BULGARIAN ACADEMY OF SCIENCES

SPACE RESEARCH IN BULGARIA, 5

Sofia . 1986

Mathematical/Statistical Methods for Classification of Objects by Means of Spectral Reflective Characteristics

T. K. Yanev, D. N. Mishev

The resolution of the problems of the earth surface and atmosphere remote sensing relates to the measurement of various quantitative parameters of the electromagnetic field. These are mainly parameters of the solar radiation, reflected from the earth surface and transformed by the atmosphere, as well as of the proper thermal radiation of the earth-atmosphere system.

We have to mention first the spectral reflective signatures (SRS) and in particular the spectral reflective coefficient (SRC).

The SRS is a photometric function representative of the solar radiation structure, dissipated from the various elements of the surface of a given natural formation. The SRS is a multiple function of variables characterizing the physico-chemical and biological properties of the studied natural formations and the conditions required to obtain the SRS (illumination, atmospheric conditions, etc.) That is why the SRS is a multifactorial function containing specific information on the objects it has been obtained from.

I. SRS features

In order to complete a SRS classifier (see point II.), it is reasonable to group the factors that define the SRS values according to some of their general features. Further 'feature' will mean a factor that affects the SRS in a clearly expressed way. The features may be grouped as follows:

A. Primary features — defining the function of reflected solar radiation, trapped at the input of the measurement system;

a) intrinsic features of the studied subject:

— subject-specifying which characterize the elementary unit representative of the subject properties;

— geometric (shape, surface, structure, etc.).

b) external features (atmospheric conditions, illumination, air and soil temperature, soil and air moisture, soil electroconductivity, orientation of the measurement system, etc.).

B. Secondary features which are the SRS values, measured in the individual channels of the multizonal measuring systems.

C. Generalized features obtained in the processing of the primary and secondary features after being treated with certain operator (mathematical, taxonomic, etc.). The generalized features may be grouped into:

a) generalized primary features, for example, type-kind and subkind, forming the taxonomic features within the natural classification of the subjects: soil, vegetation, water, etc.

b) generalized secondary features obtained after mathematical transformations of the SRS.

II. Aim of the SRS analysis

The necessity to study and classify the SRS results from the possibility of their application in agriculture, geology, ecology, etc. The formal expression of the SRS applications in this respect is: to define the relationship between the SRS and the other features (primary and secondary-generalized), in order to resolve the reverse problem; based on the SRS information to reconstruct partially or entirely the values of the primary features (and their generalizations). Such a reversible relation is the classifier. It could be compiled on taxonomic, probabilistic, set-structured, regressive, physical modelling, etc. principle.

IIII. Specific features of SRS

The SRS are random functions obtained in discrete shape (spectrum along the wavelength) under conditions of incomplete apriori information and significant accompanying noise. These specifics require the implementation of statistic-probabilistic techniques for the SRS analysis.

The larger portion of subject-specifying features (including quantitative generalizations) have joint conditional distributions in the multidimensional space of the signatures that, as a matter of fact, are not intercrossing.

Further the conditional distributions will be briefly referred to as distributions only. In this sense such features may be denoted as discrete (Fig. 1), and at the highest level of generalization as qualitative. Some of the subject-spesifying features (and their generalizations) characterize the state of the given subject only (for example, various evolution stages) and may have intercrossing distributions, but of differentiated modes, i. e. they are quazi-discrete. The external features (parameters) may have largely overlapping distributions resulting in smooth transition of the parametric surfaces into the multidimensional space. Such features may be denoted as continuous (quantitative). In the course of development or formation of the subject, for example soils, the transition into a stage (or type or kind) possibly can be performed smoothly, but for the purpose of our Classifier well-shaped states are of interest and this defines the necessity of introducing the quasi-discrete features. Regardless of the eventual smooth transition between them, the probability of distinguishing them is defined with the possibility of identifying various states of the studied subjects in their set of features.

Another important specificity of the features is that they are random magnitudes in terms of our lack of knowledge on their expected values in the progress of the experiment.

a) The primary features (and their generalizations) are random due to inaccuracy of the measuring instruments and insufficient volume of the representative samples in the course of the experiment.

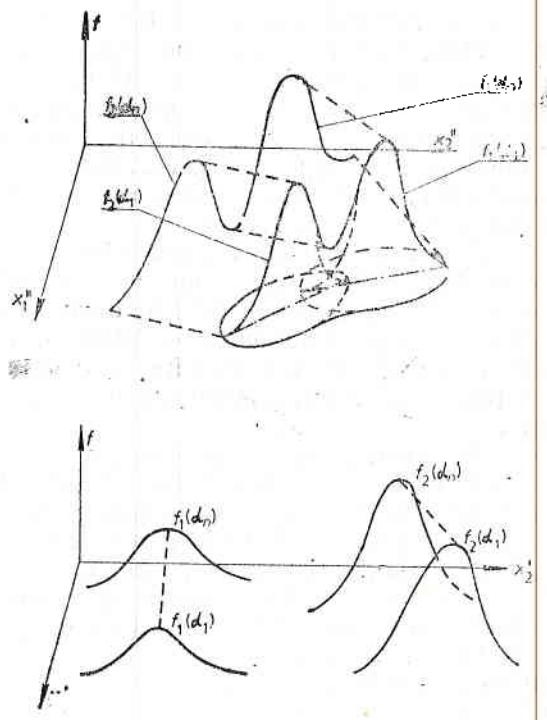

Fig. 1. Discrete and quasi-discrete feature distributions

x_1', x_2' — discrete features;
x_1'', x_2'' — quasi-discrete features;
f — distribution density;
α — parameter

b) The SRS are random functions due both to the reasons innumerated here and in a), as well as to the impossibility to study all primary features affecting the SRS. This problem will be discussed in detail in point IV.

IV. Major problems

As it was already mentioned, the main purpose is to define the relationship between the SRS and the other features. The precise analytical shape of this relationship still seems an irresolvable problem. The various stages of approximation to this task could be formulated as follows:

Physical model of SRS.

At the available state of our knowledge this is possible only for some particular relations between the SRS and the primary features. The modest results attained by now do not permit large-scale applications in agriculture.

Statistical Model of SRS.

This is realized with the determination of the regressive dependence (surface in the multidimensional space of the feature) between SRS and a given set of primary features or their generalizations. There is no available method by now to define the confidence regions of such a regression in the general case, when both the features and the dependence value (the SRS in this case) are random variables [1]. The Classifier procedure requires to know these confidence regions. Relatively easier is the problem of regression coefficients determination. This can be done using, for example, the method of the maximum likelihood [1], the method of least squares [2], etc., and the regression is traced between the i-th component of the SRS $\mathbf{r}$ vector and all the k-th components of the $\boldsymbol{\alpha}$ vector of the features, $k=1, \ldots, K$

$r_i=f(\alpha_1, \ldots, \alpha_k)$, $i=1, \ldots, m$-channel number for SRS.

The confidence intervals of the regression may be defined with certain inaccuracy with the available techniques (which require for the independent variables of the regression to be determined variables) if we can ensure that the relative variance (measured for example, with the variation coefficients) of the primary features (and their generalizations) is at least of an order smaller than that of the SRS.

