GEODESIC DISTANCE AND MST BASED IMAGE SEGMENTATION

*George Economou, Vassilios Pothos and Apostolos Ifantis**

Electronics Laboratory, Department of Physics, University of Patras, GR-26500 Patras, Greece
phone: +302610997463, fax: +302610997456, email: *economou@physics.upatras.gr*
*Technological Educational Institute of Patras

ABSTRACT

In this work, a new approach is proposed for the integration of spatial proximity information in graph based segmentation algorithms. This is done by means of the geodesic distance. Distance calculation and the implementation of the method are carried out using the minimal spanning tree (MST), constructed on a watershed image partition. Distance, defined over the MST edges, presents a measure of both spatial and feature coherence. It is incorporated in MST based color image segmentation applications, by means of a new density feature, which is computed with spatial locality restrictions.

1. INTRODUCTION

In image analysis applications, two are the usual embeddings used for the representation of visual data. One is in the feature space, mapping the image data in the Euclidean space. If intensity or color is the only feature used, this space is 1-D for intensity images or 3-D in the case of color images. The second one is by means of a planar graph structure a quite popular approach lately.

It is accepted that the image segmentation process can be handled as a clustering problem where the image basic elements (pixels or other entities) have to be partitioned in homogeneous groups. Depending on the way image data are utilized segmentation-clustering procedures differ considerably and lead to a categorization of the techniques in two basic classes. Algorithms of the first class operate in a vectorial space considering image elements as independent vectors. Often work on a pixel basis and by means of a central clustering algorithm data are grouped in the image feature space. This approach is efficient in capturing global characteristics but is not good at preserving image details since no spatial structure or connectivity information is utilized. A very serious drawback is that pixels from disconnected parts of the image can be grouped together.

Given the importance of edge information as well as the need to preserve pixels spatial relationship on the plane, there is recently a tendency of handling images in the spatial domain. This second class of algorithms, operate on the image spatial domain and map image data into a planar graph structure. Vertices or nodes of the graph represent image pixels (or small regions), while edges denote the connectivity between them. Associated with each edge there is a weight indicating the similarity or dissimilarity of the two pixels (or regions). It is a highly complex structure that needs to be partitioned into subgraphs to produce the final segmentation result. Thus the image segmentation problem is now considered as a graph-clustering (or graph-partitioning) problem [1-3]. Two of the graph partitioning methods that attracted most attention for application in image segmentation problems, are the minimal spanning tree and the n-cut method [2,3]. These are usually referred as region-based techniques and make use of pairwise data relations to perform grouping.

The minimal spanning tree (MST) preserves the connectivity of the image graph and provides a link to all nodes (regions) at a minimum total edge cost. By deleting MST edges with the largest values, isolated clusters that correspond to segmented regions, are formed. Recursive MST versions, incorporating some form of global information by recalculating links and costs, have also been proposed but at a much increased computational complexity.

In this work we attempt a different approach to make use of the important information existing in the MST. More specifically, its ability to preserve connectivity among all image elements and at the same time offer a hierarchical representation of data is used to provide a new locality preserving feature information. The new feature used is the density value. It is estimated via nonparametric techniques using pairwise distances defined on the MST. In this way information concerning image connectivity and spatial contiguity is integrated in the new feature.

2. GRAPH IMAGE SEGMENTATION

Prior to mapping the color image into a graph structure, it is first transformed to a more efficient representation. The watershed algorithm is applied to the magnitude of the gradient image, resulting to an oversegmented partition. Original image and the watershed oversegmentation are displayed in Figure 1(a,b). Many graph based image segmentation methods work directly on the pixel level. We opted to work with already formed small regions since this approach does not require any closing or connection of the edges and at the same time preserves all the important edge information. Additionally computational complexity in the remaining part of the algorithm is greatly reduced. An averaged size image has approximate 10^5 pixels and 3*10^3 regions although this depends heavily on color uniformity and texture.

