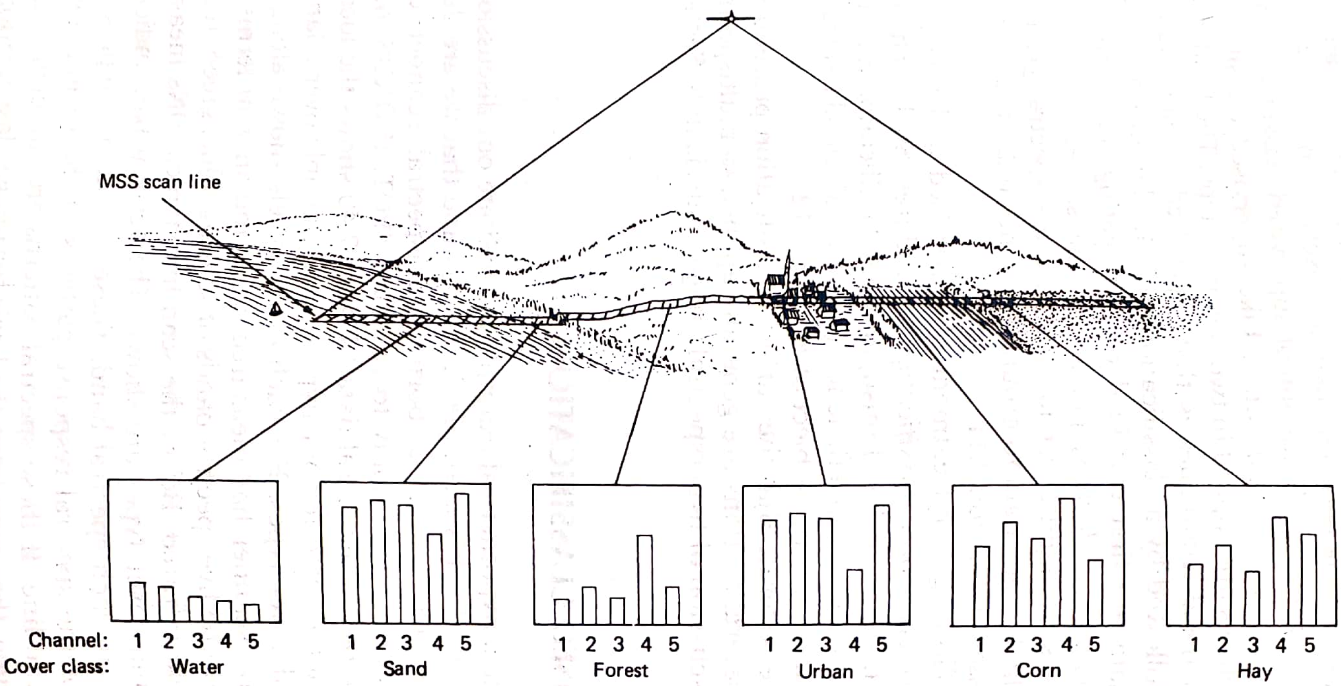

Figure 7.40 Selected multispectral scanner measurements made along one scan line. Sensor covers the following spectral bands: 1, blue; 2, green; 3, red; 4, near infrared; 5, thermal infrared.