The evaluation in this respect may be obtained, if we compare the vector of the relative errors δ_α of the vector $\boldsymbol{\alpha}$ of the primary features (including their generalizations) and the vector of the relative errors δ_r of the vector $\mathbf{r}(\boldsymbol{\lambda}, \boldsymbol{\alpha})$ of the secondary features (incl. their generalizations). If δ is assumed to be measured with the variation coefficient, then $\delta_l=S_l/\mu_l$, where μ_l is the average value of the l-th component of $\mathbf{r}$, respectively $\boldsymbol{\alpha}$, and S_l is an estimate of its mean square deviations. Basic sources for δ_α and δ_r formation are respectively (p. A for δ_α and p. B for δ_r):

A. a) Test field measurements:

i. Errors from the measuring systems.

ii. Errors from the interpolation and extrapolation of the measured data, when studied areas are large and with various conditions (and they must be such in order to receive statistical representation of the data), therefore measurements of sufficient coverage cannot be performed.

b) Remote sensing (helicopters, airplanes, balloons, satellites):

i. Errors from the measurement systems.

ii. Errors from interpolation and extrapolation of the measured values (as for p. A. b).

B. a) Test field measurements:

i. Errors from measurement systems.

ii. Errors from insufficient knowledge on SRS yielding conditions (mainly in evaluating the illumination conditions).

iii. Undetermination of $\mathbf{r}$ in the result of the fact that not all the intrinsic and external features are included in $\boldsymbol{\alpha}$.

b) Remote sensing:

i. Errors from inaccurate evaluation of the SRS yielding conditions (illumination, atmospheric state, etc.).

ii. As in B, a, iii.

All the features incorporated in $\boldsymbol{\alpha}$ (if properly selected) are independent from $\mathbf{r}$ (but they also might be interdependent) and affect the $\mathbf{r}$ values on the cause-effect scheme.

Present experience shows that the $\boldsymbol{\alpha}$ and $\mathbf{r}$ component distributions are close to normal. From the definitions in A and B and the experience acquired, the following conclusions can be made:

C. δ_{Ba1} and δ_{Ba2} are commensurable respectively with δ_{Aa1} and δ_{Bb1}, and δ_{Bb2} is commensurable respectively with δ_{Ab1}.

D. δ_{Ba3} and δ_{Bb3} are commensurable respectively with δ_{Aa2} and δ_{Ab2}.

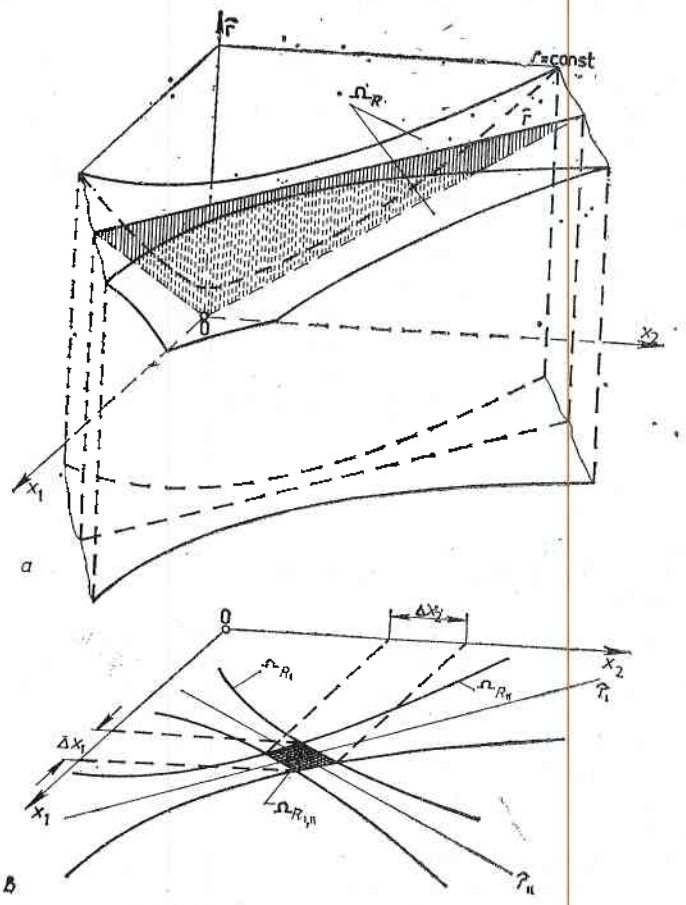

Fig. 2. Confidence regions of regression surfaces

a — three-dimensional regression surface $\hat{r}$ and its confidence region Ω_R;

b — two-dimensional projects of confidence regions Ω_{R_I} and $\Omega_{R_{II}}$ and their cross section $\Omega_{R_{I,II}}$

The errors from D are much larger than from C. The available experimental possibilities provide for a significant reduction of δ_{Aa2} only as compared to the other δ from D. Therefore, a regression with confidence intervals determination can be effectively traced (at least at present) only between r and α on test field data, if a sufficient number of well planned experiments is available with δ_{A2} is at least an order smaller than δ_{Ba3}.

The possibility to obtain regression with a confidence region still does not lead to an easy practical application, at least due to:

1. The confidence regions Ω_R (Fig. 2) are defined for the regression surface parameters and the classifier requires the confidence region of the individual result $\Omega_I \approx \Omega_R \cap \Omega_{IS}$ to be known, where Ω_{IS} is the sample value of Ω_I.

2. The determination of Ω_R and Ω_{IS} is related to significant computation difficulties. Fig. 2 shows the case of two dimensional regressions (α contains two features). Even in this relatively simplified case the analytic (and respectively the automatic) determination of Ω_R is significantly hampered. We should add to the necessity of plotting the Ω_R projections over the coordinate system axes the necessity of knowing the confidence intervals of the individual parameters (features). Additional difficulties result from the determination of Ω_{IS}. Certain facilitation may be achieved if r is obtained as a mean square from the multiple resolution elements (RE), each of a dimension representative of the least possible RE, already marked with the properties of the studied class. Then Ω_{IS} will tend to Ω_R. This case reveals the problem of determining the minimum RE dimensions. Anyway, such a problem should be resolved in general prior to the large-scale SRS measurements, in order to determine for each class or theme of classes (although approximately) the optimal RE dimension in agreement with beforehand selected criteria. In [3] there is a solution through which the optimal RE dimension is defined in the course of the experiment. Certain general methods and instrumental requirements for site-testing in the ground-based studies of SRS are given in [4,5].

Another problem related to RE of SRS is the case when mixture of two or more thematic classes is available in RE (i. e. such that will be of interest to the interpreter) and a separation of this mixture is required. In general, the separation of SRS mixture is a problem of two main aspects:

— separation of a mixture from SRS distributions obtained from a RE set (pixel), within the limits of which RE there is no mixture of SRS of different classes, and

— separation of SRS mixture, obtained within the limits of one RE.

The first problem is closely related to the selection of optimal RS dimension. Under proper selection, if possible, i. e. if a plateau exists within the limits of the studied region, the problem does not exist at the interpretation level. In the reversible case the SRS distribution from various RE is obtained in polymode. Under defined conditions, for example, identification of the distribution type, the component number determination is performed with the well known methods [6].

The second problem requires solution when two or more interpetation classes are mixed within the RE limits [7,10] and their percentage participation is sought (weight coefficient) in their collective SRS. Such case is represented, for example, with series of agricultural crops divided by soil strips. The problem is solvable in the statistical sense, if the class distribution number is known for the classes participating in the mixture and their parameters. Then the confidence intervals can be defined for the weight coefficients when sufficient channel number in the SRS measuring system is available.

In conclusion we may say that the elaboration of general regressive function $r(\lambda, \alpha)$ is still very difficult both on methodological and applicable computing level. Therefore, in near future it is possible to obtain only particular regressive dependences of largely limited α dimensionality. Nevertheless, we consider that even so the information will be useful for the SRS interpretation.

Probabilistic-set classifier

This classifier is built by comparing the confidence regions of the primary features and their generalizations with the confidence regions of the SRS or their transformations. The adequacy of this comparison with the available apriori information, as well as the updating with the experience acquired is controlled by

the interpreter. This approach is suitable for the discrete and quasi-discrete features mainly. The major stages in the compilation and operation of such a classifier are:

a) training — with and without supervisor,
b) classification and new information accumulation,
c) updating,
d) dimensionality reduction of the features space in a given thematic class set.

