the deer detection system occur as single isolated pixels in a 1 by 1728 scan. In fact, the ratio of single false counts to two or more neighboring false counts ranged from about 5 to 9. Because the smallest dimension of a deer is larger than the width of a 10-cm square pixel, a simple spatial constraint was incorporated to reduce further the frequency of false counts. This was done by incrementing the counters c_j and n only if two or more adjacent pixels are classified as deer. Single pixels classified as deer are ignored. Such a spatial constraint is easy to implement electronically, and it reduced the number of false counts in our small-scale experiments by a factor of 5 to 9. It should be noted that, except for the information in Table 3, a false count will be taken to mean two or more adjacent pixels incorrectly classified as "deer."

DETAILED EXAMPLE

A brief example will further illustrate the general approach. On 14 October 1982 a test was conducted in a snowless area near the summit of Sardine Canyon, about 20 miles southwest of Logan, Utah. The prototype instrument initially constructed by Voorheis (1982) using only two ratios ($t=2$) of three spectral bands, [0.764]/[0.981] and [0.981]/[0.725], was used. The classifiers were set using thresholds of 1.50 and 0.688, respectively. (Note; the first threshold is more conservative than that shown in Figure 2, in a further effort to reduce the possibility of false counts.)

A deer hide was placed in a vertical position at approximately 445 m, and 50 scans were made to simulate a population of 50 deer ($N = 50$). The classification of each pixel by each of the two classifiers was output by a small thermal printer. This allowed careful interpretation of single, and multiple adjacent, false counts and deer that were missed. The number of deer (N) could be estimated, rather than only N_p, by visually examining the pixel patterns from the printed output.

Forty-one of the 50 deer were classified correctly by the first classifier ($c_1 = 41$) and 43 of the 50 deer were correctly classified by the second classifier ($c_2 = 43$) while only 34 deer were detected by both classifiers ($n_{12} = 34$). Of course, both c_1 and c_2 are substantially less than N, indicating the need for capture-recapture theory to reconstruct an estimate of N from the incomplete counts.

The thresholds T_1 and T_2 were established empirically in an attempt to avoid false counts, while risking an incomplete count of deer. In spite of the rather extreme thresholds, one false count was made by the second classifier (so c_2 was actually 44). The single observed false count was expanded to five based on our knowledge of *a priori* probabilities of nondeer. With this realism accounted for, the relevant statistics are

TABLE 4. SUMMARY OF RESULTS OF FIELD TEST WHERE THE NUMBER OF DEER WAS 50 AND THREE CLASSIFICATION ALGORITHMS WERE USED TO ANALYZE THE DATA AND MAKE ESTIMATES OF THE NUMBER OF DEER.

Classifier	No. of deer actually detected	No. of false counts (expanded)[1]	Estimated no. of deer	Percent error
Multistage[2]	34	5	39	−22
Multistage and spatial algorithm[3]	34	0	34	−32
Multistage with capture-recapture estimation	41,43	20,40[4]	103	106
Multistage with capture-recapture estimation and spatial algorithm[3]	41,43	0,5[4]	58	16

[1]Expanded by 5 based on an assumption of the *a priori* probability of a deer of 10^{-5}.
[2]"Deer" detected only if both classifiers are "deer."
[3]Classification as "deer" only if at least two contiguous pixels are "deer" (i.e., single pixels as "deer" are ignored).
[4]For the first and second classifier (reflectance ratio), respectively.

$$c_1 = 41,\ c_2 = 48 \text{ (i.e., } 43 + 5\text{), and } n_{12} = 34.$$

Chapman's (1951) estimator is known to be better than the jackknife for only two classifiers (Table 2). Thus,

$$\hat{N} = \frac{(41+1)(48+1)}{(34+1)} - 1$$

$$= 58$$

$$\text{se}(\hat{N}) = \sqrt{\frac{(41+1)(48+1)(41-34)(48-34)}{(34+1)^2(34+2)}}$$

$$= 2.1$$

$$95 \text{ percent confidence interval} = 54 - 62.$$

This represents a 16 percent relative bias, which seems quite encouraging. If the false count could have been eliminated, $\hat{N} = 52$, with only a 4 percent relative bias. This again shows the critical importance of eliminating false counts.

In this example, the detection probabilities were 41/50 = 0.82 and 43/50 = 0.86 for the first and second classifiers, respectively. The magnitude of these probabilities is quite sufficient for capture-recapture estimates to perform well, in the absence of false counts. This suggests setting the thresholds at even more extreme levels, reducing the detection probabilities, but further eliminating false counts. Estimation could be markedly improved if data on a third or fourth classifier were available (our eventual intention in the deer censusing program).

