1-1-1975

An Incisive Neighborhood Discerning Iteratively Aggregating Nomenclator (Indian) System for Pattern Recognition in Unsupervised Environments

Belur V. Dasarathy

Follow this and additional works at: http://docs.lib.purdue.edu/lars_symp

Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

June 3 - 5, 1975

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
75CH1009-0 -C

AN INCISIVE NEIGHBORHOOD DISCERNING
ITERATIVELY AGGREGATING NOMENCLATOR
(INDIAN) SYSTEM FOR PATTERN RECOGNITION
IN UNSUPERVISED ENVIRONMENTS

Belur V. Dasarathy
Computer Sciences Corporation
Huntsville, Alabama 35802

ABSTRACT

AUTOMATIC CLASSIFICATION OF AIRCRAFT

AND SATELLITE IMAGES USING

MIXED INTEGER PROGRAMMING

M. Rebollo and L. F. Escudero

IBM Center
IBM Espana
International Business Machines, S. A. E.
Madrid, Spain

A novel system for pattern recognition in
unsupervised environments, which combines the
conceptual elegance of clustering schemes based
on inter-sample distance measures with the
computational simplicity of histogram approaches,
is presented in this study. The multi-dimensional
histogram of the entire data set is first derived
and by scanning this histogram space, the signifi-
cant hills therein are identified. The centroids of
these hills are deemed to be representative of the
given input sample set. This representative
pseudo-sample set is then input to the CURRY
system (International Journal of System Science,
Vol. 6, No. 1, January 1975, pp 23-32), which
has the innovative capability of self learning the
number of clusters inherent in the environment,
to derive the nuclei of these inherent clusters.
The total input data set is then clustered with
these cluster nuclei as prototypes. The two major
advantages of this new approach are:

- The conceptual satisfaction of lessening
  the sensitivity, of the clustering
  approaches based on inter-sample
  distance measures, to individual out-
  liers of the sample distributions through
  selection of a representative pseudo-
  sample set.

- The computational economy achieved
  in processing large data sets, such as
  those arising in remote sensing envi-
  ronments, through the choice of a
  significantly smaller but representative
  subset of pseudo-samples.