Further, a largely simplified scheme of the SRS classifier is applied. This is compiled within the attempt to balance the two basic contradictory requirements: application simplicity and sufficient accuracy of the classificaton.

The following exposure contains in brief and in a time-generalized form the fundamental information for the structural units of the suggested scheme and also a brief comparative analysis of some available methods and approaches related to the discussed problem.

V. SRS Classifier

1. SRS Transformations

The basic criterion of effectiveness for a given transformation is the risk function R (Appendix 1). In a fixed set of classes the transformation which reduces R to a higher degree should be adapted as more efficient. The evaluation of R in most of the cases is difficult. Some indirect criteria are known that are used as alternatives in first approximation of R linear and nonlinear functions of the intra- and interset class distances (clusters respectively), enthropic criteria, etc. In the case when the class distributions are close to normal we assume as the clearest and in many cases directly proportional the relationship between the variation coefficient V of $\vec{r}$ (by components) and $R: V_i = \sigma_{r_i}/r_i$, $i=1, \ldots, m$ — channel number of SRS. Further on, when possible, it will be preferred as a measure of effectiveness for the transformation replacing R. The main types of transformations that are available are:

1.1. Orthogonal (expanded by orthogonal functions).

a) non-eigen systems of orthogonal functions: Fourier, Walsh, Hadamard, Vilenkin-Chrestenson, orthogonal polynomials, etc. No direct data are available as to whether the use of these expansions coefficients as $\vec{r}$ components in the transformed space results in V reduction. It may be demonstrated in particular that for the Fourier transformations such a reduction is not guaranteed (see Appendix 2).

Another defect of these transformations is that when steep SRS sectors are available (for example, vegetation), a large number of expansion components is necessary which increases the dimension of the transformed space. Also these transformations assume that SRS are periodic functions, i. e. a compromise with reality which results in description inaccuracy.

b) eigen systems — transformations of Karhunen-Loeve (in the discrete version a method of the main components). The advantage of these transformations is that no SRS periodicity is required here. The disadvantage being that the eigen systems of the j-th class is optimal with respect to the economic class description only for this class, and not for the other classes from the classifier set examined. That is why, the switch-off of some SRS channels after the analysis of the given class in its eigen system may negatively affect

R within a given situation of the class description in the set (for example, it is possible to identify well some of the classes precisely in these channels).

1.2. Autocorrelative transformations (some formulae are given in Appendix 3).

The advantages of these transformations are: dimension of the transformed space not larger than that of the primary space; the variation coefficient of r in the transformed space is smaller than that of the primary space; simple calculation operations for transformation. As a disadvantage we may point out the ambiguity of the transformations II and IV in the sense that equal autocorrelative functions may be derived from various functions. But this ambiguity includes a class of symmetrical functions to the coordinate axes, and such symmetries cannot be found in SRS of the natural formations over the Earth surface.

1.3. Enthropic transformations (Appendix 4).

As it is already known, they are realized in eigen coordinate systems [9, 13], and, therefore, the summary of p. 1.1. refers to them also.

1.4. Divergency transformations (Appendix 4).

They lead to the optimal results in the sense of minimizing the defined divergency only for a given couple of classes and, therefore, the disadvantage of the eigen transformations from p. 1.1. In addition, they involve a very large number of calculations, and information losses in the real case in unequal covariant matrix of the classes [9].

1.5. Metric measures for similarity and distance optimization bewteen vectors-realizations.

According to the type of the distance metrics, these transformations divide into: (A) Euclidian; (B) Mahalanobissian, etc.; and by the optimization criterion into: (a) minimizing intraset distances; (b) maximizing interset distances, and (c) mixed.

i. Complex (in the sense of a certain criterial function) distance optimization [9, 11, 12, 13] and ii. Serial realization of a) and b). Usually (A) methods are based on linear transformations and relate directy to (a) and (b). (A, a) transformations lead also to enthropic transformations. The non-parametric Patrick-Fisher transformations [13] use the exponential function as a criterion that increases with the expansion of the interset distance and the reverse, and relates to (c).

The advantages of these transformations are mainly two: they have a completed analytical technique for determination of the transformation matrix, and provide possibility to improve the class separation in the sense of the defined criteria, but without a direct link with R.

The disadvantages are grouped into: i. criteria (a) do not guarantee separately the obligatory decrease of R; ii. criteria (c) make it possible to evaluate in explicit form the increase of the class separability only when there are two classes, while in the general case of many classes (especially when the primary space is populated with a large density, it is difficult to evaluate beforehand the effect of the global distance change between the realizations. Description of other metric and non-metric clusters-algorithms is given in [9, 11]. The possibility for their application in SRS classification is probably smaller than of the techniques discussed here.

This brief comparative analysis provides certain advantages of the autocorrelative transformations, because it is possible to obtain with them universal (for the complete class set studied) improvement of the class separability, measured in the first approximation with the variation coefficient V and estimated by now for the primary space only, and not for the other types of transformations (These transformations also relate to simple computing operations)

As a general disadvantage of all the mentioned transformations, we shall note the absence of a clear analytical link in the general case of unequal covariance matrices of classes between the risk function in the primary and transformed space. Due to this the transformation effectiveness can be evaluated only approximately in an analytic mode. After the transformation, direct calculations of R must be performed for each beforehand given class set with a theme compiled in the Classifier.

2. Transformations to avoid ill-conditioned covariation matrix

When the channel number of the system for the SRS yield is large, for example 32, there is a possibility for the determinant of the covariation matrix of certain classes either to become smaller than the computer zero (in the case when the SRS are reduced to spectral illumination coefficients smaller than a unit, and therefore, with dispersions of the 10^{-1} order), or to exceed the upper computer limit (for example, when the quantization level number of r_i is of the order of 100-200). A possibility to avoid this ill conditioning due to scale effect is provided, if the primary SRS data are this multiplied with a suitable and equal for all SRS number or through the division of each SRS to its mean arithmetic: this means a reduction to a relative coordinate system and loss of information for the mean class vector, except if it is not introduced as an additional feature. The transformations discussed are not effective when the determinant is ill-conditioned in structure.

3. Subtraction of submatrix

A possibility of reducing the input matrix from dimensions $(m \times n)$ to dimension $(m' \times n)$ is foreseen when it is possible to decrease dimensionality of the primary space of the features.

4. Cluster analysis

It is applied for grouping the input data by some formal criterion of similarity. The quantitative expression of this criterion is the measure of similarity. Usually it is selected as 'distance' in space apriori selected metrics (for example, Euclidian). The following procedures divide the cluster-algorithms into two groups: (a) subjecting the measure of similarity to threshold values through the realization of non-equalities, and (b) optimizing the selected function of this measure, in order to define a criterion (most often for this purpose the transformations from 1.5. are applied). There are tens of well-known cluster-algorithms (nearest neighbouring maximum distance, ISODATA, etc.) [9, 12, 13]. The computations of inter- and intraset distances is performed mainly for algorithms from (b). These operations make difficult the algorithmic application when the number of vector-observations that are subject to clustering is larger than 10^3 [12]. Since in the real case of remote sensing data (and even test field data) for the SRS their number will be larger than 10^3, it is recommendable to fix on the application of more simplified cluster-algorithms) mainly those from (a).

5. Automatic controls

5.1. For cluster population: if population is less than the critical (given by the user) the cluster is not analysed and is entered into the memory.

5.2. For normal cluster distribution: in satisfying the requirements of a given criterion of normality the vector-observations forming the cluster are passed for the calculation of an average vector and covariation matrix.

