Fuzzy C- Means Clustering with Kernel Metric and Local Information for Image Segmentation

Pallavi Khare
Department of Electrical & Electronic Engineering
Padmashree Dr. D. Y. Patil Institute of Engineering & Technology,
Pimpri, Pune-18

Anagha Gaikwad
Department of Electrical & Electronic Engineering
Padmashree Dr. D. Y. Patil Institute of Engineering & Technology,
Pimpri, Pune-18

Pooja Kumari
Department of Electrical & Electronic Engineering
Padmashree Dr. D. Y. Patil Institute of Engineering & Technology,
Pimpri, Pune-18

ABSTRACT
Image segmentation has been an intriguing area for research and developing efficient algorithms, playing a paramount role in high caliber image interpretation and image analysis. Segmentation of images plays an imperative role in medical diagnosis. Such segmentation demands a robust segmentation algorithm against noise. The legendary orthodox fuzzy c-means algorithm is proficiently exploited for clustering in medical image segmentation. FCM is highly sensitive to noise due to the practice of only intensity values for clustering. Thus this paper aims to apply the 'kernel method', instituted on the conventional fuzzy clustering algorithm (FCM) to swap the Euclidean metric norm to a novel kernel-induced metric in the data space. Images can be segmented by pixel classification through clustering of all features of interest. In unsupervised methods of clustering algorithms utilizing kernel method, a nonlinear mapping is operated initially in order to map the data into a much higher space feature, and then clustering is executed. The integer of clusters in the multidimensional feature space thus represents the number of classes in the image. As the image is sorted into cluster classes, segmented regions are obtained by examination of the neighborhood pixels for the same class label. Since clustering produces disjointed regions with holes or regions with a single pixel, a post processing algorithm such as region growing, pixel connectivity, or a rule-based algorithm is applied to obtain the final segmented regions.

Keywords
Segmentation, clustering, kernel, nonlinear, FCM

1. INTRODUCTION
Imaging science has long-drawn-out primarily along three distinct but related lines of research: segmentation, registration and visualization. Registration involves finding the transformation that brings different images of the same object into strict spatial (and/or temporal) congruence. And visualization involves the display, manipulation, and measurement of image data. Finally, segmentation is defined as the process of partitioning an image into a set non-overlapping regions whose union is the entire image where these regions should ideally correspond to objects and their meaningful parts, and background. Most image segmentation algorithms are based on two basic properties that can be extracted from pixel values-discontinuity and similarity-or a combination of them. Segmentation of nontrivial images is a grim problem-made even rigid by non- uniform lighting, shadows, overlapping objects, poor contrast between objects and background, and so on with some degree of success to this date. Image segmentation can be approached as three philosophical perspectives - region, boundary and edge. Image processing techniques for quantitative analysis are primarily used in computational medical analysis. Computer analysis, if performed with the appropriate care and logic, can potentially add objective strength to the interpretation of the expert. Thus, it becomes possible to improve the diagnostic accuracy. It is the important yet elusive capability to accurately recognize and delineate all the individual objects in an image scene.

2. EXISTING METHODS
Rudimentally we can cerebrate of several rudimental concepts for segmentation. Pixel-predicated methods only utilize the gray values of the individual pixels. Region-predicated methods analyze the gray values in more immensely colossal areas. Conclusively, edge-predicated methods detect edges and then endeavor to follow them. The prevalent constraint of all these approaches is that they are predicated only on local information. Even then they utilize this information only partly. Pixel-predicated techniques do not even consider the local neighborhood. Edge-predicated techniques look only for discontinuities, while region-predicated techniques analyze homogeneous regions. In situations where we ken the geometric shape of an object, model-predicated segmentation can be applied.