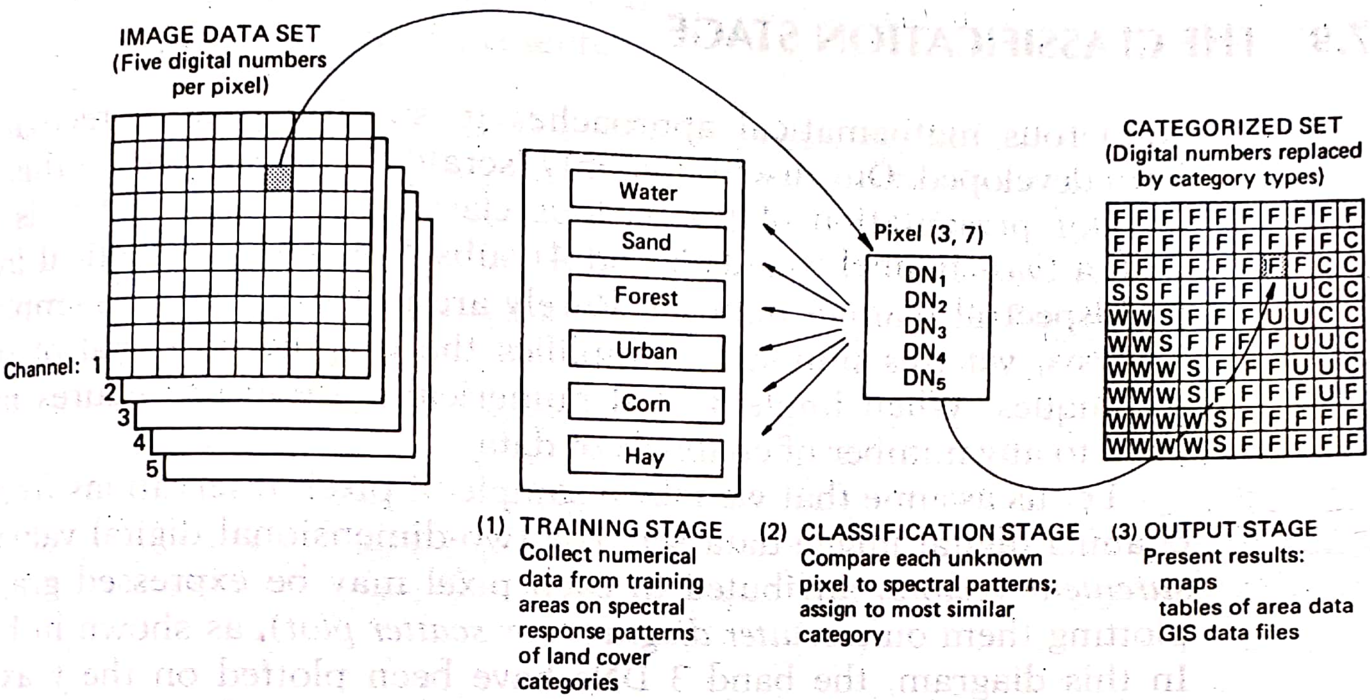

Figure 7.41 Basic steps in supervised classification.

