Jaskirat Kaur, Sunil Agrawal, Dr. Renu Vig/ International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 2,Mar-Apr 2012, pp.664-667

A Methodology for the Performance Analysis of Cluster Based Image Segmentation

Jaskirat Kaur
Student UIET
Chandigarh

Sunil Agrawal
Assistant prof.,UIET
Panjab university
Chandigarh

Dr. Renu Vig
Director, UIET
Panjab university
Chandigarh

Abstract- Partitioning of an image into several constituent components is called image segmentation. Numerous algorithms using different approaches have been proposed for image segmentation. Today many data clustering algorithms are being used for segmenting images. A major challenge in segmentation evaluation comes from the fundamental conflict between generality and objectivity. As there is a glut of image segmentation techniques available today, customer who is the real user of these techniques may get obfuscated. In this paper to address the above described problem a review is done on different types of clustering methods used for image segmentation. Also a methodology is proposed to classify and quantify different clustering algorithms based on their consistency in different applications. This paper also describes the various performance parameters on which consistency will be measured in the proposed methodology.

I. INTRODUCTION

Partitioning of an image into several constituent components is called image segmentation. Segmentation is an important part of practically and automated image recognition systems, because it at this moment extracts the intensity objects, for further processing such as description or recognition [1]. It is widely used in exploratory pattern-analysis, grouping, decision making, machine learning situations, including data mining, document retrieval and pattern classification [2]. In many such above mentioned cases, there is little a priori information available about the data and we need to make as many assumptions as possible. Under all these restrictions clustering methodology is particularly appropriate for the exploration of interrelationship among the data points to make an assessment of their structure [2].

Numerous algorithms using different approaches have been proposed for image segmentation. These approaches include local edge detection, morphological region based approaches etc. Some intensity based methods such as thresholding and histogram based models are easy to be formulated and are fast. However they fail to segment objects with low contrast or noisy images with varying background [3].

Today many data clustering algorithms are being used for segmenting images. They are termed as unsupervised methods for segmentation of images. In such techniques, image is separated into a set of disjoint regions with each region associated with one of the finite number of classes that are characterized by distinct parameters [3]. Therefore till date many types of segmentation techniques have been developed and many data clustering techniques are being used for segmentation of images [4].

A potential problem for a measure of consistency between different segmentations available is that there is no unique segmentation of an image. For example two people may segment an image differently because they either perceive the scene differently, or they segment at different granularities. If two different segmentations arise from different perceptual organizations of the scene, then it is fair to declare the segmentations inconsistent [5].

A major challenge in segmentation evaluation comes from the fundamental conflict between generality and objectivity. For general-purpose segmentation, segmentation accuracy may not be well defined, while embedding the evaluation in a specific application, the evaluation results may not be extensible to other applications. Reliable segmentation performance evaluation for quantitatively positioning image segmentation is extremely important. In many prior works, segmentation performance is evaluated by subjectively or objectively judging several sample images. Such evaluations lack statistical meanings and may not be generalized to other images and applications [4].

As there is a glut of image segmentation techniques available today, customer who is the real user of these techniques may get obfuscated. In this paper to address the above described problem a review is done on different types clustering methods used for image segmentation. Also a methodology is proposed to classify and quantify different clustering algorithms based on their consistency in different applications. This paper also describes the various performance parameters on which consistency will be measured in the proposed methodology [3].

The rest of the paper is organized as follows: Section II gives introductory information about cluster based image segmentation. Section III reveals the information about advancements to the basic. Proposed methodology for classification of algorithms is described in Section IV. Some conclusion is given in Section V.

II. CLUSTER BASED IMAGE SEGMENTATION

Extracting information from an image is referred to as image analysis. It is one of the preliminary steps in pattern recognition systems. Each region of the image is made up of set of pixels. Partitioning an image into several disjoint segments is what is termed as image segmentation. It simplifies and changes the representation of an image, image is transferred into something more meaningful and easier to analyze. Typically it is used to locate objects of interest and boundaries like lines, curves in an image [1]. Segmentation algorithms are based on two basic properties of an image intensity value: discontinuity and similarity. To study discontinuities in an image we divide image based on the abrupt changes in intensity such as edges.