3. PIXEL-PREDICATED DIRECT CLASSIFICATION
The pixel-predicated direct relegation methods use histogram statistics to define single or multiple thresholds to relegate an image pixel by pixel. The threshold for relegating pixels is achieved from the investigation of the histogram of the image. A humble line of attack is to examine the histogram for bimodal dispersal. If the histogram is bimodal, the threshold can be set to the gray assessment analogous to the inmost point in the histogram valley. If not, the image can be fenced off into two or more constituencies utilizing some heuristics about the assets of the image. The histogram of every partition can then be utilized for decisive thresholds. The image f(x, y) can be segmented into two classes using a gray value threshold T such that

$$G(x, y) = 1 \text{ if } f(x, y) > T \text{ ;else is } 0 \text{ if } f(x, y) \leq T$$

Where G(x, y) is the segmented image with two classes of binary gray values, "1"and "0", and T is the threshold selected at the valley point from the histogram.

4. THRESHOLDING

Thresholding is one of the simplest methods to attain a crusty segmentation a uni-spectral image. Thresholding engenders a binary image in which the pixels belonging to objects have the value 1 whereas the pixels belonging to the background have the value 0. Images are normally acquired as gray-scale images. Idyllically, objects in the image should appear steadily brighter (or darker) than the background. Using Xi for location (i, j), the thresholded image is given by

$$IT(X_i) = \{1;\ I(X_i) \geq \tau \quad 0;\ I(X_i) < \tau\}$$

Where τ is the threshold value

Detriments

- Little tolerance to intensity rescaling.
- Difficult to set threshold.
- Slight use of spatial information

5. REG I O N: BASED SEGMENTATION

Region-growing based segmentation algorithms inspect pixels in the vicinity grounded on a predefined resemblance criterion. The neighborhood pixels with similar properties are merged to form closed regions for segmentation. The region-growing approach can be extended to merging regions instead of merging pixels to form larger meaningful regions with similar properties. Such a region-merging approach is quite effective when the original image is segmented into a large number of regions in the preprocessing phase. Large meaningful regions may provide better correspondence and matching to the object models for recognition and interpretation.

Difficulties

- Low restraint to intensity rescaling.
- Challenging to set mounting criteria and preventing criteria.
- Needs human intervention for defining seed point

6. EDGE-BASED IMAGE SEGMENTATION

Edge-based methodologies use a spatial filtering method like the Laplacian mask to work out the first-order or second-order gradient statistics of the image. The segmentation of an image into separate objects can be achieved by finding the edges of those objects. This method involves computation of an edge image, containing all (plausible) edges of an original image, then processing the edge image so that only closed object boundaries remain, and finally transforming the result to an ordinary segmented image by filling in the object boundaries

Disadvantages

- Transforming an edge image to closed boundaries often requires the removal of edges that are caused by noise or other artifacts,
- Intelligent decisions should be made to connect the edge parts that make up a single object where detection of edges remains ambiguous

7. LEVEL SET METHOD

The level set method was devised by Osher and Sethian to embrace the topology ups and downs of curves. The level set method has been very prosperous in computer graphics and vision, widely used in medical imaging for segmentation and shape recovery. Based on geometric deformable model, the level set scheme translates the tricky evolution 2-D (3-D) close curve (surface) into the evolution of level set function in the space with sophisticated dimension to obtain the benefit in handling the topology changing of the shape.

Drawbacks

- It is difficult to obtain a perfect result when there is a fuzzy or discrete boundary in the region, and the leaking problem is inevitable.
- Solving the partial differential equation of the level set function requires numerical processing at each point of the image domain which is a time consuming process.
- The iteration time increases greatly for too large or too small contour causing the convergence of evolution curve to the contour of object incorrectly.

8. CLUSTERING

Clustering is a process which partitions a given data set or data points into homogeneous groups predicated on given features such that kindred objects are kept in a group whereas dissimilar objects are in different groups. A common approach to image clustering involves addressing the following issues: Image representation, Organizing data, classification of image to a group. The similarity of feature vectors can be represented by an appropriate distance measure such as Euclidean or Mahalanobis distance. Each cluster is represented by its mean (centroid) and variance (spread) associated with the distribution of the corresponding feature vectors of the data points in the cluster. The materialization of clusters is optimized with reverence to an objective function involving pre-specified distance and similarity measures; along with additional constraints such as smoothness. It is the most paramount unsupervised learning quandary. It deals with finding structure in an accumulation of unlabeled data