The image can be represented as a planar weighted undirected graph $G(V,E)$ where V are the nodes of the graph and an edge E is formed between every node and it's connected neighbours. Each node V_i represents a small quasi-homogeneous watershed defined region. These regions are undersized to capture texture, shape or other characteristics that require a certain spatial extend. As a result, the mean color vector can be considered as sufficient feature descrip-

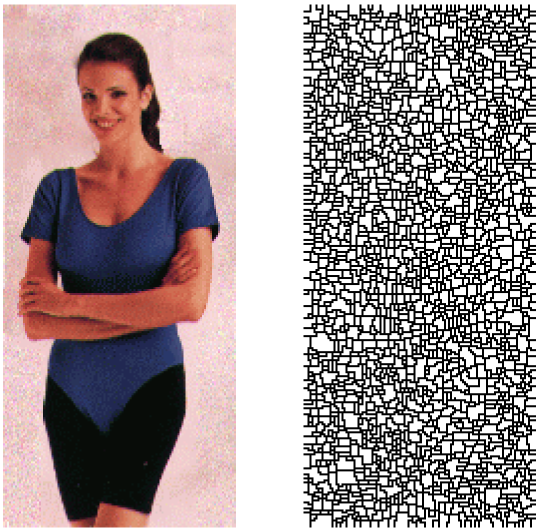

Figure 1. (a) Original image "woman", 261x116 pixels (b) watershed segmentation of the gradient image (~1800 regions)

tor. Using this transformation, each node V_i of the graph, also represents a vector X_i in the RGB feature space. Associated with each edge $E_{i,j}$, there is a positive weight $w_{i,j}$, indicative of the visual dissimilarity of the two nodes i, j. Its value is defined here using the Euclidean distance $d_{i,j}$ between the corresponding vectors in RGB feature space as the L_2 norm

$$w_{i,j} = d_{i,j} = \left\| X_i - X_j \right\| \quad (1)$$

This type of graph image representation is often called region adjacency graph (RAG) [4]. It should be noticed that although edge weights are defined in Euclidean space the graph is planar.

2.1 MST based segmentation

A simplification of the RAG structure by means of a spanning tree that runs over all nodes, at a minimum total length cost, results in the MST graph displayed in Figure2. The MST is constructed using well known greedy algorithms and the upper bound of its computational complexity is $O(N^2 \log N)$, N the number of graph nodes. Although among graph based segmentation algorithms, the computational efficiency of the MST is recognized, there are several problems associated with it. These are related to the fact that the MST does no include global information in the merging operations, which are decided locally by weight values. Nevertheless results are in some cases quite satisfactory. A situation where MST fails is when a series of regions, which differ slightly from each other, bridge a large gap in the feature space. Since the edge weights are small, the graph will not be cut and very different objects in terms of feature space characteristics will be merged. Recursive implementations of the MST algorithm, recalculating edge costs, can tackle the problem increasing complexity [2].

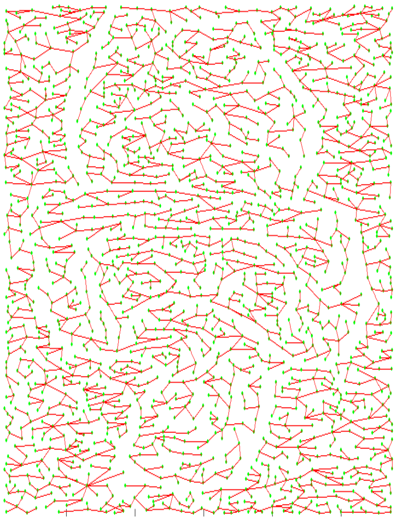

Figure 2. Planar MST resulting from the region adjacency graph of Fig.1(a,b).

2.2 Feature space clustering

In pattern recognition applications, the basic algorithmic tool for manipulating data in the feature space is clustering. This is in done using distances, either explicitly, between vectors in a coordinate system or using pairwise data dissimilarities. Image data are strongly correlated and form clusters with complex shape and structure that makes clustering and subsequent segmentation a difficult task.

A serious problem related to cluster based image segmentation is that classical algorithms neglect the information contained in the pixels' spatial coordinates. Regions that are located at close distance, have a higher probability to be located close in feature space. The inclusion of the spatial coordinates as additional features, has not proved as a good solution to the problem since it mixes heterogeneous attributes and is difficult to apply the proper scaling.

The problem of spatial proximity and region homogeneity is quite complicated. For example there are image features like the background pixels, which usually belong to the same class and which can be situated far apart on the image plane. On the contrary, nearby pixels that are separated by a large color or intensity difference, belong to different classes. Thus for segmentation purposes, x-y coordinate distance is not a good criterion to be associated with region homogeneity. A better measure to describe spatial coherence is by means of a geodesic distance on the image relief since; it describes distance spanned in the feature space.

3. GEODESIC DISTANCE AND MST

Given the region adjacency graph of the watershed segmented image, the geodesic distance between regions p and q, is defined as the shortest path from p to q. Suppose $P = \{p_1, p_2 ...., p_n\}$, is a path between p_1 and p_n, where (p_i and p_{i+1} are connected neighbours).

