International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249–8958, Volume-1 Issue-3, February 2012

Various Image Segmentation Techniques through clustering and Markovian Model: A Survey

Ram Kishan Dewangan, Tripti Sharma

Abstract*: *Image segmentation is the identification and separation of homogeneous regions in the image, has been the subject of considerable research activity. Many algorithms have been elaborated for gray scale images. This paper is a survey on different clustering techniques to achieve image segmentation. Clustering can be termed here as a grouping of similar images in the database. Clustering is done based on different attributes of an image such as size, color, texture etc. The purpose of clustering is to get meaningful result, effective storage and fast retrieval in various areas.

Keywords*: *Clustering, image segmentation, markovian model, relevance feedback.

I. INTRODUCTION

A. Image Segmentation:

Image segmentation is the separation of the image into regions and is one of the first steps leading to image analysis and interpretation[5]. It is used in many practical applications in machine vision, biometric measurements, medical imaging etc. for the purpose of detecting, recognition or tracking of an object. Image segmentation approaches can be classified according to features or type of the technique used.

Segmentation is the process of partitioning a digital image into multiple segments based on pixels[1],[3]. It is a critical and essential component of image analysis system. The main process is to represent the image in a clear way. The result of image segmentation is a collection of segments which combine to form the entire image. Real world image segmentation problems actually have multiple objectives such as minimize overall deviation, maximize connectivity, minimize the features or minimize the error rate of the classifier[9].

B. Clustering

Clustering in image segmentation is defined as the process of identifying groups of similar image primitive. Clustering techniques can be classified into supervised clustering and unsupervised clustering[7],[12].

Supervised clustering demands human interaction to decide the clustering criteria Unsupervised clustering decides the clustering criteria by itself. Supervised clustering includes hierarchical approaches such as relevance feedback techniques and unsupervised clustering includes density based clustering methods. These clustering techniques are done to perform image segmentation.

Clustering is a process of organizing the objects into groups based on its attributes. A cluster is therefore a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters. An image can be grouped based on keyword or its content[12].

In keyword based clustering, a keyword is a form of font which describes about the image keyword of an image refers to its different features. The similar featured images are grouped to form a cluster by assigning value to each feature.

In content based clustering, content refers to shapes, textures or any other information that can be inherited from the image itself. The tools, techniques and algorithms that are used originate from fields such as statistics, pattern recognition, signal processing etc. Clustering based on the optimization of an overall measure is a fundamental approach explored since the early days of pattern recognition. The most popular method for pattern recognition is K-means clustering.

In K-means clustering a centroid vector is computed for every cluster. The centroid must be chosen such that it should minimize the total distance within the clusters.

Both supervised and unsupervised clustering techniques are used in image segmentation[4],[6]. In supervised clustering method, grouping is done according to user feedback. In unsupervised clustering, the images with high features similarities to the query may be very different in terms of semantics. This is known as semantic gap. To overcome this novel image retrieval scheme called as cluster based retrieval of images by unsupervised learning can be used. This works based on a hypothesis: semantically similar images tend to be clustered in some feature space.

A variety of clustering techniques[2], have been introduced to make the segmentation more effective.

Manuscript published on 28 February 2012.
***** Correspondence Author (s)

1Ram Kishan Dewangan*, Computer Science & Engineering, Chhattishgarh Swami Vivekanand Technical University/ Chhattrapati Shivaji Institute of Technology, Durg, India, 9301209363, (e-mail: rkdewangan@csitdurg.in).

2Ms. Tripti Sharma*, Computer Science & Engineering, Chhattishgarh Swami Vivekanand Technical University/Chhatrapati Shivaji Institute of Technology, Durg, India, Mobile No.8966000053, (e-mail: triptisharma@csitdurg.in).

The clustering techniques which are included in this paper are relevance feedback[8], log based clustering, hierarchical clustering, graph based, retrieval-dictionary based, filter based clustering etc.

II. CLUSTERING TECHNIQUES

An image may contain more than one object and to segment the image in line with object features to extract meaningful object has become a challenge to the researches in the field. Segmentation can be achieved through clustering[11]. This paper critically reviews and summarizes different clustering techniques.

A. Relevance feedback:

A relevance feedback[8] approach allows a user to interact with the retrieval algorithm by providing the information of which images user thinks are relevant to the query. Keyword based image retrieval is performed by matching keyword according to user input and the images in the database.

Some images may not have appropriate keywords to describe them and therefore the image search will become complex. One of the solution in order to overcome this problem is "relevance feedback" technique that utilize user feedback and hence reduces possible errors and redundancy.
This technique uses a Bayesian classifier which deals with positive and negative feedback. Content based clustering methods cannot adapt to user changes, addition of new topics due to its static nature. To improve the performance of information retrieval log-based clustering approaches are brought into the application.