CALCULATING DISTANCE BETWEEN CLUSTERS

- Centroid is defined as the distance between the centroids of two clusters, i.e., dis (Ki, Kj) = dis(Ci , Cj)
- Centroid: the "middle" of a cluster

$$\frac{C_m = \sum_{i=1}^{N} (t_{ip})}{N}$$

- Medoid is defined as the distance between the medoids of two clusters, i.e., dis (K i, K j) = dis (M i , Mj) where medoid is defined as one chosen, centrally located object in the cluster.
- Distances are normally used to measure the similarity or dissimilarity between two data objects
- Minkowski distance

The Minkowski metric favors the prime scaled feature, which dominates others. The problem can be addressed by proper normalization or other weighting schemes applied in the feature space where d is the dimensionality of the data.

d (i ,j)=q√(|xi1-xj1|q+|xi2-xj2|q+…)

Where i = (xi1, xi2, …, xip) and j = (xj1, xj2, …, xjp) are two p-dimensional data objects, and q is a positive integer.

- If q = 1, d is Manhattan distance

d(i ,j)=|xi1-xj1|+|xi2-xj2|+…+|xip-xjp|

If q = 2, d is Euclidean distance which is defined as the distance between two points as the length of the line segment connecting them. The Euclidean distance has an intuitive appeal as it is commonly used to evaluate the proximity of objects in 2- or 3-D space. It works well when a data set has "compact" or "isolated" clusters. The advantage of Euclidean distance is that it is intuitively obvious. The disadvantages are costly calculation due to the square root, and its Hon-integral value.

Mahalanobis distance

Linear correlation among features can also distort distance measures. This distortion can be alleviated by applying a whitening transformation to the data by using the Mahalanobis distance measure dM(x, y) as

d M (x, y) = ((x - y)A-1 (x – y)T)1 /2 4

Where A is a transposition matrix

Kernel based similarity measure

Mercer Kernel functions map data from input space to high, possibly infinite, dimensional feature space. For a finite sample of data X, the kernel function yields a symmetric N*N positive definite matrix K, where the K (i, j) entry corresponds to the dot product between f (x i) and f (x j) as measured by the kernel function. In feature space, the distance measure between any two patterns is given by:

$$\| f(x_i) - f(x_j) \|^2 = \langle f(x_i), f(x_j)\rangle + \langle f(x_i), f(x_j)\rangle - 2\langle f(x_i), f(x_j)\rangle$$
$$\| f(x_i) - f(x_j) \|^2 = k(i,i)\ k(i,j)\ 2k(i,j)$$

CLUSTERING CLASSIFICATION

Clustering relegation techniques can be grouped into two main types: supervised and unsupervised.

Supervised relegation relies on having example pattern or feature vectors which have already been assigned to a defined class. By contrast, unsupervised relegation does not rely on possession of subsisting examples from a kenned pattern class. The examples are not labeled and we seek to identify groups directly within the overall body of data and features which enables us to distinguish one group from another. Clustering techniques are an example of unsupervised relegation.

8.1 K-means clustering

K-means (MacQueen, 1967) is one of the simplest unsupervised learning algorithms. It outlines a conceptually simple way to partition a data set into a specified number of clusters k. The algorithm aims to iteratively minimize a simple squared error objective function of the form

$$J = \sum_{j=1}^{k} \sum_{\substack{\text{all } i \\ \text{in class } j}} | x_i^j - c_j |^2 ,$$

Where cj denotes the coordinate vector of the jth cluster and {xij} are the points assigned to the jth cluster. Minimizing J equivalently means reaching that configuration at which switching any point to a cluster other than its currently assigned one will only increase the objective function.