The path length $l(P)$ is defined as

$$l(P) = \sum_{i=1}^{n-1} d(p_i, p_{i+1}) \qquad (2)$$

that is the sum of the neighbour distances d between adjacent points in the path.

Having constructed the MST of the graph image $G(V,E)$, the previously described geodesic distance between any of its nodes V_i, V_j, can be computed, in a unique way. The new distance connecting every two nodes of the MST is defined as

$$D_{(i,j)MST} = \sum_i^j w_{t,t+1} \qquad (3)$$

where, $w_{t,t+1}$, are the weights linking adjacent nodes V_t, V_{t+1} on the MST. According to equation (1), each weight is equal to the Euclidean distance $d_{t,t+1}$, between the corresponding feature vectors X_t, X_{t+1}.

The distance defined in this way through the MST can be considered as an approximation of the shortest path distance and its value is related to the feature space coherence of the two regions. A very similar idea, defined according to fuzzy logic principles is the notion of fuzzy connectedness.

3.1 MST based local density estimation

It is known that given a collection of sample vectors X_i, i.e. corresponding to a watershed image regions, using nonparametric techniques we can estimate the density function at any point X of the color space

$$\hat{f}(X) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{X - X_i}{h}\right) \qquad (4)$$

The kernel $K(X)$ is taken to be a radially symmetric, non-negative function centered to zero and integrating to one. A common selection is the Gaussian kernel and in this work the space dimensionality is $d = 3$. The window radius h (bandwidth) is very critical for the smoothness of the density estimate. By locating local maxima of $\hat{f}$ we can perform clustering and therefore image segmentation. As mentioned in 2.2, the density estimate computed in this way does not consider any spatial contiguity between pixels.

In this work, the previously defined geodesic distance based on the MST,

$$D_{(i,j)MST} = \sum_i^j d_{t,t+1} \qquad (5)$$

substitutes in the kernel of equation (4), the $\|X - X_i\|$ norm, and is used to obtain a new density estimate at every graph node. Although the n value in the summation of (4), in principle runs over all image regions, in practical terms only nodes that are not separated by a large color or intensity discontinuity will contribute to the density average. Thus the new density estimate takes into consideration locality criteria defined by the structure and length of MST edges.

In order to visualize results, using equations (3) and (4), the density value estimated at every image graph node is displayed in the vertical axis of the plot of Figure 3. It is observed that the density modes are peaked in the spatially uniform parts of the image. The variability in values is due to the 'random walk' type of MST distance, as it runs over uniform areas. When the application of interest is image segmentation a variation of the mean shift algorithm [5], is appropriate to determine modes in this peaked type of density structure. To reduce density variability in image 'flat areas', a more appropriate distance measure to use is

$$D_{(i,j)MST} = \max\{d_{t,t+1}\} \qquad (6)$$

where, the max of the individual distances $d_{t,t+1}$, forming the path between nodes V_i, V_j, is selected. In this case there is a flattening of the density estimate. An additional useful constraint imposed to avoid the well-known MST limitation, namely the 'chaining effect', is to consider also the distance in the feature space between nodes i, j and use the maximum of the two

$$D_{(i,j)} = \max\{D_{(i,j)MST}, d_{(i,j)}\} \qquad (7)$$

The estimated density in this case is displayed in Figure 4. A filtering of local variations takes place and the density levels corresponding to separate objects or regions of the image are quite distinct.

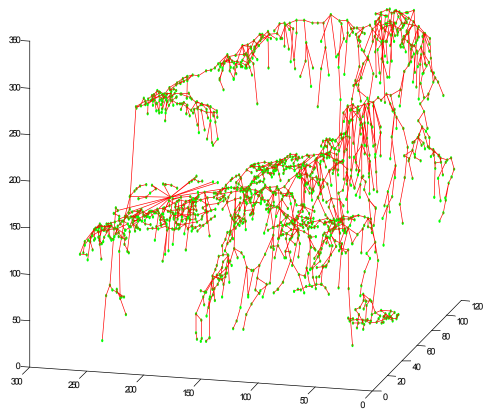

Figure 3. Estimated density (vertical axis, arbitrary units), at each node of the MST, using equations (4,5). The horizontal plane corresponds to the image spatial coordinates.