5.3. For ill-conditioned covariation matrix.

6. Normal distribution simulation

Based on average vectors given by the user, a set of normally distributed vectors is obtained, i. e. classes of normally distributed vectors are modelled. Such data file is necessary for the comparative study of the effectiveness of various transformations or other programs when the real data available are insufficient.

7. Compilation of classifier

7.1. All clusters satisfying the controls (incl. the operator control of clustering accuracy – see explication in 7.5) and all classes formed by the user at the input of the main program are passed for computation of average vector and covariation matrix. Afterwards the information is storaged in the Classifier. The storage may have various variants in dependence on what type of transformation has been applied for its compilation.

7.2. Storage classes limits: (a) with Bayes approach—the limits depend on the set of classes which at a given stage of classifier compilation are compared in determining the risk function. This approach ensures minimum risk function; (b) in beforehand limitation of classes.

i. Multidimensional confidential parallelepipeds with axes parallel to the coordinate axes and multidimensional confidence ellipsoides whose axes have been obtained with sufficient statistic material for the classes (after a training procedure); in this approach the computing operations are largely facilitated, but the value obtained of the risk function is not minimum. Anyway, solution may be looked for at apriori given admissible upper limit of R.

ii. With linear and non-linear decision functions; in this case — the region of a given class is localized with limitations of some of the multidimensional planes and non-linear surfaces. This approach is related to the following significant disadvantages:

— even when the discriminators are linear at a class number larger than 20-30, the oblique multidimensional linear surfaces are complex and result in an inadmissibly large number of computations; in the case of SRS classification, the class and subclass number subject to identification should hardly be smaller than several tens;

— probabilistic evaluation is difficult for the classification quality, moreover that the larger part of these algorithms do not permit probabilistic evaluation and become cycling when the classes intercross (an exception is the algorithm of Ho-Kashyap [9]); iii. In the statistic algorithms for obtaining a decision functions: stochastic approximation, perception approach, potential functions, etc. In this case the convergence of the algorithms to the Bayes classifier is very slow [9]. This disadvantage will be manifested particularly strongly in the SRS classification when the dimension of the feature space and the class number are of the order of several tens and this is the real situation.

Due to the disadvantages demonstrated, the further procedure will consider the limits defined in (a, b. i).

7.3. Risk function computations. Case 7.2.a is computed when apriori information is available on the class distribution, their apriori probabilities and

the loss matrix. Advantages: minimum R_{Bayes} is guaranteed. Disadvantages: usually the apriori information is not available, assumptions in simplified version are needed to be adapted and to be updated in the process of Classifier compilation. Anyhow, when the form of distribution can be assumed with sufficent certainty (in this case the experience acquired in SRS is encouraging [14, 15, 16]) and there is a possibility to increase the representative sample to the volume required, the Bayes approach is recommended [17].

Case 7.2.b has the following advantages as compared with 7.2.a: R constant limits simplified computation; and disadvantages: R value obtained is not minimum at the expense of the cutoff 'tails' of the distributions in the formulation of the class constant limits. The upper limit of $R_{const.\ limit}$ may be calculated with the 'tails.'

7.4. Classification of newly added vector-observations. The referring of the newly added observations to a certain class in the Classifier storage (to none in particular) is defined with the verification of the non-equality system: in the case 7.2 as obtained with the principle of maximum likelyhood, and in the case 7.2.b, i — from the class limits.

7.5. Classifier learning. In the scheme suggested the learning phase is realized as follows:

Unsupervized learning: through the cluster programs in agreement with the introduced formal criterion in them for similarity between observations.

Supervised learning: (a) at the output of the cluster program where in interactive mode the operator controls the clustering quality and corrects the erroneously grouped observations, based on the available apriori information; (b) At the input of the major program: through apriori classified observation files, for example, from test field measurements; c) updating of the Classifier storage. The verification of the normal conditions guaranteed the necessary closeness of class distributions to the normal in the classifier [17] and, therefore, guarantees the Bayes mode of learning of the average vector at sample volume tending to infinity.

The supervised learning is performed with apriori information, and comparison between the confidence SRS regions (and their transformations) with those of the generalized features.

8. Reduction of the feature space dimension

In many cases the identification of a given set of classes can be performed with a beforehand given identification quality (for example, through the permissible maximum R value), with a portion of the information obtained in the experiment (for example, with the SRS values of some channels only). The resolution of this problem results in both direct economy of computations and in optimization of the technical problem with regard to the measurement systems and their exploitation capacity. The result from the resolution of such a problem may serve as a criterion for the effectiveness of certain SRS transformations. Some Basic methods for the reduction of dimensionality of feature space are given in [9].

The precise resolution requires a study of the complete combinatorics of subsurfaces formed with the subset of features (primary and generalized) and a definition of those combinations that satisfy the selected criterion of the classification quality. In the general case of unequal covariation matrices and non-linear decision functions this problem is difficult to be realized in regard to the computation efforts involved. But if we assume that the covariation matrices, though unequal, are diagonal (of diagonals equal to

those of the real matrices) and the classes are limited with multidimensional parallelepipeds, including the confidence ellipsoids whose axes are parallel to the coordinates due to the diagonality of their covariation matrices, then the problem reduces to a problem of technical diagnostics. A possible solution is suggested in [18], where the V value is the criterion for the classification quality. In the scheme thus suggested the problem is resolved with the combinatorics program. The restrictions introduced in the resolution of the problem in this case result in the following: if the program provides an answer that given subset of features is sufficient for the classification of the studied set of classes, then this conclusion is preserved in the general case also for the nondiagonal covariation matrices, but the risk function in the general case will be smaller. The reverse answer that the given feature subset is not sufficient for the set class recognition at a fixed value, is not a guarantee that at nondiagonal matrices the same result will be preserved.

9. Coupling of classified realizations with geographic coordinates and contouring of the spectral homogeneous regions

In order to resolve this problem, it is necessary to provide accompanying code for each vector-observation through which the geographic coordinates of RE can be compared. With the help of plotter-programs in a two-dimensional coordinate system, the coordinates of all the observations of a given class are plotted. When simultaneous visualizations are needed (graphs) of more than one class, several well distinguished symbols are used. The computation of the geometric characteristics of the contoured regions can be provided with upplementary programs: perimeter, surface, formfactor, etc.

10. Classifier storage updating

All classified and unclassified realizations entered in the Classifier are memorized according to the type of the classification programs (Bayes, multidimensional ellipsoid, etc.), and the type of the apriori SRS transformations.

After accumulating a given quantity of such data the updating is performed (on user's request) in two directions:

A. With the already classified observations the average vector and the covariation matrix of each available class in the storage is updated. It is advisable for the purpose to use numerical models for brief computations [9]. This approach assumes the implementation of nonmarked training sequences of SRS in the classification of new observations, i. e. for which it is not apriori known to which class they belong. Therefore, if in the classification mode at the input of the program the maked training sequence is fed, the information from the unclassified observations cannot be used for direct updating of resolution surface parameters. This may be performed, if the techniques of stochastic approximation are applied for the iterative determination of these parameters.

B. With the unclassified realizations to any classes in the storage and also with those that have not satisfied the automatic controls. For the purpose, an input file is composed from them and the complete training cycle is performed with it. Here it is possible to produce new classes.