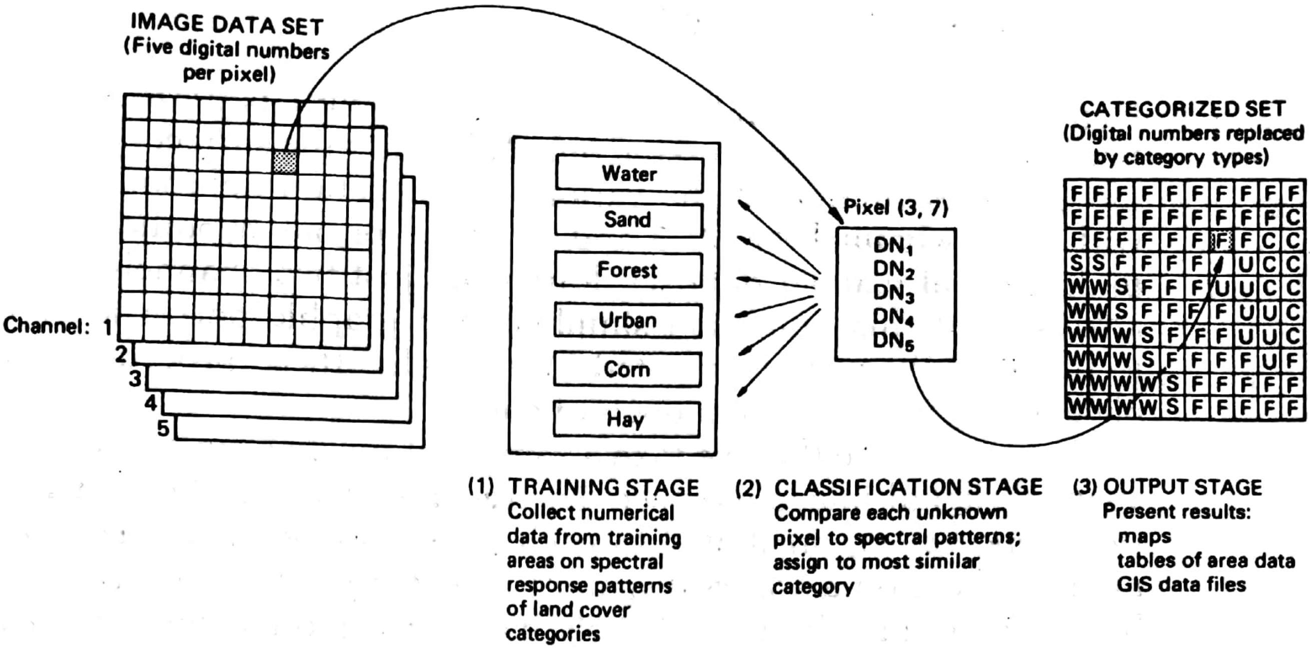

Figure 7.41 Basic steps in supervised classification.

Figure 7.41 summarizes the three basic steps involved in a typical supervised classification procedure. In the *training stage* (1), the analyst identifies representative training areas and develops a numerical description of the spectral attributes of each land cover type of interest in the scene. Next, in the *classification stage* (2), each pixel in the image data set is categorized into the land cover class it most closely resembles. If the pixel is insufficiently similar to any training data set, it is usually labeled "unknown." The category label assigned to each pixel in this process is then recorded in the corresponding cell of an interpreted data set (an "output image"). Thus, the multidimensional image matrix is used to develop a corresponding matrix of interpreted land cover category types. After the entire data set has been categorized, the results are presented in the *output stage* (3). Being digital in character, the results may be used in a number of different ways. Three typical forms of output products are thematic maps, tables of full scene or subscene area statistics for the various land cover classes, and digital data files amenable to inclusion in a GIS. In this latter case, the classification "output" becomes a GIS "input."

We discuss the output stage of image classification in Section 7.14. Our immediate attention is focused on the training and classification stages. We begin with a discussion of the *classification* stage because it is the heart of the supervised classification process—during this stage the spectral patterns in the image data set are evaluated in the computer using predefined decision rules to determine the identity of each pixel. Another reason for treating the classification stage first is because familiarity with this step aids in understanding the requirements that must be met in the training stage.

7.9 THE CLASSIFICATION STAGE

Numerous mathematical approaches to spectral pattern recognition have been developed. Our discussion only "scratches the surface" of this topic.

Our presentation of the various classification approaches is illustrated with a two-channel (bands 3 and 4) subset of our hypothetical five-channel multispectral scanner data set. Rarely are just two channels employed in an analysis, yet this limitation simplifies the graphic portrayal of the various techniques. When implemented numerically, these procedures may be applied to any number of channels of data.

Let us assume that we take a sample of pixel observations from our two-channel digital image data set. The two-dimensional digital values, or *measurement vectors*, attributed to each pixel may be expressed graphically by plotting them on a *scatter diagram* (or *scatter plot*), as shown in Figure 7.42. In this diagram, the band 3 DNs have been plotted on the y axis and the band 4 DNs on the x axis. These two DNs locate each pixel value in the two-dimensional "measurement space" of the graph. Thus, if the band 4 DN for

