

A SPATIAL CLUSTERING ALGORITHM
FOR IMAGERY DATA

Its'hak Dinstein
Robert M. Haralick

Space Technology Laboratories
University of Kansas
Lawrence, Kansas 66044

Abstract: Cluster analysis has been applied to imagery data in different disciplines. In many cases, clustering algorithms do not take into account the spatial relation between vector measurements. A clustering algorithm was developed with the assumption that vector measurements of a spatial neighborhood tend to belong to the same class except at class boundaries. The algorithm consists of two basic steps: (1) detection of homogeneous areas and (2) assignment of those areas to clusters. The detection of homogeneous areas is accomplished by computing the gradient image and then thresholding the gradient image by its mean. Areas with gradient values less than the mean are considered to be homogeneous. After a cleaning process eliminates small and fuzzy homogeneous regions, the homogeneous areas are assigned to clusters. The largest homogeneous area is assigned to the first cluster, and then other homogeneous areas, the means of which are within a threshold distance from the cluster mean, are assigned to that cluster. Out of the unassigned homogeneous areas, the largest one is assigned to the next cluster and the process continues till all the homogeneous areas are assigned to clusters. Non-homogeneous areas are left unassigned.

This algorithm was applied to two multispectral images. The first one was a four band ERTS image (Monterey Bay, California - North). The image was partitioned into 8 x 8 resolution cell sub-images and the means of those sub-images were spatially clustered into four clusters corresponding to Water, Coastal Forest, Urban Area, Wood and Grass Lands. The second image consisted of multi-band aerial photographs taken by the University of Kansas, CRINC, over portion of the Missouri River near Parkville, Missouri. Three clusters were generated, corresponding to Water, Cultivated Land and Trees. In both cases, the areas assigned to a cluster were most representative of their respective categories.