Mathematically the regions we obtain after partitioning an image into regions is considered to be homogeneous with respect to some image property of interest. Image property can be intensity, color, or texture.

If

$$I = \{x_{ij}, i = 1 \ldots N_r, j = 1 \ \ldots N_c\} \qquad (1)$$

is the input image with N_r rows and N_c columns and measurement value x_{ij} at pixel (i, j), then the segmentation can be expressed as $l = \{S_1, \ldots., S_k\}$ with the l_{th} segment

$$S_l = \{(i_{l_1}, j_{l_1}), \ldots, (i_{l_{N_1}}, j_{l_{N_1}})\} \qquad (2)$$

consisting of a connected subset of the pixel coordinates. No two segments share any pixel locations and the union of all the segments covers the entire image.

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 researchers in the field. Segmentation can be achieved in a more efficient manner through clustering.

Clustering is an interesting approach for finding similarities in data and putting similar data into groups. Cluster partitions data set into several groups such that the similarity within a group is larger than that among the groups. Clustering algorithms are used extensively not only to organize and categorize data, but are also useful for data compression [7].

The segmentation of images presented to an image analysis system is critically dependent on the scene to be sensed, the imaging geometry, configuration, and sensor used to transduce the scene into a digital image, and ultimately the desired output of the system [7].

The applicability of clustering methodology to the image segmentation problem was recognized over three decades ago, and the paradigms underlying the initial pioneering efforts are still in use today. It defines feature vectors at every image location called as pixel component of both functions of image intensity and functions of pixel location itself.

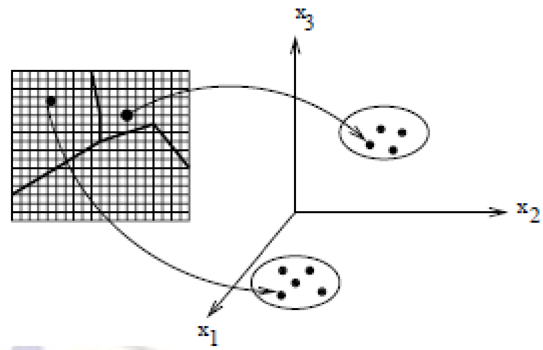

Fig.1. Feature representation for clustering

The basic idea of assigning pixel values is depicted in figure 1. In the above figure image measurements and positions are transformed to features. Also clusters in feature space correspond to image segments [7].

Historical records show that clustering is a powerful tool for obtaining classifications of image pixels. The key issues in the design of any clustering – based segmenter are the choice of pixel measurements (features) and dimensionality of the feature vector, a measure of similarity which is appropriate for the selected features and the application domain, the identification of the clustering algorithm, the development of strategies for feature and data reduction, and the identification of necessary pre- and post-processing techniques. The new variations of clustering using segmentation continue to emerge. Challenges to the more successful use of clustering include the conflict between the two parameters defined below.

Good objectivity means that all the test images should have an unambiguous segmentation so that segmentation evaluation can be conducted objectively. Good generality means that test images should have a large variety so that the evaluation results can be extended to other images and applications. There always exists a well known dilemma between objectivity and generality in segmentation evaluation [4]. There is no such unique clustering technique which can segment all types of images uniquely and unambiguously.

III. ADVANCEMENTS TO THE BASIC

There is a large literature on the image segmentation and clustering algorithms have been developed in the past few decades, with application in many areas.