In the process of compiling a thematic classifier of finite number M_{max} of forming classes it is possible to obtain as an intermediate result $R_{Bayes} < R_{const.\ limits}$ when the nonequality $M < M_{max}$ exists. This is due to the following: Bayes SRS classifier operates with theoretical normal distributions that are de-

fined in the interval $(-\infty, \infty)$ for each of the features. After the first learning stage (no updating was made) the class number is $M<M_{max}$ (in the real case at the recent stage of knowledge on SRS not always the apriori precise classifier content may be available even for relatively limited themes). On the other hand, in the next stage, classification, again for the real case of space and airborne data processing observations that may belong to $M+j$-th classes, $j=M+1, \ldots, M_{max}$ may be entered. If $M \ll M_{max}$, the distance between the centers of M-th classes is relatively large with regard to the available filling of the thematic feature space. Due to this and also to continuity of the normal distributions in the intervals $(-\infty, \infty)$, the Bayes decision functions limiting the available classes pass away from the class centers. Thus probability increases for observations belonging to the $M+j$-th class to be classified as belonging to some of the M-th classes if $M+j$-th class is close to some of the M-th classes. Such an error of II gender can be reduced, if the performed M classes are limited with, for example, confidence ellipsoids or parallelepipeds of smaller size and not with Bayes functions, but anyhow such that the sum of the cutoff tails of the normal distributions would be admissible, i. e. the error of I gender would not exceed the apriori limit. Then the possibility for realization from the $M+j$-th class to enter some of the M-th classes is considerably reduced, except for if in general the $M+j$-th class does not intercross largely with some of the M-th classes. After repeated updating the classifier will be refilled with new classes and when $M \to M_{max}$ then $R_{Bayes} \sim R_{const}$ will be ensured. Therefore, when the updating is performed: a) under condition $M < M_{max}$; b) with data of no apriori information as to whether they belong to the M-th formed class only, it is recommendable to complete the classifier, using confidence ellipsoids or parallelepipeds and only under $M \to M_{max}$ (the criterion for this may be the absence of new class formation after repeated updating within the theme) to apply the Bayes decision functions.

For the given confidence decision functions it is possible at fixed value M (by number and class content) to minimize the risk function [19]. Of course, the minimum R_{min} thus obtained will be larger than R_{Bayes}, if realizations belonging to the M-th class are entered only.

VI. Program package for classification of spectral reflectance signatures

When determining the structure and the content of a program package designed for the classification of spectral reflectance signatures (SRS), multiple considerations from both general and particular nature must be taken into account in view of the package effectiveness: scientific, economic and applicable. Some of these considerations were discussed in the previous chapters. Considerations, related to the experimental specifics of SRS obtaining and affecting the volume of the computation efforts under application of program package for SRS classification, may be generalized as follows:

1. In the real case of satellite and nonsatellite information use for classification or for learning files, thousand or tens of thousand SRS are applied (vector-realizations in multidimensional signature space) as obtained from individual elements of solution.

2. Class and subclass number that is interpreted within the limits of the thematical classifiers, for example, for agricultural purposes, is of the order of several tens.

3. The dimension of the primary feature space (the channel number of the multispectral devices) in which SRS are obtained is also several tens for the contemporary technical provisions.

4. The SRS distribution is in most of the cases sufficiently close to the normal, due to which there are grounds to apply the respective statistical methods elaborated for normal distributions.

The major operational modes for the SRS Classifier usually are: training, classification and updating. According to the selected algorithms for determination of the class characteristics, these three modes may be realized both parallel or in sequence, for example, the techniques of the stochastic approximation require parallel performance of training and updating to achieve more complete information use for the incrorrectly classified marked realizations in the training sample, but this involves larger computation efforts, that must be taken into account due to the considerations in 1, 2 and 3.

It follows from the considerations 1, 2 and 3 that the necessity of decrease in the dimension of the feature space, where the SRS classification or one of their transformations is performed, should be also considered. This can be realized through appropriate SRS transformations and determination of the minimal feature combination (under beforehand given class set) and thus to attain the given threshold value of the risk function or of another criterion on the classification quality [9]. In the general case when the covariation matrices of the classes are not equal, the relationship between the risk function R and the various criteria for classification quality is established rather nonreliably after reduction of the feature space.

Consideration 1 should be taken into account when selecting the cluster-algorithms with preference to those where the training matrix is used in series, column by column, with no necessity to be kept entirely in the operational memory, as is the case with the cluster-algorithms that optimize the criterion functions from intra- or intergroup distances, etc. In cases similar to the latter, the acceptable dimension of the training matrix is no more than 1000-2000 vector-realizations.

Consideration 4 makes possible the use of probabilistic methods for classification that are more precise than the cluster-algorithms or the determined discriminant functions (hiperplanar, etc.) and provide possibility to compute or evaluate the function of the average risk.

Accelerated computing procedures are applied to obtain separate mathematical functions as the probabilistic integral in the multidimensional feature space, the average vector and the covariational matrix of each class and the algorithms for their updating, etc. In [27] an accelerated procedure is suggested for classification by Bayes through replacement of some of Bayes decision functions in the course of the procedure with simpler criterion nonequalities, and also through appropriate transformations of the covariation matrix.

Significant alleviation of the computation efforts may be obtained through the application of the confidence hiperparallelepipeds as a prefilter to Bayes procedure of classification. In this case, the verification of the belonging of vector $\mathbf{X}$ to a given class is initiated with a system of nonequations

$$a_{ij} \leq X_i \leq b_{ij}, \quad i=1, \ldots, m, \quad j=1, \ldots, n, \tag{1}$$

where m is the dimension of x, n is the number of classes in the Classifier, and a_{ij} and b_{ij} are the limits of the hiperparallelepiped of the j-th class. These limits may be determined differently, for example, so that the respective confidence ellipsoid should be inscribed in it (determined at a given confidence level) or so that the shaping edges of the parallelepiped would equal the res-

pective main axes of the ellipsoid (then the latter will be incorporated into the hiperparallelepiped) and so on. In agreement with consideration 4, the dimensions of the ellipsoid can be determined on the basis of the normal distribution. If under a fixed j at least one of the non-equations (1) is not fulfilled, then the vector $\mathbf{x}$ does not belong to the j-th class (under selected confidence level). In the reverse case the answer that $\mathbf{x}$ belongs to the j-th class is not absolutely positive, because it is possible that the j-th hiperparallelepiped would have section $\Omega_{jk} \neq 0$ with another k-th hiperparallelepiped. Under a built thematical classifier the indices k of the classes for which $\Omega_{jk} \neq 0$ should be kept in its memory for a given j and also when it is established that the system (1) under fixed j is satisfied by $\mathbf{x}$, Bayes classification should be performed for those values of k only for which $\Omega_{jk} \neq 0$. Similarly, the classification with the help of hiperellipsoid may be performed. System (1) requires simple computation operations without necessity to compute the Mahalanobissian distance in a prestage and, therefore, we may expect significant fast action of the procedure described. It is possible when we have significant amount of closely located classes in the Classifier to perform a combination of the upper procedure with the one sugested in [27] at the second stage (when it is established that $\mathbf{x}$ satisfies (1) and $\Omega_{jk} \neq 0$ for several values of k) that will result in a greater acceleration of the classifying procedure, especially in the case when the covariational matrices of the classes have eigen numbers that differ in-between.

An example of the structure of a program package for SRS classification, taking into account the above-mentioned considerations, is shown in Fig. 3. The denomination of the subprograms and their destination are as follows:

1 — input matrix of data; *2* — multiplication of the input matrix with constant number; *3* — cluster-algorithms; *4* — SRS transformations; *5* — multiplication of vector-column by normal law; *6* — verification of normality of the cluster-formed distributions; *7* — determination of a mean vector and covariation matrix of the normally distributed sets of SRS vectors or of their transformations; (7_1 — through initial training file, 7_2 — through the method of stochastic approximation, 7_3 — through updating with accumulated data for correctly classified vector-realizations); *8* — control of ill-determined covariation matrix; *9* — computation of the risk function by Bayes; *10* — computation of the risk function under decision surfaces, composed with hiperparallelepipeds and hiperellipsoids; *11* — classification by Bayes; *12* — classification with hiperparallelepipeds and hiperellipsoids decision surfaces; *13* — subprogram for dimension decrease of the feature space; *14* — subprogram for yielding in *X-Y* spatial coordinate system (plotter) of the vector-realizations, classified to one class (spectrally homogenious regions).