The counts c_1, c_2, and n_{12} were incremented in this example only if two or more contiguous pixels were classed as deer. The use of this simple spatial constraint eliminated four single false counts by the second classifier.

Further results are summarized in Table 4. A simple multistage classification scheme is not satisfactory (22 to 32 percent underestimate). If thresholds T_1 and T_2 are set conservatively to minimize false counts, then more serious underestimates can be expected (Table 4). Alternatively, if T_1 and T_2 are set less conservatively, then even more serious error (overestimation) can be expected.

Treating the classification problem in the context of capture-recapture theory is promising when combined with the simple spatial constraint. In any case, false counts must be avoided. In general, N_p must be estimated first. Then $\hat{N} = \hat{\beta}\,\hat{N}_p$ where $\hat{\beta}$ is a constant, obtained empirically, roughly equivalent to the inverse of the average number of pixels per deer classified.

CONCLUSIONS AND RECOMMENDATIONS

Remote sensing systems for enumerating rare objects in heterogeneous spectral backgrounds must avoid false counts while providing fast and accurate counts of the objects of interest. The spectral information in two to four bands can be expected to be less than adequate, causing classification errors; at least several background objects will be spectrally similar to the object of interest, leading to false counts.

Multistage classification algorithms based on capture-recapture theory have merit, especially when combined with spatial information to further reduce false counts. These algorithms are fast and can be implemented in real time with existing hardware. Ideally, three or four classifiers can be employed to provide the **X** matrix data. Proper spectral band selection is critical to avoid the problem of pixel mixtures in some applications.

The classification algorithm recommended provides an estimate of the number of pixels N_p containing the object of interest, not the number of objects of interest N. Estimation of the constant that links these $N = \beta N_p$ has not been addressed here.

ACKNOWLEDGMENTS

This research was supported jointly by the Arizona Game and Fish Department, the Oregon Department of Fish and Wildlife, and the Utah Division of Wildlife Resources with matching funds from the Federal Aid in Wildlife Restoration Act. K. P. Burnham reviewed a draft manuscript and offered several important suggestions.

REFERENCES

Burnham, K. P., and W. Overton, 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65(3):625–63.

———, 1979. Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60(5):927–936.

Chapman, D. R., 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. *Univ. Calif. Publ. Stat.* 1:131–160.

Conners, R. W., M. M. Trivedi, and C. A. Harlow, 1984. Segmentation of a high resolution urban scene using texture operators. *Comput. Vision Graph. & Image Process.* 25:273–310.

Darroch, J. N., 1958. The multiple-recapture census: I. Estimation of a closed population. *Biometrika* 45(3/4):343–359.

Magnusson, W. E., G. J. Caughley, and G. C. Grigg, 1978. A double-survey estimate of population size from incomplete counts. *J. Wildl. Manage.* 42(1):174–176.

Maxim, L. D., L. Harrington, and M. Kennedy, 1981. A capture-recapture approach for estimation of detection probabilities in aerial surveys. *Photogram. Engin. & Remot. Sens.* 47(6):779–788.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson, 1978. Statistical inference from capture data on closed animal populations. *Wildl. Monogr.* 62:1–135.

Petersen, C. G. J., 1896. The yearly immigration of young plaice into the Limfjord from the German Sea. *Rep. Danish Biol. Sta.* 6:1–48.

Seber, G. A. F., 1982. *The estimation of animal abundance and related parameters.* 2nd Ed. Macmillan, N.Y., 654 p.

Trivedi, M. M., 1979. *Feature selection and classifier design with applications to remote sensing of deer.* Ph.D. Dissertation, Utah State University, Logan. 121 p.

Trivedi, M. M., C. L. Wyatt, and D. R. Anderson, 1982. A multispectral approach to remote detection of deer. *Photogram. Engin & Remot. Sens.* 48(12):1879–1889.

Trivedi, M. M., C. L. Wyatt, D. R. Anderson, and H. T. Voorheis, 1984. Designing a deer detection system using a multistage classification approach. *Phtogram. Engin. & Remot. Sens.* 50(4):481–491.

Voorheis, H., 1982. *Development of an electro-optical classifier system for remote sensing of deer.* Master of Science Thesis, Utah State University, Logan. 63 p.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis, 1982. *Capture-recapture and removal methods for sampling closed populations.* Los Alamos National Laboratory, Report LA-8787-NERP. 235 p.

(Received 22 June 1984; revised and accepted 11 February 1986)