In this section, some of the related work is presented that is most relevant to the approach of the paper. Most clustering based segmentation methods attempt to segment images from one particular type of application [1]. Most previous works are developed to compare different clustering based image segmentation algorithms based on characteristics such as correctness, stability with respect to parameter choice and stability with respect to image choice [5]. The effectiveness of a new algorithm over the previous one is demonstrated only by the presentation of the few segmented images and the expert decides the consistency of the algorithms based on the application requirements. To overcome the problem of objective numerical evaluations, [5] introduces some parameters like global consistency error and local consistency error based on which the effectiveness is decided of the two image segmentation algorithms. Other method presented in [4] tries to resolve the conflict between the generality and objectivity. They present a new benchmark to evaluate five different image segmentation methods according to their capability to separate a perceptually salient structure from the background with a relatively small number of segments. In this paper, the evaluation is done based on the some image segmentation algorithms. Not all soft computing methods which are considered efficient over the traditional computing methods [1] are used. Also the images which are taken are assumed to have any unambiguous salient objects, or they have multiple equally salient objects. Comparison is done based on the number of segments used. It shows as number of segments is less watershed method works worst and if they are near 20 Normalized-cut method outperforms other methods. Still, the problem of general purpose segmentation remains far from well solved. In [6] four data clustering methods are taken and are implemented and tested against a medical problem of heart disease diagnosis. Data is given as input to measure the performance and accuracy of four techniques are presented. The methods used are K means, Fuzzy c means, mountain and subtractive clustering methods. Comparison is made based on root mean square error (RMSE), accuracy, regression line slope and time required for the algorithms. But still no comparison is made on images from different fields. Performance analysis based on quality measures [3] like structural content, peak signal to noise ratio, average difference, image fidelity and normalized correlation coefficient is done on three types of medical images such as MRI, X-ray and ultrasonic images. The performance of quality measures is checked through K means and Expectation maximization model.

Most of the methods discussed above use segmentation algorithms to segment images but objectivity and generality are not defined clearly for clustering algorithms used for image segmentation.

In this paper, a methodology is proposed to test the consistency of different cluster based image segmentation algorithms. Quantification of different algorithms will be based on some objective parameters used for all algorithms.

IV. PROPOSED METHODLOGY

In this section to study the relative performance of cluster based image segmentation methods the following methodology is proposed.

A. Test image database construction

At the first stage in testing the consistency we collect around 100 different real natural images from internet, different repositories and some well known image databases such as corel. A particular requirement is that each image should have different area of interest to be analyzed. This database consists of images from different fields and each image is totally different from other. Fields chosen can be medical, remote sensing, architectural, industrial images etc. To make this benchmark for testing consistency suitable for evaluating large variety of cluster based image segmentation algorithms, color information is removed and all the images will be unified to 256-bit gray scale image in JPEG format.

B. Selected cluster based image segmentation methods

We will evaluate the following cluster based image segmentation methods.

1) K means
2) Expectation Maximization model
3) Fuzzy c means
4) Normalized cut method
5) K means or fuzzy c means combined with particle swarm optimization technique [8][9]

We chose these methods based on the following considerations:

1) All the above methods are relatively new methods for image segmentation.
2) They represent the different categories of segmented images well.

C. Performance Measure

To study the relative performance of cluster based segmentation methods the following quality measures are calculated.

1) Structural content (SC)
The method to calculate is

$$SC = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N} r(j,k)^2}{\sum_{j=1}^{M}\sum_{k=1}^{N} t(j,k)^2}$$

The large value of SC means image is of poor quality.

2) Normalized Correlation Coefficient (NK)
The method to calculate is

$$NK = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N}[r(j,k)t(j,k)]}{\sum_{j=1}^{M}\sum_{k=1}^{N}[r(j,k)^2]}$$

3) Peak signal to noise ratio (PSNR)
The method to calculate is

$$PSNR = 10.log_{10}\left[\frac{\max(r(x,y))^2}{\frac{1}{n_x.n_y}\left[\frac{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)]^2}{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)-t(x,y)]^2}\right]}\right]$$

The above equation calculates PSNR in decibels. The small value of PSNR means the image is of poor quality.

4) Root mean square error (RMSE)

The method to calculate is

$$RMSE = \sqrt{\frac{1}{n_x.n_y}\left[\frac{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)]^2}{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)-t(x,y)]^2}\right]}$$

The above four parameters are calculated based on the input image $r(x,y)$ and the segmented image $t(x,y)$.