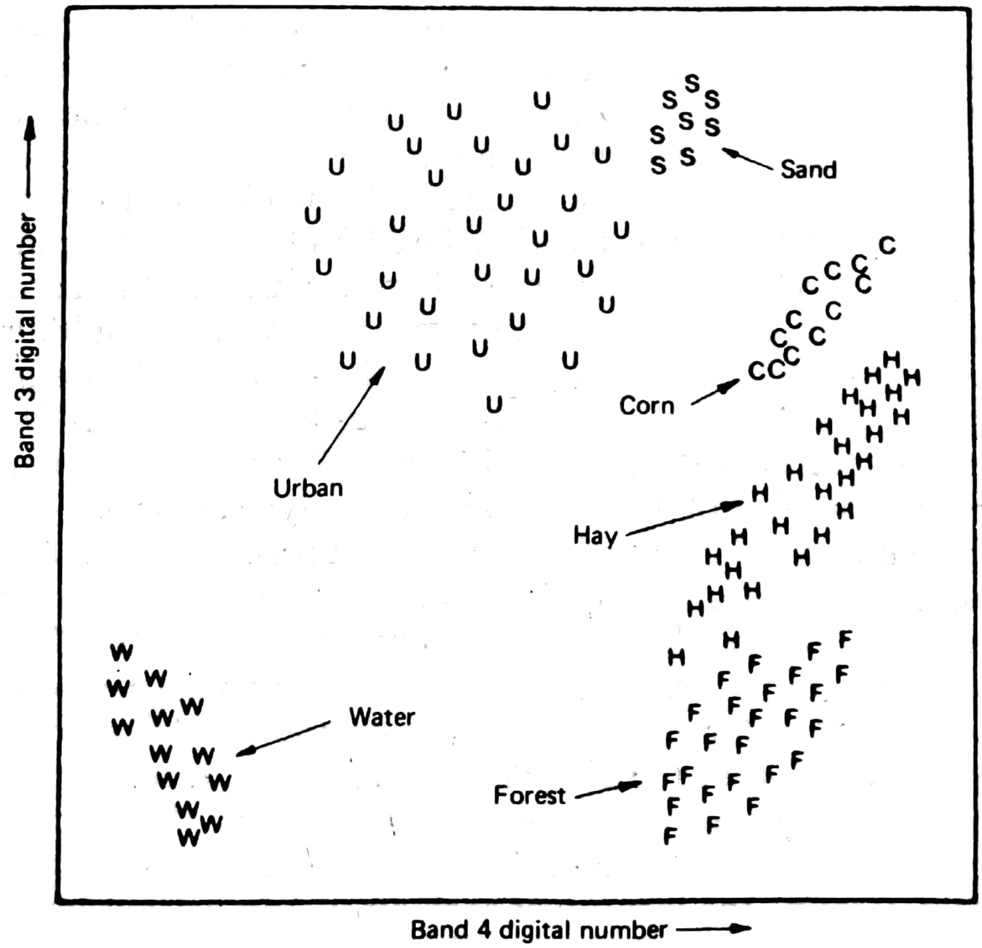

Figure 7.42 Pixel observations from selected training sites plotted on scatter diagram.

a pixel is 10 and the band 3 DN for the same pixel is 68, the measurement vector for this pixel is represented by a point plotted at coordinate (10, 68) in the measurement space.[1]

Let us also assume that the pixel observations shown in Figure 7.42 are from areas of known cover type (that is, from selected training sites). Each pixel value has been plotted on the scatter diagram with a letter indicating the category to which it is known to belong. Note that the pixels within each class do not have a single, repeated spectral value. Rather, they illustrate the natural centralizing tendency—yet variability—of the spectral properties found within each cover class. These "clouds of points" represent multidimensional descriptions of the spectral response patterns of each category of cover type to be interpreted. The following classification strategies use these "training set" descriptions of the category spectral response patterns as interpretation keys by which pixels of unidentified cover type are categorized into their appropriate classes.

Minimum-Distance-to-Means Classifier

Figure 7.43 illustrates one of the simpler classification strategies that may be used. First, the mean, or average, spectral value in each band for each category is determined. These values comprise the *mean vector* for each category. The category means are indicated by +'s in Figure 7.43. By considering the two-channel pixel values as positional coordinates (as they are portrayed in the scatter diagram), a pixel of unknown identity may be classified by computing the *distance* between the value of the unknown pixel and each of the category means. In Figure 7.43, an unknown pixel value has been plotted at point 1. The distance between this pixel value and each category mean value is illustrated by the dashed lines. After computing the distances, the unknown pixel is assigned to the "closest" class, in this case "corn." If the pixel is farther than an analyst-defined distance from any category mean, it would be classified as "unknown."

The minimum-distance-to-means strategy is mathematically simple and computationally efficient, but it has certain limitations. Most importantly, *it is insensitive to different degrees of variance in the spectral response data*. In Figure 7.43, the pixel value plotted at point 2 would be assigned by the distance-to-means classifier to the "sand" category, in spite of the fact that the greater variability in the "urban" category suggests that "urban" would be a more appropriate class assignment. Because of such problems, this classifier is not widely used in applications where spectral classes are close to one another in the measurement space and have high variance.

[1]Pattern recognition literature frequently refers to individual bands of data as *features* and scatterplots of data as *feature space plots*.