File β contains the mean vectors μ and the covariation matrices K of the formed classes as follows: β_1 — without pretransformation on the input matrix; β_2 — after transformation of the input matrix with some of the transformations in subprogram *4*; β_3 — after fulfilment of subprograms *5*; β_4 — after fulfilment of subprograms *4* and *5*; β_5 — through training matrix in which the classes are formed by blocks.

File α contains the vector-realizations for updating.

In α_1 the information from following groups is stored:

a) clusters, where the quantity of realizations is smaller than a critical number pregiven by the user, the value of which is determined from the type of the criterion of normality in subprogram *6*;

b) clusters that have not satisfied the requirements for normality in *6*;

c) clusters or sets of vector-realizations, determined as classes of subprogram *5*, or through the training matrix, determined for file β. These clusters or sets have ill-determined covariation matrix.

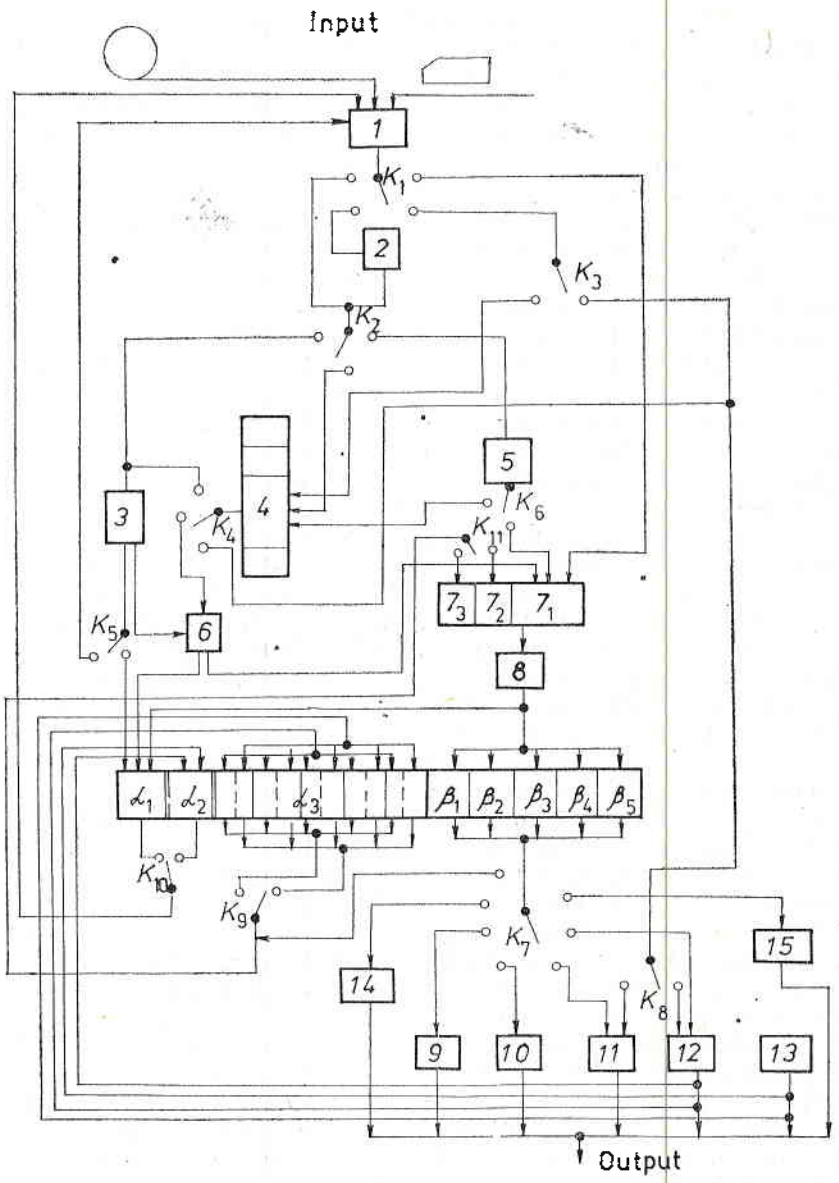

Fig. 3. An example of structure of program package for a spectral reflective signatures (SRS) classifier

Information on the vector-realizations that is not classified to any of the classes in file β is stored in α_2.

File α_3 contains vector-realizations classified to some of the classes in file β (information of the field number of file β together with information by which

subprogram the classification was performed: 11, respectively 12, is stored and this information is used in the updating).

Through the various states of variables $K_1, \ldots, K_{11}$, the mode of package operation is determined on the user's request.

Principal modes of operation of the package are:

Training.

a) With teacher: this is realized with marked training input matrices and the final result is the formation of the classes in β_1, β_2, β_5 and the computation of the risk function for the set of classes obtained;

b) Without teacher: this is realized through unmarked input matrices (the real case suggests operational conditions of incomplete apriori information) that pass through subprogram *3* (possibly in combination with *4*); through the variable K_5 the operator may verify the correctness of the clustering, based on the apriori information available, and then to exclude the incorrectly introduced vector-realization in a given cluster from the input matrix and to start again the selected sequence of subprograms with the new matrix thus formed; the final result is the formation of the classes in β_1 and β_2;

c) For research purposes files β_3 and β_4 are used in which model classes are formed through subprograms *5*, respectively *5* and *4*, in order to verify the possibilities of the various types of transformations, both with respect to the size of the risk function and to the decrease in the dimension of the feature space (for the purpose subprogram *15* is used).

Classification.

This mode of operation performs the classification of the unmarked vector-realizations through subprograms *11*, respectively *12*. If the classification is performed in field β_2, the vector-realizations are beforehand transformed with the transformation from *4* through which the classes in β_2 have been obtained. Finally decisions for the belonging of the vector-realizations are obtained in print, and the vectors are sent to α_1, respectively to α_2, for updating of the classes characteristics (using the stochastic approximation, the updating can be performed in the real time mode).

Updating.

The updating of the classes characteristics (mean vector and covariation matrix) from file β is performed in two modes:

a) through formation of an input matrix from data in α_1, α_2 or mixture of the two files after which operation is performed as in the 'training' mode;

b) from file α_3 in 7_b, respectively 7_c, the vector-realizations are introduced considering the field of file β that is updated. File α_3 contains the fields, corresponding to those of β and additional subfields in respect to which subprogram *11* or *12* was used to perform the classification. These are shown with dotted-line in Fig. 3.

The first version of the above-described program package is realized in the Central Laboratory for Space Research at the Bulgarian Academy of Sciences on FORTRAN IV and ASSEMBLER languages and under the control of OC/EC.

Chapters I to IV are composed based on works [25] and [28] and chapter V is based on work [26].

Conclusion

The approaches examined (major tasks) are of downward sequence of complexity, resulting from available possibilities: physical models, regression models, probabilistic-set classifier. Undoubtedly, the physical models are the most de-

sirable result for the experimenter and the interpreter, but they are resolvable in long-term perspective. The general regressive models are irresolvable in practice by now, although partial solution is acceptable in principle with the limitations as given above. The closest relative perspective for direct practical applications on a larger agricultural scale is that of the third approach, although the absence of sufficient apriori information and the disturbing atmospheric effect and other noise sources reduce its effectiveness.

Appendix 1

Risk function [9, 12]

Each SRS is represented as m-dimensional vector $\mathbf{r}:\{\mathbf{r}_i\}$, $i=1, \ldots, m$-SRS channel number. The compatible by a given criterion of similarity vectors $\mathbf{r}_{jk}$ form the k-th class of population: N_k, $k=1, \ldots, M$, $j=1, \ldots, N_k$.