B. Log –Based Clustering:

Images can be clustered based on the retrieval system logs maintained by an information retrieval process. The session keys are created and accessed for retrieval[15]. Through this the session clusters are created. Each session cluster generates log –based document and similarity of image couple is retrieved. Log –based vector is created for each session vector based on the log-based documents. Now, the session cluster is replaced with this vector. The uncased documents create its own vector.

A hybrid matrix is generated with at least one individual document vector and one log-based clustered vector. At last the hybrid matrix is clustered. This technique is difficult to perform in the case of multidimensional images. To overcome this hierarchical clustering is adopted.

C. Hierarchical Clustering:

One of the well- known technologies in information retrieval is hierarchical clustering. It is the process of integrating different images and building them as a cluster in the form of a tree and then developing step by step in order to form a small cluster[16].
The steps involved in this process are as follows: the images from various databases are divided into X-sorts. The classification will be calculated by modifying the cluster centers, sorts of the images and stored in the form of matrix m*m continuously which also includes dissimilarity values. At first it calculates the similarities between the queried image and the retrieved image in the image database.

D. Retrieval Dictionary Based Clustering

A rough classification retrieval system is formed. This is formed by calculating the distance between two learned patterns and these learned patterns are classified into different clusters followed by a retrieval stage. The main drawback addressed in this system is the determination of the distance[14].
To overcome this problem a retrieval system is developed by retrieval dictionary based clustering. This method has a retrieval dictionary generation unit that classifies learned patterns into plural clusters and creates a retrieval dictionary using the clusters. Here, the image is retrieved based on the distance between two spheres with different radii. Each radius is a similarity measure between central cluster and an input image. An image which is similar to the query image will be retrieved using retrieval dictionary[6].

E. K-Means Algorithm

In K-means algorithm[10],[13],[14] data vectors are grouped into predefined number of clusters. At the beginning the centroids of the predefined clusters are initialized randomly. The dimensions of the centroids are same as the dimension of the data vectors. Each pixel is assigned to the cluster based on the closeness, which is determined by the Euclidian distance measure. After all the pixels are clustered, the mean of each cluster is recalculated. This process is repeated until no significant changes result for each cluster mean or for some fixed number of iterations.

The K-Means is a nonhierarchical clustering technique that follows a simple procedure to classify a given data set through a certain number of K clusters that are known a priori. The K-Means
algorithm updates the space partition of the input data iteratively, where the elements of the data are exchanged between clusters based on a predefined metric (typically the Euclidian distance between the cluster centers and the vector under analysis) in order to satisfy the criteria of minimizing the variation within each cluster and maximizing the variation between the resulting K clusters[14]. Steps of the classical K-Means clustering algorithm:

1. Initialization – generate the starting condition by defining the number of clusters and randomly select the initial cluster centers.
2. Generate a new partition by assigning each data point to the nearest cluster center.
3. Recalculate the centers for clusters receiving new data points and for clusters losing data points.
4. Repeat the steps 2 and 3 until a distance convergence criterion is met.

The assignment of the data

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

points may not be unique (a data point can be equally distanced from two or more cluster centers) a case when the

K-Means algorithm doesn't find the optimal solution corresponding to the global objective function J. In addition, it is sensitive to the initialization process that selects the initial cluster centers (usually randomly picked from input data). If the initial cluster centers are initialized on outliers, the algorithm will converge to local minima and this is one of the major drawbacks of this space partitioning technique.
Although this clustering is convergent and its aim is to optimize the partitioning decisions based on a user-defined initial set of clustering that is updated after each iteration, K-Means algorithm produces accurate segmentation results only when applied to images defined by homogenous regions with respect to texture and color since no local constraints are applied to impose spatial continuity. In addition, the initialization of the K-Means algorithm is problematic and usually the initial cluster centers are randomly picked.

F. Markov Random Field-Based Techniques

The true image is assumed to be a realization of a Markov or Gibbs random field with a distribution that captures the spatial context of the scene. Given the prior distribution of the true image and the observed noisy one, the segmentation problem is formulated as an optimization problem. The commonly used estimation principles are maximum a posteriori (MAP) estimation, maximization of the marginal probabilities and maximization of the posterior marginal. However, these methods require fairly accurate knowledge of the prior true image distribution and most of them are quite computationally expensive [17], [18].

III. CONCLUSION

In this paper a comparative study has been performed on the existing methods for the image segmentation. A comprehensive survey highlighting different clustering techniques used for image segmentation has been presented. Clustering concepts and image segmentation concepts have been analyzed. Through clustering algorithms, image segmentation can be done in an effective way. Spectral clustering technique can be used for image clustering because images that cannot be seen can be placed into clusters very easily than other traditional methods. In general, clustering is a hard problem. Clustering techniques helps to increase the efficiency of the image retrieval process.