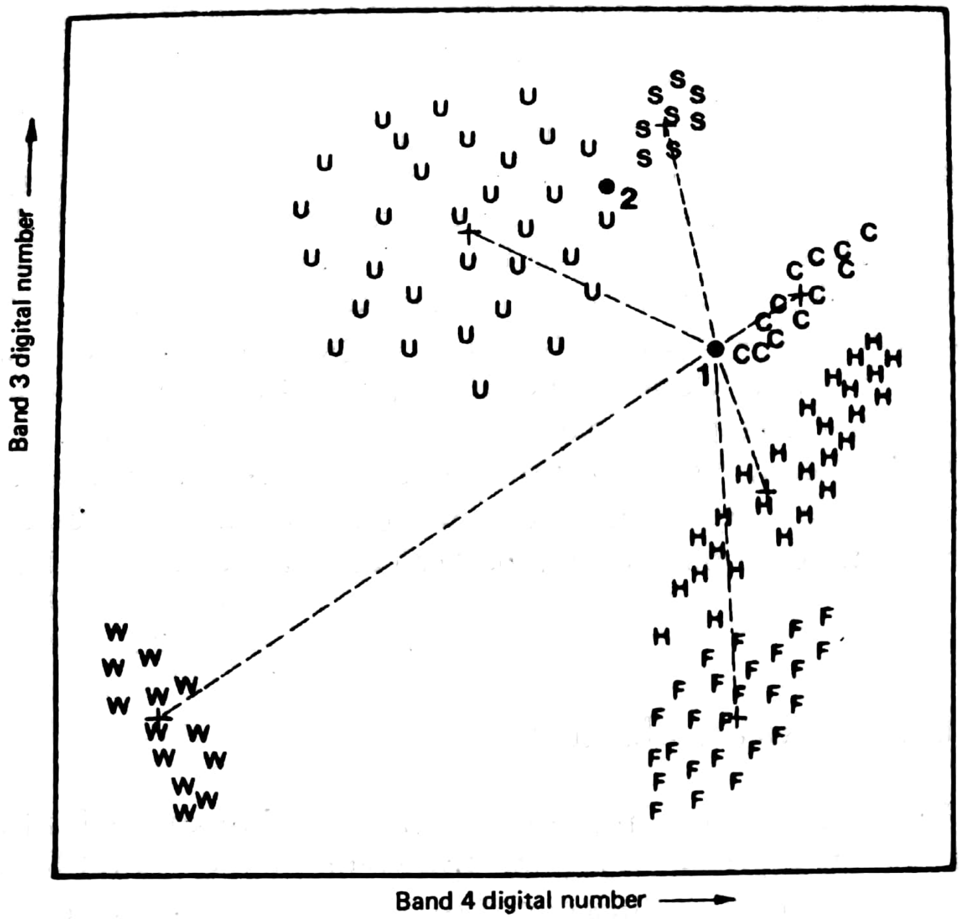

Figure 7.43 Minimum distance to means classification strategy.

Parallelepiped Classifier

We can introduce sensitivity to category variance by considering the *range* of values in each category training set. This range may be defined by the highest and lowest digital number values in each band and appears as a rectangular area in our two-channel scatter diagram, as shown in Figure 7.44. An unknown pixel is classified according to the category range, or *decision region,* in which it lies or as "unknown" if it lies outside all regions. The multidimensional analogs of these rectangular areas are called *parallelepipeds*, and this classification strategy is referred to by that tongue-twisting name. The parallelepiped classifier is also very fast and efficient computationally.

The sensitivity of the parallelepiped classifier to category variance is exemplified by the smaller decision region defined for the highly repeatable "sand" category than for the more variable "urban" class. Because of this, pixel 2 would be appropriately classified as "urban." However, difficulties are encountered when category ranges overlap. Unknown pixel observations that occur in the overlap areas will be classified as "not sure" or be arbitrarily placed in one (or both) of the two overlapping classes. Overlap is caused largely because category distributions exhibiting *correlation* or high *covariance* are

