IMAGE SEGMENTATION ALGORITHMS: GRAPH AND MEAN SHIFT

Ptandon1

Reading assignment 5

SUMMARY

The first paper devises a new mean shift based procedure for analysis of feature space by modeling it as a pdf and using the pattern recognition techniques like Kernel density estimation to find the peaks of the feature space with the aim to identify the modes to generate a successful and effective segmentation of the images at hand. The second paper on the other hand takes the approach of graphs and model an input image as a graph and applying undirected graph algorithms and properties such as Minimum Spanning Tree and min cut, graph connectivity properties. Both the papers first discuss the methods used for segmentation to date (of publication) and how the approach they mentioned fits into the problem followed by problem formulation, algorithm description and testing on real world example images for performance evaluation.

INTRODUCTION/BACKGROUND

The mean shift procedure is backed by the idea of a feature space. The feature space is mapping of the input based on time based subset processing. Significant objects or more subjectively 'regions' correspond to dense parametrized regions on the feature space such that given the parameters, the subset of the image can be generated using the mapping. For the mean shift procedure to work, we have the task of clustering based on density estimation. This is based on the fact that the feature space can be considered as a regular probability density function such that the dense regions in the features (the significant features in the original input image) correspond to local maxima in the p.d.f. or more clearly, the modes of the distribution. The identification of the location of a mode initiates the local search around the area for analysis of the local structure of feature space.

Prior to the usage of graph methods for segmentation, there were a whole range of methods pertaining to the concept of eigen matrices and eigen vectors which very slow and not applicable to the problems of scalability to large scale databases. Methods other than based on eigen vectors also existed which were faster but they failed to capture important non local properties of the image under consideration. The technique shown here not only captures those non local characteristics but also runs in O(n log n) time where n representing the number of pixels.

DETAILED SUMMARY: Mean Shift Procedure

This detection of the modes and clustering are the main objectives of the mean shift procedure. The kernel density estimator given n data points is given by

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) \text{ where } K_H(x) = |H|^{-\frac{1}{2}}K(H^{-\frac{1}{2}}x) \text{ and } H \textit{ is the bandwidth matrix}.$$

The mean shift method lays its focus on kernels that are radially symmetric. The choice for the bandwidth matrix H is made with the aim to reduce the number of required parameters and is thus chosen as $H = h^2 I$ where I is the Identity matrix. This leads to the KDE transforming into: $\hat{f}(x) = \frac{1}{nh^d}\sum_{i=1}^{n} K_H\left(\frac{x-x_i}{h}\right)$. To judge the performance, an asymptotic approximation of the mean squared error between the density and the estimator integrated over the domain is computed which is minimized by the Epanechnikov kernel yielding a radially symmetric profile. Due to nondifferentiability of the kernel at the boundary. Thus a multivariate normal kernel is generated using the profile $k_N(x) = \exp\left(-\frac{1}{2}\right)$ to finally incorporate in the KDE as: $\widehat{f_{h,K}}(x) = \frac{c_{k,d}}{nh^d}\sum_{i=1}^{n} k\left(\left\|\frac{x-x_i}{h}\right\|\right)$. Having the density estimator, the density gradient estimator is calculated by calculating the gradient of the estimator. This yields the expression

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}}[\sum_{i=1}^{n} g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)]\frac{\sum_{i=1}^{n} x_i\, g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}$$

where the first term $\frac{2c_{k,d}}{nh^{d+2}}[\sum_{i=1}^{n} g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)]$ represents proportionality to the KDE whereas the second term $\frac{\sum_{i=1}^{n} x_i\, g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}$ represents the mean shift. Rearranging we get: $m_{(h,G)}(x) = \frac{1}{2}h^2 c\frac{\nabla f_{h,K}(x)}{f_{h,G}(x)}$. This reveals that the mean shift always is directed towards the direction of increase of the density.

With the information available about the increasing direction of density, the modes can be obtained by checking for stationary points (while pruning the points that are not local maxima) and repeated computation of mean shift vector and the KDE. The paper gives full-fledged proofs about the convergence of the algorithm and applicability of Kernel Regression and M-estimators to mean shift procedure is discussed. The applicability of the procedure to segmentation is very much similar to the main procedure. We run the mean shift algorithm to get information about the d-dimensional convergence point For all the points that are closer than h_S in the spatial domain and h_R in the range domain, cluster them linking the corresponding convergence points and then followed by assigning $L_i = \{p|z_i \in C_p\}$ where L_i *is the label of th ith pixel*.

The performance of the procedure is demonstrated at different parameter levels of h_S *and* h_R and it proves that the mean shift procedure is not computationally intensive and performs very well with respect to segmentation. The only limiting point of this algorithm is the KDE which does not perform well at very high dimension levels.

Graph based Segmentation

The method of graph based segmentation is very much similar to the process of clustering, in terms of the fact that both are based on selecting edges from graph. Each pixel of the image corresponds to a node in the graph. Nodes are connected using undirected edges since modelling is done only for the relation of pixels and not the directionality of relations. The graph is weighted with the weights representing the level of dissimilarities between pixels which is calculated based on some segmentation criterion. This criterion can be changed as per the variability of neighboring area in the image. The proposed algorithm makes greedy decisions in the process of segmenting the image (modelled as graph) but is still able to capture the non-local characteristics of the image.

The main idea behind the calculations backend of the algorithm are the concepts of intensity difference and internal differences between neighboring pixels. The goal of performing segmentation is to achieve a state such that elements in a component of the image are identified as similar and elements from different components are identified as dissimilar.

The paper defines a term called as Pairwise Region Comparison Predicate is defined as a quantity representing inter-component differences compared to intra-component differences. The internal difference is set to be corresponding to the maximum weight in the MST of the component whereas the inter component difference is set as the weight of the least cost edge connecting the two components. Following the definitions, the criterion is defined that a boundary exists if the inter-component difference is significantly greater than internal difference for any component. This quantity is defined as $MInt(C_1, C_2) = \min\big(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)\big)$ where τ is a threshold function controlling the amount by which the difference between components must be greater to judge the existence of a boundary. The algorithm is as follows:

Given an input graph G = (V,E) with n vertices and m edges, sort the edges in the increasing order of their weights. Start with each vertex in its own component and start merging the segments into a larger segment if their internal difference $\geq$ inter component difference. Repeatedly cluster them as such till all the vertices have been visited and terminate the algorithm when it happens return the set of clusters so formed.

The paper proves that the segmentation obtained is neither too coarse nor too fine, which the main aim of any segmentation algorithm. The method proposed is proven to be very fast and its applicability to large scale image databases is also shown.

DISCUSSIONS

Both methods, the mean shift algorithm as well as the graph based segmentation offer promising ways to effectively segment the image without loss of detail and most importantly not very computationally intensive. Not only segmentation, these algorithms have a much wider range of applicability in the field of computer vision and feature space analysis including cases like discontinuity preserving smoothing without blurring the edges.

REFERENCES

[1] D. Comaniciu & P. Meer, Mean shift: a robust approach toward feature space analysis, *IEEE. Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 603-619, 2002.

[2] P. F. Felzenszwalb & D. P. Huttenlocher, Efficient graph-based image segmentation, *International Journal of Computer Vision,* vol. 59, pp. 167-181, 2004.