

Reprinted from

**Symposium on
Machine Processing of
Remotely Sensed Data**

June 29 - July 1, 1976

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
76CH1103-1 MPRSD

Copyright © 1976 IEEE
The Institute of Electrical and Electronics Engineers, Inc.

Copyright © 2004 IEEE. This material is provided with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the products or services of the Purdue Research Foundation/University. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

HINDU: HISTOGRAM INSPIRED NEIGHBORHOOD

DISCERNING UNSUPERVISED PATTERN

RECOGNITION SYSTEM*

Belur V. Dasarathy
Computer Sciences Corporation
Huntsville, Alabama 35802

ABSTRACT

A newly developed system for pattern recognition in unsupervised environments, capable of processing large volume data sets with minimal computational resources and human intervention, which is currently operational at the Marshall Space Flight Center, is presented in this study. Here, in this system, the problem of unsupervised learning is viewed as one of clustering the large volume multidimensional data sets and is approached through the novel gambit of terrain development in the multidimensional histogram space. The terrain is developed by connecting each histogram cell to all of its higher density neighbors. This process leads to amalgamation of all the cells belonging to each of the clusters. Certain of these cells, being connected to more than one cluster, define fuzzy boundaries between the clusters. Discriminant hyperplanes, which not only separate these clusters but also form least square fits to the centroids of the cells defining the fuzzy boundaries, are derived. The design of these discriminant functions is through a new algorithm developed specifically for catering to this problem environment of discriminating between clusters with fuzzy boundaries. The dimensionality curse, an often encountered problem of computational complexity arising from higher dimensionality of data sets, is tackled here by a relatively simple pre-processing technique of ordering and selecting features on the basis of an adhoc histogram information measure sensed for each of the features. The conceptual and computational claims of this HINDU system, presently in operational status on an IBM 360/65, have been verified by simulation tests using remotely sensed multispectral LANDSAT data.

*This work was performed under Contract NAS8-31639 for the Data Systems Laboratory National Aeronautics and Space Administration, George C. Marshall Space Flight Center, Huntsville, Alabama.