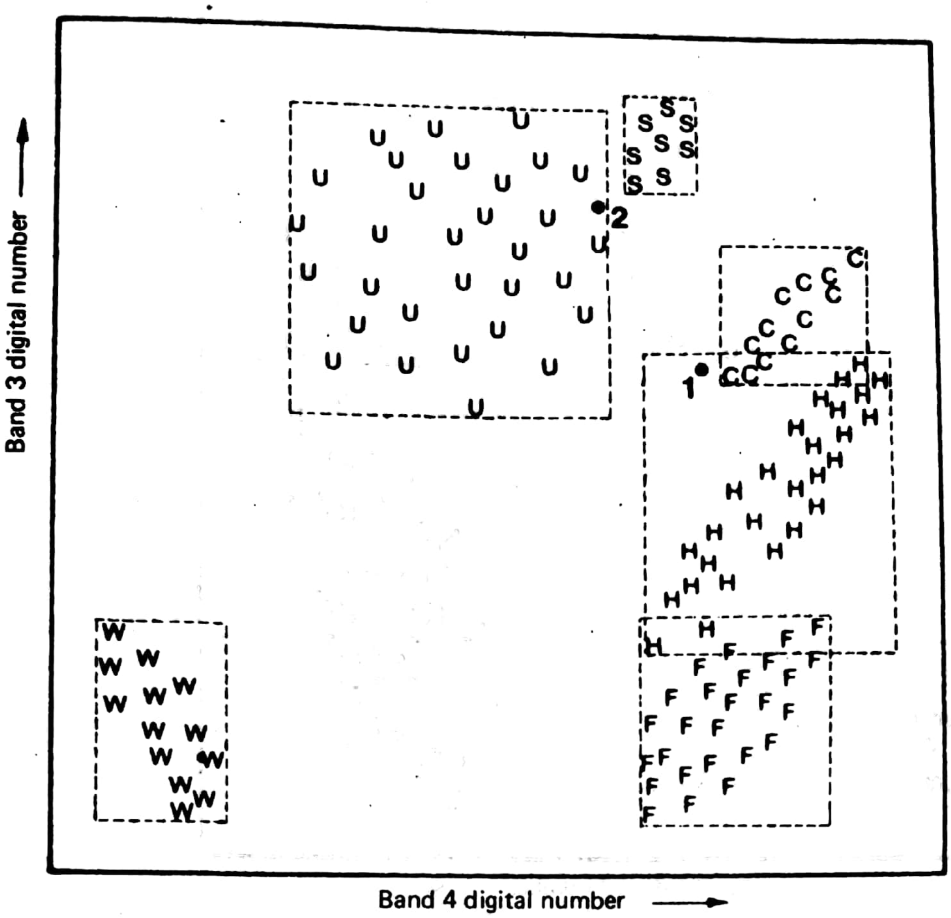

Figure 7.44 Parallelepiped classification strategy.

poorly described by the rectangular decision regions. Covariance is the tendency of spectral values to vary similarly in two bands, resulting in elongated, slanted clouds of observations on the scatter diagram. In our example, the "corn" and "hay" categories have positive covariance (they slant upward to the right), meaning that high values in band 3 are generally associated with high values in band 4, and low values in band 3 are associated with low values in band 4. The water category in our example exhibits *negative covariance* (its distribution slants down to the right), meaning that high values in band 3 are associated with low values in band 4. The "urban" class shows a lack of covariance, resulting in a nearly circular distribution on the scatter diagram.

In the presence of covariance, the rectangular decision regions fit the category training data very poorly, resulting in confusion for a parallelepiped classifier. For example, the insensitivity to covariance would cause pixel 1 to be classified as "hay" instead of "corn."

Unfortunately, spectral response patterns are frequently highly correlated, and high covariance is often the rule rather than the exception. The resulting problems can be somewhat alleviated within the parallelepiped classifier by modifying the single rectangles for the various decision regions into a series of rectangles with stepped borders. These borders then describe tion for each spectral category. 3) MAXIMUM LIKELIHOOD:

The probability density functions are used to classify an unidentified pixel by computing the probability of the pixel value belonging to each category. That is, the computer would calculate the probability of the pixel value occurring in the distribution of class "corn," then the likelihood of its occurring in class "sand," and so on. After evaluating the probability in each category, the pixel would be assigned to the most likely class (highest probability value) or be labeled "unknown" if the probability values are all below a threshold set by the analyst.