4. SEGMENTATION RESULTS

The computed density can be either used as a stand-alone feature for segmentation or in conjunction with color. As an example, segmentation results using matlab's kmeans algorithm and 4 clusters are displayed in Figure 5 when, (a) RGB components are used, (b) data are clustered in a 4-D space, by using the additional local density feature. An alternative way of segmenting the image is by using, a combination of density and color uniformity restrictions, on the MST structure. Results are displayed in Figure 6(a,b). In all cases, kernel bandwidth is critical for smoothing noise variations. This has to be exercised with caution to avoid smearing of useful signal features. It was observed that, as expected, prior image filtering improved results, increasing class separability. The above-described technique is not applicable to heavily textured images. A possible solution in this case would be to use a texture dimensionality reduction technique and subsequently construct the MST in the new space. Similar results were obtained with other, more complex types of images.

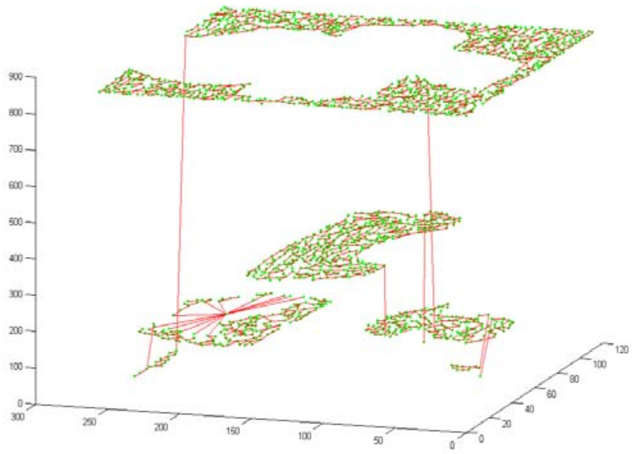

Figure 4. Estimated density, using equations (4,7).

5. CONCLUSIONS

Density estimation is useful and important in statistical approaches to various problems of image processing, pattern recognition and artificial intelligence. In this work a novel way of its estimation is proposed by means of the minimal spanning tree and applied for image segmentation. It is in general helpful in detecting clusters that are internally homogeneous and constrained by a spatial neighbourhood structure. Besides image segmentation, this novel spatially local density estimate can find other applications as i.e. in density-based sampling for data mining.

6. ACKNOWLEDGEMENTS

This work was funded by the National Ministry of Education in the context of program "Archimidis".

REFERENCES

[1] C. T. Zahn. "Graph-theoretical methods for detecting and describing gestalt clusters", *IEEE Trans. on Computers*, C-20: pp. 68-86, Jan. 1971.

[2] O. J. Morris, M. de J. Lee and A.G. Constantinides, "Graph Theory for Image Analysis: An Approach Based on the Shortest Spanning Tree", *IEE Proceedings*, Part F., vol. 133, No. 2, pp. 146-152, 1986.

[3] J. Shi and J. Malik, "Normalized cuts and image segmentation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22(8), pp. 888–905, 2000.

[4] K. Haris, S. N. Efstratiadis, N. Maglaveras, and A.K. Katsaggelos, "Hybrid image segmentation using watershed and fast region merging," *IEEE Trans. Img. Proc.*, vol. 7, no. 12, Dec. 1998.

[5] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24(5), pp. 603–619, 2002.

Figure 5. (a) Segmentation results in space using kmeans and 4 clusters, (b) same results using an additional local density feature defined using the MST.

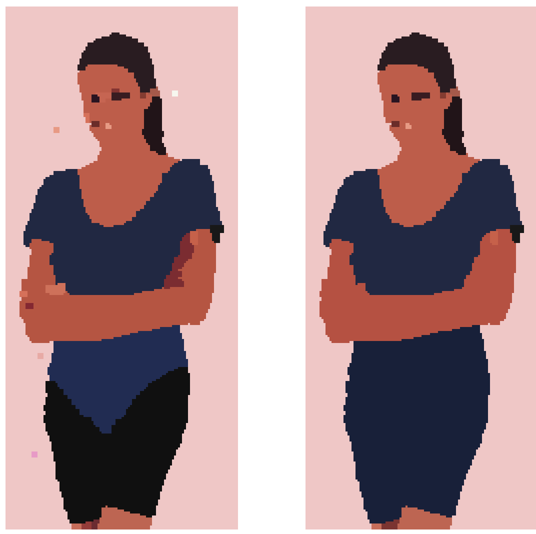

Figure 6. MST segmentation results using density and color distances, for different MST-edge and kernel bandwidth parameters.