REFERENCES

1. Kearney,Colm and Patton, J. Andrew, "Survey on the image segmentation", Financial Review, 41: 29-48 (2000).
2. H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, 34: 2259–2281, (2001).
3. W. K. Pratt, *"Chapter 17: Image Segmentation,"* in *Digital Image Processing*, 3rd Edition New York: John Wiley and Sons, 2001, pp. 551-588.
4. R. M. Haralick and L. G. Shapiro, *"Image segmentation techniques," Computer. Vision Graphics, Image Processing*, vol. 29, pp. 100-132, 1985.
5. R. C. Gonzales and R. E. Woods, *"Chapter 10: Image Segmentation,"* in *Digital Image Processing*, 2nd ed Upper Saddle River, N.J.: Prentice Hall, 2002, pp. 567-642.
6. K. S. Fu and J. K. Mui, "*A survey of image segmentation," Pattern Recognition*, vol. 13, pp. 3-16, 1981.
7. N. R. Pal and S. K. Pal, *"A review on image segmentation techniques," Pattern Recognition*, vol. 26, pp. 1277-1294, 1993.
8. Zhou XS, Huang TS. Relevance feedback in image retrieval: A comprehensive review. Multimedia Syst;8:536-544, 2003.
9. Shirakawa, S., and Nagao, T., *"Evolutionary Image Segmentation Based on Multiobjective Clustering".*Congress on Evolutionary Computation (CEC '09),Trondheim,Norway,2466-2473,2009.
10. S. Bhattacharya, *"A Brief Survey of Color Image Preprocessing and Segmentation Techniques"* Journal of Pattern Recognition Research 1 (2011) 120-129.
11. Thrasyvoulos N. Pappas, An Adaptive Clustering Algorithm for Image Segmentation, IEEE Transactions on Signal Processing Vol 40 no 4 April 1992.
12. Hoel Le Capitaine, Carl Fr´elicot, *"On selecting an optimal number of clusters for color image segmentation",* International Conference on Pattern Recognition,2010.
13. Irani, A.A.Z. Belaton, *"A K-means Based Generic Segmentation System"* B.Dept. of Comput. Sci., Univ. Sains Malaysia, Nibong Tebal, Malaysia Print ISBN: 978-0-7695-3789-4 On page(s): 300 – 307, 2009.
14. Li Wenchao Zhou Yong Xia Shixiong China Univ. of Min. & Technol., Xuzhou, *"A Novel Clustering Algorithm Based on Hierarchical and K-means Clustering"* Print ISBN: 978-7-81124-055-9, On page(s): 605, 2009.
15. Huiyu Zhou, Abdul H. Sadka, Mohammad R. Swash, Jawid Azizi and Abubakar S. Umar., *"Content Based Image Retrieval and Clustering: A Brief Survey"* school of Engineering and Design, Brunel University, Uxbridge, UB8 3PH, UK
16. Huang Min,Sun bo,Xi Jianqing *"An Optimized image retrieval method based on Hierarchical clustering and genetic algorithm*" forum on Information technology and applications, 978-0-7695-3600-2/09-IEEE,2009.
17. Zhuowen Tu, Song-Chun Zhu, *"Image Segmentation by Data Driven Markov Chain Monte Carlo", IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, May 2002, pp 657-673.
18. H. Choi, R.G. Baraniuk, *"Multiscale Image Segmentation Using Wavelet-Domain Hidden Markov Models" IEEE Transactions on Image Processing*, Vol. 10, No. 9, Sept 2001, pp 1309-1321.

AUTHOR PROFILE

Ram Kishan Dewangan, received B.E. (Computer Science & Engineering) in year 2005 and in pursuit for M.Tech. (Computer Sc.) from Chhatrapati Shivaji Institute of Technology (CSIT), Durg, Chhattisgarh, India, His interests are Digital Image Processing, Compiler Design, Data Structures, Theory of Computation and Database Management System. Also he is having Life Membership of Indian Society of Technical Education, India (ISTE) membership No-LM 62936 and Institutional Member of Computer Society of India (CSI) Member ship No-00175227.

Prof. Tripti Sharma, received B.E. (Computer Sc.) and M.Tech. (Computer Sc.) in the years 2002 and 2010 respectively. Currently working as Associate Professor and Head in the Department of Computer Science & Engineering at Chhatrapati Shivaji Institute of Technology (CSIT), Durg, Chhattisgarh, India, Her interests are Digital Image Processing and Data Mining. Also she is having Life Membership of Indian Society of Technical Education, India (ISTE), Membership No-LM 74671 and Institutional Member of Computer Society of India (CSI). Membership No-N1009279.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)