Consider the set of M classes, $k=1, \ldots, M$ each with multidimensional normal distribution $f(\mathbf{r}/\mu_k)$ in the space of $\mathbf{r}_{jk}$, i. e. each class is characterized with an average vector $\boldsymbol{\mu}_k$ and a covariation matrix K_k.

With c_{kl} we denote the losses due to the case when the classifier makes decisions for availability of subject from the l-th class, while in reality there is a subject from the k-th class. The elements c_{kl} form the matrix C of the losses, $l=1, \ldots, M$.

We introduce decision surfaces S_{kl} for the classes k and l. The possibility to have error in unifold classification (of one vector $\mathbf{r}$) is:

$$p_{kl}=\int_{V_{kl}} f(\mathbf{r}/\mu_l)\, d\mathbf{r},$$

where V_{kl} is the volume, in which (through S_{kl}) the k-th class is defined. The average value of c_{kl} losses from all combination (k, l) and for multifold repeated classification (for many vectors $\mathbf{r}$) is called risk function R and is obtained from

$$R=\sum_{k=1}^{M} \int_{V_{kl}} \sum_{l=1}^{M} p_k c_{kl} f(\mathbf{r}/\mu_l)\, d\mathbf{r},$$

where p_k is apriori probability of the k-th class.

Usually it is assumed that $c_{kk}=0$ and $c_{kl}=\text{idem}=c$, i. e. no loss of correct classification occurs and in incorrect classification all the classes are assumed to be of equal weight with regard to the losses. Under this simplifying condition ensuing from the principle of maximum similarity it follows that the equations for s_{kl} are:

$$(1) \qquad S_{kl}: ln \frac{p_k}{p_l}+l_n \frac{|K_l|}{|K_k|} - \frac{1}{2}[(\mathbf{r}-\mu_k)' K_k^{-1} (\mathbf{r}-\mu_k)-(\mathbf{r}-\mu_l')K_l^{-1}(\mathbf{r}-\mu_l)]=0.$$

Equation (1) ensures the minimum R value.

Appendix 2

Variation coefficient of the Fourier series coefficients under linear approximation of discrete SRS

The real SRS are obtained in discrete shape: $\mathrm{r}(\lambda_i)$, $i=1, \ldots, m$. Under a large channel number, for example, $m=32$, the linear approximation is acceptably correct. In this case the coefficients a_k of the basic orthogonal functions $\sin \frac{k2\pi\lambda}{T}$ are expressed with [21]

$$a_k = -\frac{1}{\pi} \sum_{i=1}^{m} \frac{\beta_i}{k^2} \left(\cos \frac{k2\pi}{T} \lambda_i - \cos \frac{k2\pi}{T} \lambda_{i-1} \right) \tag{1}$$

$$= \frac{2 \sin \frac{k\pi}{T} \Delta\lambda}{\pi k^2 \Delta\lambda} \sum_{i=1}^{m} (r_i - r_{i-1}) \sin \left[\frac{k2\pi}{T} (\lambda_i + \lambda_{i-1})/2 \right].$$

We assume that r_i are normally distributed random values of dispersion $\sigma_i^2 = q^2 \mu_i^2$ where $q \ll 1$, for example, 0,03. Such an assumption of rigid proportion between σ and the average vector μ we introduce for convenience [18, 20] of the comparative analysis. We assume also that $r(\lambda)$ is a periodic function of period T. In reality this is not true, and in order to apply the Fourier series of $r(\lambda)$, we have to assume that $r(\lambda)$ performs a jump, for example, in λ_m, and attains the value of r_{λ_1}. This assumption in principle does not introduce difficulties in the classification, but requires the introduction of additional terms in the transformation to describe the nonlinearity of the jump. The periodicity of the function thus defined makes the coefficients a_k independent between themselves, for example, [9]. Then the mean square deviation of $\bar{a}_k$ from (1) equals to

$$\sigma_k = q \frac{2 \sin \frac{k\pi}{T} \Delta\lambda}{\pi k^2 \Delta\lambda} \sqrt{\sum_{i=1}^{m} \left\{ (\mu_i^2 + \mu_{i-1}^2) \sin^2 \left[\frac{k2\pi}{T} \frac{\lambda_i + \lambda_{i-1}}{2} \right] \right\}}. \tag{2}$$

The coefficient of variation of a_k is respectively equal to

$$V_k = \frac{\sigma_k}{\mu_k} = q \frac{\sqrt{\sum_{i=1}^{m} \left[(\mu_i^2 + \mu_{i-1}^2) \sin^2 \frac{k\pi}{T} (\lambda_i + \lambda_{i-1}) \right]}}{\sum_{i=1}^{m}{}' (\mu_i - \mu_{i-1}) \sin \frac{k\pi}{T} (\lambda_i + \lambda_{i-1})}. \tag{3}$$

For comparison we may use the variation coefficient of r_i

$$V_i = \sigma_i / \mu_i = q. \tag{4}$$

The comparison between (3) and (4) shows that it is possible even for small k values, i. e. for main harmonics, to obtain $V_k > 1$, because the expression in the denominator of (3) is a sum of terms of different signs. For example, for the discrete series of 32 values (linearly approximated) for $r(\lambda)$: 53, 51, 48, 46, 43, 42, 41, 41, 40, 40, 41, 41, 42, 38, 36, 28, 29, 48, 53, 53, 98, 85, 65, 50, 38, 33, 35, 30, 27, 45, 53 under $k=5$ we obtain: $V_k = V_5 \approx 8{,}8\ q$.

When the discrete Fourier transformation is applied to obtain V_k the well-known relationship between the Fourier series coefficients and its discrete version has to be applied [21].

Appendix 3

I. Autocorrelative function (for example [22])

$$K(\tau_j)=\sum_{i=1}^{m}[r(\lambda_i)-\bar{r}(\lambda)]\,[r(\lambda_i+\tau_j)-\bar{r}(\lambda)],$$

where

$$\bar{r}(\lambda_i)=\sum_{i=1}^{m} r(\lambda_i)/m, \quad j=0, 1, 2, \ldots, m-1, \tau_j=(\lambda_i-\lambda_{i-1})\,j=\Delta\lambda\,.\,j.$$

II. Power autocorrelative function [8, 18, 23]

$$C^{(n)}(\tau_j)=\sum_{i=1}^{m}|r(\lambda_i)-r(\lambda_i+\tau_j)|^n$$

n is given by the user, for example, $n=0{,}5\,;\ 1\,;\ 2\,;\ldots$ (under $n=2$, the Kolmogorov function is obtained [22])

$j=1, \ldots, m/2$ for m even,

$j=1, \ldots, (m+1)/2$ for m odd,

τ_j as for $K(\tau_j)$.

III. Integral (mean arithmetic) transformation [24]

$$I(\lambda_l)=\frac{1}{l}\sum_{j=1}^{l} r(\lambda_j), \quad l=1, \ldots, m.$$

IV. Combined transformation [24]

$$L(\tau_j)=C_j\,.\,I[r(\lambda)]=\sum_{i=1}^{m}|I(\lambda_i)-I(\lambda_i+\tau_j)|^n,$$

$j=1, \ldots, m/2$ for m even,

$j=1, \ldots, (m+1)/2$ for m odd.

Transformations I, II and IV are invariant with respect to an additive constant, i. e. they are filter of systematic, additive and apriori unknown errors.