5) Compression Ratio (CR)
The method to calculate is

$$CR = \frac{n_1}{n_2},$$

where n_1 and n_2 denote the number of information carrying bits in the original and encoded images

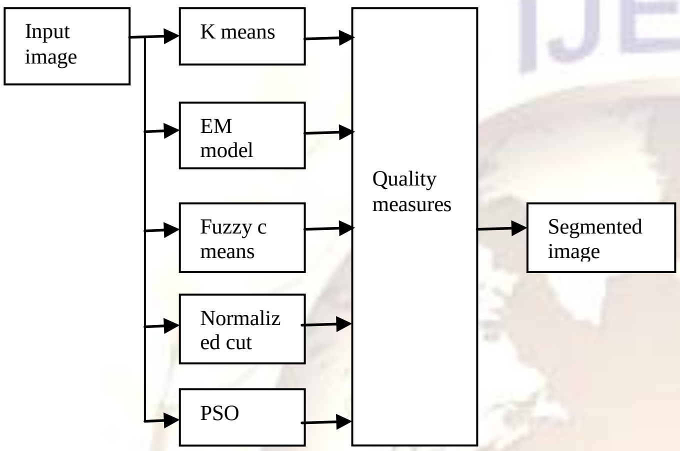

Fig. 2. Block diagram of the proposed methodology

As shown in figure 2, each image will be inputted to each algorithm and the output from each algorithm will be tested for five different quality measures. This depicts that for one image we have 25 results. Generality for particular algorithm will decided after testing images from all fields. Generality over here means we need to decide that for all type of medical images one particular algorithm will work better. As we know hundred percent results are unlikely to obtain, so we decide one threshold [4] for segmentation performance. Seventy percent can be considered as the performance threshold, it means if particular algorithm works fine for seventy percent of the images we say it can be generalized to other images of the same field.

V. CONCLUSION

In this paper, we review the application of different image segmentation and clustering algorithms. These approaches solve the problem of checking the consistency of different algorithms based on some small number of images or images from one particular field. Also the parameters taken for evaluation are small in number. As the generality in cluster based image segmentation is missing, it means they lack statistical meanings. To solve the above problem of generality, a methodology is proposed based on which the performance analysis of different cluster based image segmentation methods will be done.

REFERENCES

[1] B.Sowmya and B.Sheelarani, "Colour Image Segmentation Using Soft Computing techniques," *International Journal of Soft Computing Applications*, Issue-4(2009), pp.69-80.

[2] Shilpa Agarwal, Sweta Madasu, Madasu Hanmandlu and, Shantaram Vasikarla, "A Comparison of some Clustering Techniques via Color Segmentation," *International Conference on Information Technology: Coding and Computing (ITCC'05)*, 0-7695-2315-3/05, IEEE.

[3] R.Harikumar, B. Vinoth Kumar and, G.Karthick and I.N. Sneddon, "Performance Analysis for Quality measures Using K means Clustering and EM models in segmentation of medical Images," *International Journal of Soft computing and Engineering*, vol. 1, Issue-6, pp.74-80, January 2012.

[4] Feng Ge, Song Wang and, Tiecheng Liu, "New Benchmark for image segmentation evaluation," *Journal of Soft Electronic Imaging*, 16(3), 033011, pp.033011-1-16, Jul-sep 2007.

[5] Alina Doringa, Gabriel Mihai, Liana Stanescu, Dumitru Dan Burdesce, "Comparison of Two Image Segmentation Algorithms," *Second International Conferences on Advances in Multimedia*, 978-0-7695-4068-9/10, 2010 IEEE.

[6] Khaled Hammouda, "A Comparative Study of Data Clustering Techniques," *Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada*

[7] A.K. Jain, M.N. Murty and, P.J. Flynn, "Data clustering: A Review," *ACM computing Surveys*, vol.31, No. 3, September 1999

[8] Li Wang, Yushu Liu, Xinxin Zhao and, Yuanqing Xu, " Particle Swarm Optimization for Fuzzy c- Means Clustering," 6th *World Congress on Intelligent Control and Automation*, pp. 6055-6058, June2006.

[9] Sheng Su, "Image Classification based on Particle Swarm Optimization combined with K-means," *International Conference on Test and Measurement*, 978-1-4244-4700-8/09, IEEE 2009