Algorithm for K-means clustering

1. Select the number of desired clusters k. Place the k cluster centers at different initial locations in the image.
2. Assign each data point to the cluster whose center is nearby.
3. Re-compute the cluster centers; the cluster center should be at the average coordinates (center of gravity) of the data points that make up the cluster.
4. Go to step 2 until no more changes befall or a determined number of iterations is reached

Advantages

1. Fast, robust and more facile to understand.
2. Gives best result when data set are distinct or well disunited from each other

Shortcomings

1. The learning algorithm requires apriori designation of the number of cluster centers.
2. If there are two highly overlapping data then k-denotes will not be able to resolve that there are two clusters.
3. The learning algorithm is not invariant to non-linear transformations i.e. with different representation of data we get different results (data represented in form of Cartesian co-ordinates and polar co-ordinates will give different results).
4. Euclidean distance measures can unequally weight underlying factors.
5. The learning algorithm provides the local optima of the squared error function.
6. Randomly culling of the cluster center cannot lead us to the fruitful result.
7. Applicable only when mean is defined i.e. fails for categorical data.
8. Unable to handle strepitous data and outliers

8.2 FUZZY C-MEANS CLUSTERING

Fuzzy c-means (FCM) is a scheme of clustering which allows one section of data to belong to dual or supplementary clusters. This method was developed by Dunn in 1973 and enriched by Bezdek in 1981 and it is habitually used in pattern recognition. Main objective of fuzzy c-means algorithm is to minimize:

$$J(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (u_{ij})^m \| x_i - v_j \|^2$$

Where, '||xi – vj||' is the Euclidean distance between ith data and jth cluster center.

Algorithm for Fuzzy c-means clustering

1. Randomly select cluster centers.
2. Calculate the fuzzy membership.
3. Compute the fuzzy centers.
4. Repeat step 2) and 3) until the minimum value of the objective function is achieved

ADVANTAGES

1. FCM gives best result for overlapped data set and is comparatively better then k-means algorithm.
2. Data point is assigned membership to each cluster center as a result of which data point may belong to more than one cluster center

WEAKNESSES

1. Apriori measurement of the number of clusters.
2. With subordinate value of β we get the better result but at the overhead of extra number of iteration.
3. Euclidean distance measures can inequitably weight underlying factors.

9. PROPOSED METHOD

KERNEL FUZZY C MEANS CLUSTERING

The kernel metric Fuzzy C-Means minimizes the following objective function.

$$J\phi = \sum_{i=1}^{k} \sum_{j=1}^{n} u_{ij}^{m} ||\phi(x_j) - \phi(v_i)||^2$$

where, uij denotes the membership of xj in cluster i, ϕ(vi) is the center of cluster i in the feature space, and ϕ is the mapping from the input space X to the feature space F. Minimization of the function has been proposed only in the case of a Gaussian kernel.

KFCM Algorithm

1. Select initial class prototype $\{V_i\}_{i=1}^{c}$
2. Update all memberships Uij
3. Obtain the prototype of clusters in the forms of weightedaverage Repeat step 2-3 till termination. The termination criterion is ‖Vnew - Vold‖≤ε

Where ‖.‖ is the Euclidean norm. V is the vector of cluster centers ε is a small number that can be set by user (here ε =0.01).

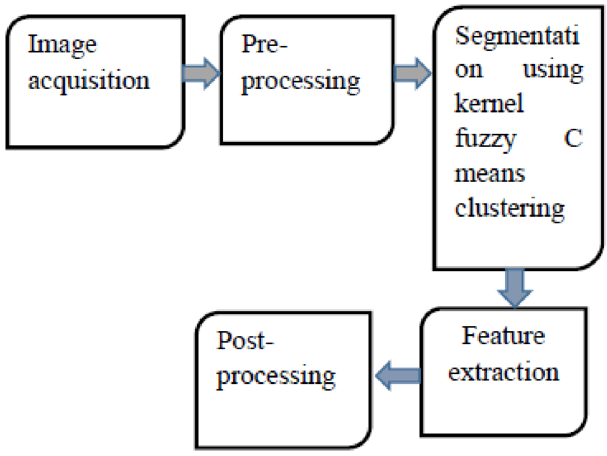
Kernel-based clustering algorithms have the following main advantages.

1. We can obtain a linearly separable hyper-plane in the high-dimensional, or even in an infinite feature space.
2. They can identify clusters with arbitrary shapes.
3. Kernel-based clustering algorithms, have the capability of dealing with noise and outliers.
4. There is no requirement for prior knowledge to determine the system topological structure.
5. The kernel matrix can provide the means to estimate the number of clusters.