All the transformations shown are irreversible. The ratios between the coefficients of variation of I, II and untransformed SRS are obtained in [18, 20] under the following limiting conditions: a) $\sigma_i=q\,.\,r_i$, $q\ll 1$; b) random values r_i and r_j are interindependent; c) the differences between the uniform components of the mean vectors r_k and r_l of two classes are small and are measured with $\theta\ll 1$. Under these conditions, from [18, 20] it follows that the variation coefficient V_c of $C^{(n)}(\tau_j)$ is minimum at $n=1$ (n-integer) and it follows from [23] that $V_C^{(n=1)}$ is smaller than V_k and with $m>8$ also from $V_r\,.\,V_r$ is the

variation coefficient of untransformed SRS. The last estimate: $m>8$ is approximative.

The upper ratios are deduced under conditions that differ from reality (for example, the covariation matrices are not diagonal). They may serve as orientation in the effectiveness evaluation of the transformations exposed.

In [24] the evaluation of transformations III and IV is performed which exhibits their effectiveness with reference to the reduction of the risk function R value. The evaluation is performed under limiting conditions similar to the above-mentioned.

Appendix 4

Enthropic and divergence transformations

I. Enthropic transformation [9, 12].

The matrix A of the linear transformation $\varkappa=A\mathbf{r}$ is determined so as to minimize the enthropy

$$H_j=-\int_{(\mathbf{r})} p(\mathbf{r}/\omega_j)\ln p(\mathbf{r}/\omega_j)d\mathbf{r},\quad j=1,\ldots,M,$$

where the integration is in the $\mathbf{r}$ space of SRS, ω_j is denomination of the j-th class, and $p(\mathbf{r}/\omega_j)$ is density of probability distribution for the j-th class. Maximum uniformity (structuring) of set $\{p(\mathbf{r}/\omega_j)\}$ corresponds to $H_{\min}$. This corresponds to minimization of the dispersion in various distributions and may be expected to improve the identification of the classes.

Under equal covariation matrices K of the M-th classes, A is obtained as a matrix of eigen vectors.

II. Divergence transformation [9, 12].

The matrix A is determined for the linear transformation $\varkappa=A\mathbf{r}$, so as to maximize the divergence (difference in information) between the j-th and k-th classes

$$I_{jk}=\int_{(\mathbf{r})}[p_j(\mathbf{r})-p_k'(\mathbf{r})]\ln\frac{p_j(\mathbf{r})}{p_k(\mathbf{r})}d\mathbf{r}.$$

Under equal covariation matrices $K_j=K_k$ the maximum divergence thus obtained equals the risk function for j and k classess under the condition that $c_{jk}=c_{kj}$ (see Appendix 2). Therefore, it is assumed that I_{jk} here may act as a measure of distinguishing between classes j and k.

In the real case $K_j\neq K_k$ and $\mu_j\neq\mu_k$. Therefore, the definition of A relates to large computing difficulties. Moreover, there is no general method to maximize I for a set of M classes so that secure reduction of R may be obtained.

References

1. Himmelblau, D. M. Process Analysis by Statistical Methods. Moscow, 1973 (Russ.).
2. Yanev, T., D. Mishev. Experimental planning for complex test site studies of spectral reflectance signatures of natural formations. — Seminar on Test Site Studies in Bulgaria, Asenovgrad, 1976.
3. Yanev, T., A. Stoimenov. Optimizing the multispectral volume on terrestrial studies of natural formations. — Compt. rend. Bulg. Acad. Sci., 31, 1978, 11.
4. Mishev, D., R. Kancheva. Test sites. — UN/FAO Remote Sensing Training Seminar. Sofia, 1981.
5. Mishev, D. Remote Sensing of the Earth. Sofia, 1981.

6. Milenkiy, A. Classification of Signal Under Indeterminancy. Moscow, 1975 (Russ.).
7. Mishev, D. Spectral reflective characteristics of natural formations and their application for remote sensing purposes. — Rem. sensing of the earth, 1980, 1.
8. Yanev, T., G. Astardgiyan, Ch. Nachev. Feature set yield and decision-making in analysing monodimentional objects. Large systems of information and control. — 4th Polish-Bulgarian Seminar, Varna, 24-25 Oct., 1973.
9. Tou, J., R. Gonzalez. Pattern Recognition Principles. Moscow, 1978 (Russ.).
10. Mishev, D., T. Yanev. Separation of mixture from spectral reflective characteristics within a pixel. — University Annual. Technical Physics, 12, 1984, 1.
11. Classification and Clustering. Edit. Ryzin van J. Moscow, 1980 (Russ.).
12. Young, T., Th. Calvert. Classification, Estimation and Pattern Recognition. New York, 1974.
13. Patrick, E. Fundamentals of Pattern Recognition. Moscow, 1980 (Russ.)
14. Fu, K. S., D. A. Langrebe, T. L. Phillips. Informational Processing of Agricultural Data Obtained through Distance Measurements. — TIEP, 1968, MX.
15. Kondratiev, I., A. Grigoriev, O. Pokrovski. Informational Content of Remote Sensing Data on Environmental and Natural Resources Parameters. Leningrad, 1975.
16. Remote Multispectral Sensing in Agriculture, Laboratory for Agricultural Remote Sensing. — Annual Rep., Lafayette, Indiana, 1970.
17. Duda, R., P. Hart. Pattern Classification and Scene Analysis. Moscow, 1976. (Russ.)
18. Yanev, T., D. Mishev. Discriminant analysis of natural formation reflective characteristics by a minimal number of wavelengths. — XVIII COSPAR Mtg., Varna 29 May—7 June, 1975.
19. Yanev, T. Optimization of a parallelepiped discriminant function at a multidimensional analysis. — Space Res. in Bulgaria, 1978, 1.
20. Yanev, T., D. Mishev. Multidimensional classification analysis of spectral reflective characteristics obtained by aircraft, satellites and earth surveys. — XIX COSPAR, Philadelphia, June 1976.
21. Hamming, R. Numerical Methods for Scientists and Engineers. Sofia, 1974 (Bulg.).
22. Romanenko, A., G. Sergeev. Problems of the applied analysis of random processes. — Sov. Radio, Moscow, 1968 (Russ.).
23. Yanev, T. Power autocorrelative function. — Space Res. in Bulg., 1978, 1.
24. Mishev, D., T. Yanev. Comparison of discriminating possibilities of some transformations of spectral reflective characteristics of natural formations. — Compt. rend. Bulg. Acad. Sci., 31, 1978, 5.
25. Yanev, T., D. Mishev. Mathematical-statistical methods for object classification through spectral reflective characteristics. — UN/FAO Regional Training Seminar on Remote Sensing Application for Land Resources. 16-26 Sept., 1981, Sofia.
26. Yanev, T., D. Mishev. Program package for classification of spectral reflectance signatures. — First Bulgarian-Soviet Seminar on the Results from Bulgaria-1300 Project. Sofia, December, 1981.
27. Minskiy, D. E., A. M. Chizhevskiy. Fast Classification of Multizonal Imagery by Bayes. — Remote Sensing of the Earth. 1982, 1.
28. Yanev, T., D. Mishev. Methodics for the analysis of the spectral characteristics. — Mtg. of the Working Group of the Socialist Countries on Remote Sensing of the Earth with Air- and Space-Borne Techniques within the Framework of the Intercosmos Program. 25-30 May, Smolenište, Czechoslovakia, 1981.

Математико-статистические методы классификации спектральными отражательными характеристиками

Т. Янев, Д. Мишев

(Р е з ю м е)

Проведен сравнительный анализ существующих методов классификации спектральными отражательными характеристиками (СОХ) природных образований. Предложены методика и блоковая структура пакета программ для классификации СОХ и актуализации параметров классификатора.

Методика основана на подходе Бейеса в режиме с учителем и на кластерном анализе в режиме без учителя. Выполняются предварительные преобразования СОХ с целью уменьшения функции риска. Определяется минимальное число каналов прибора получения СОХ, достаточных в данном тематическом классификаторе для достижения заданной функции риска.