In essence, the maximum likelihood classifier delineates ellipsoidal "equiprobability contours" in the scatter diagram. These decision regions are shown in Figure 7.47. The shape of the equiprobability contours expresses the sensitivity of the likelihood classifier to covariance. For example, because of this sensitivity, it can be seen that pixel 1 would be appropriately assigned to the "corn" category.

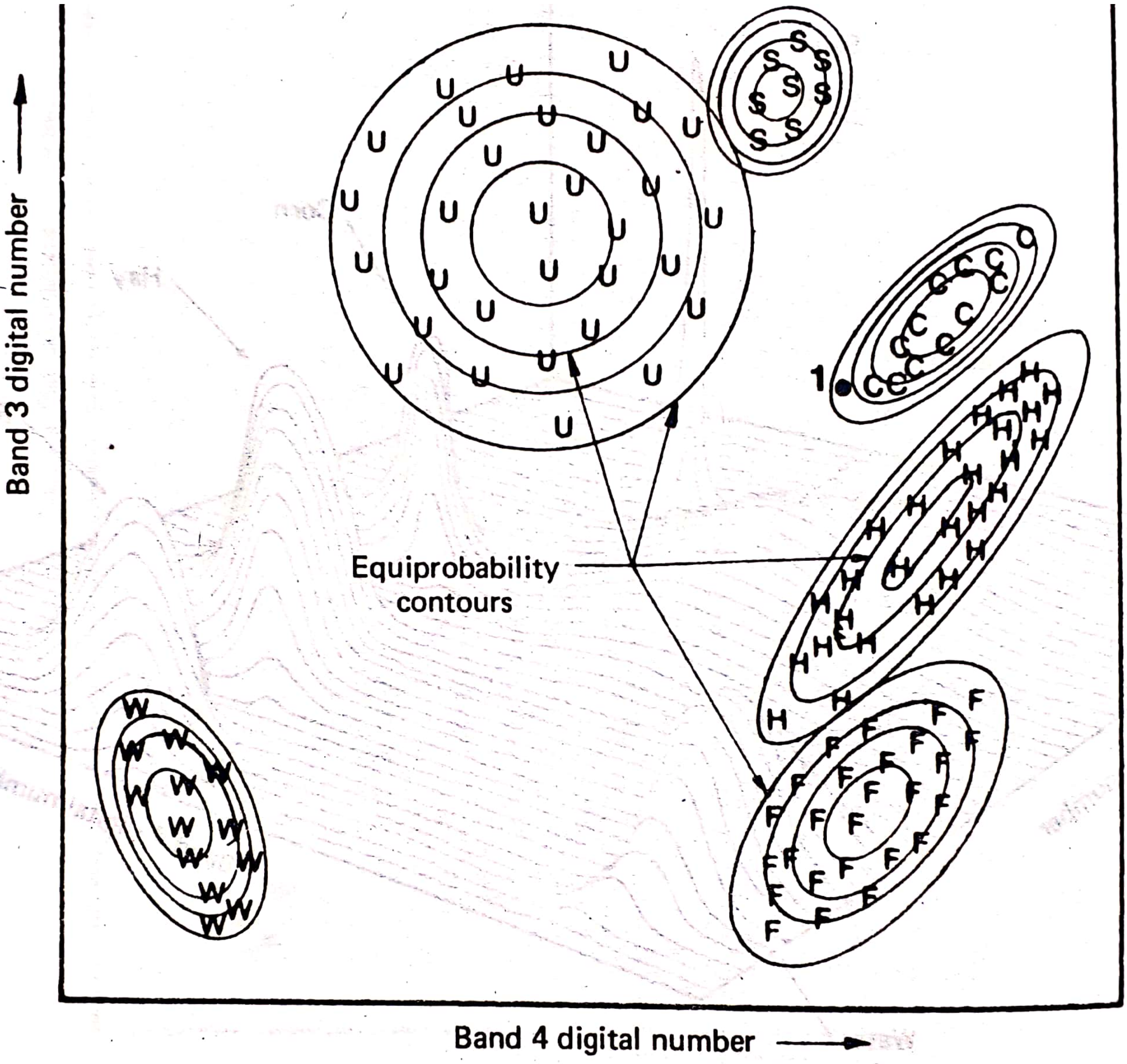

Figure 7.47 Equiprobability contours defined by a maximum likelihood classifier.