10. DOWNSIDES OF KFCM

A precarious issue related to KFCM clustering is the selection of an "optimal" kernel for the problem at hand and on the setting of the involved parameters. The kernel function in use must conform to the learning objectives in order to obtain meaningful results for un-labeled data

BLOCK DIAGRAM OF PROPOSED METHOD



The system consists of the following blocks. The raw data is passed through the system as numerical data or in the form of waves. Applicable techniques are applied to get the preprocessed data. Further, the data is passed through the clustering phase, which returns the cluster centers. Feature extraction is then performed to obtain the attributes that can downright exemplify a given instance. Next a post processing is used to enhance the quality of the final segmented image

Pre-processing

The pre-processing stage is performed to convert all attributes of the data into a numeric form that can be used by the clustering process. This is extremely useful for reduction in dimension of the dataset using normalization. If the values of some attributes vary in different ranges then to reduce the effect of such attributes, all values of the attributes are normalized to lie in some common range, like [0, 1].Pre-processing enhances the visual appearance of images and manipulates datasets.

Clustering

The clustering is an important step, as it is an essential precursor to the feature extraction. The input for feature extraction is the pre-processed data, where the labels are stripped off. Clustering is a form of unsupervised learning that helps to find the inherent structure in the data.

Feature extraction

It is the process by which certain features of interest within an image are detected and represented for further processing. It marks the transition from pictorial to non-pictorial (alphanumerical, usually quantitative) data representation which can be subsequently used as an input to a number of pattern recognition and classification techniques, which will then label, classify, or recognize the semantic contents of the image or its objects.

Post-processing

Image post processing enhance the quality of the finished image, by filtration and other treatments. Here algorithm such as region growing, pixel connectivity or a rule-based algorithm is applied to obtain the final segmented regions

11. Applications

1. Quantitative or semi-quantitative diagnostic image analysis.

2. Surgical planning.
3. Computer assisted surgery

12. REFERENCES

[1] Fast and Robust Fuzzy C-Means Clustering Algorithms Incorporating Local Information for Image Segmentation by Weiling Cai, Songcan Chen and Daoqiang Zhang.

[2] S. Krinidis and V. Chatzis, “A robust fuzzy local information C-means clustering algorithm,” IEEE Trans. Image Process., vol. 19, no. 5, pp. 1328–1337, May 2010.

[3] P.Sivasangareswari, K.Sathish Kumar, “Fuzzy C-Means Clustering With Local Information and Kernel Metric For Image Segmentation”, International Journal of Advanced Research in Computer Science & Technology, ISSN - 2347 –8446, 2014.

[4] Ajala Funmilola A, Oke O.A, Adedeji T.O, Alade O.M, Adewusi E.A, “Fuzzy k-c means Clustering Algorithm for Medical Image Segmentation”, Journal of Information Engineering and Applications, Vol. 2, ISSN - 2225-0506, 2012.

[5] J.C.Bezdek ,Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press ,New York,1981

[6] S. Chen and D. Zhang, “Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure,” IEEE Trans. Syst., Man, Cybern., B, Cybern., vol. 34, no. 4, pp. 1907–1916, Aug. 2004.

[7] Y. Tolias and S. Panas, “Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions,” IEEE Transactions on Systems, Man and Cybernetics, vol. 28, no. 3, pp. 359–369, March 1998.

[8] L.Szilagyi ,Z. Benyo, S. Szilagyii, and H. Adam, “MR brain image segmentation using an enhanced fuzzy C-means algorithm,” in proce.25th Annu.I nt. Conf. IEEE EMBS, Nov. 2003, pp. 17-21.

[9] F.Masulli and A.Schenone, “A fuzzy clustering based segmentation system as support to diagnosis in medical imaging,” med, vol. 16, no 2, pp. 129-147, 1999.Fu, S.K.—Mui, J.K.: A Survey on Image Segmentation. Pattern Recognition, Vol. 13, 1981, pp. 3–16.

[10] Hung M, D. ang D, 2001 "An efficient fuzzy c-means clustering algorithm". In Proc. the 2001 IEEE International Conference on Data Mining.

[11] “Medical image analysis” , Atam P. Dhawan, Second edition, Published by John Wiley & Sons, Inc., Hoboken, New Jersey