RESEARCH ARTICLE OPEN ACCESS

A Review of Image Segmentation of Underwater Images Using Fuzzy C- Means Clustering

Vikram Dwivedi[1] ,Dr. Paresh Rawat[2]
[2]Member IEEE,
Truba college of Science and Technology Bhopal

----------**********----------

Abstract:

The underwater image segmentation is a challenging field of research due to poor illumination condition. Therefore this paper reviews the various issues and problems based on the existing researches in underwater segmentation field. Since underwater images have many uses in general purposes like in marine engineering, analysis and monitoring of underwater animals and plants and for monitoring of oil wells. Thus, underwater segmentation problem can be considered as fuzzy problem. Therefore this paper reviews the fuzzy clustering algorithm for image segmentation. Fuzzy C-Means clustering based underwater image segmentation methods separates the underwater objects using clustering and thresholding. It is found that fuzzy based approaches are not such efficient in underwater environment and thus not much explored. The efficiency of the fuzzy based methods various with different images and objects. Paper pointed out various challenges after reviewing the previous work. It is require improving the efficiency of the fuzzy based systems for segmenting the underwater images.

Keywords—**Image segmentation, clustering, Edge-based segmentation, Hybrid segmentation.**

----------**********----------

1. INTRODUCTION

Image segmentation is initial step in image analysis and pattern recognition. Segmentation is used for classifying the image into many groups. Image segmentation methods can be classified based on histogram thresholding [1], edge-based [2] or hybrid segmentation [3] and Fuzzy segmentation [4] approach. Histogram thresholding assumes that image composed of one color corresponding to one region. Edge based segmentation utilizes the high frequency image features for segmentation. Region based segmentation is based on region thresholding, clustering, region growing, merging splitting come inside this. Hybrid segmentation [3] takes advantage of both global and region based segmentation approaches.

Underwater images suffer from various types of problems due to underwater environment. First is density, since the water is a denser medium as light enters deeper in the water they get refracted. As light travels in water from air meets denser medium, if light crosses the line at right angle we get appropriate image, if light travels in water at obtuse angle we get refraction in light. This causes the loss of light illumination. Another problem is loss of true color. The pure water has a strong cyan or blue green color and absorbs different color at different rate. The fourth problem is contrast of image, underwater poor illumination and non uniform illumination reduces contrast of images. All these problems make the underwater image segmentation a bit difficult task.

Fuzzy clustering based methods are basically used for medical imaging field. Clustering is the group of organization of objects of similar in color, similar in shape in size, clustering deals with finding the similar objects.[7] There are mainly type of clustering are Fuzzy C-Means clustering and K-Means clustering. We are using Fuzzy C-Means clustering for image segmentation. Fuzzy C-Means clustering is the clustering which allows one object for many clusters. Image segmentation is the method of portioning a digital image into multiple segments (sets of pixels also known as super pixels) in computer vision-means clustering is popular for cluster analysis in data mining. The Fuzzy C-Means clustering is also known as soft K-Means algorithm. Image segmentation methods are of following types, grouping algorithm, Normalized cuts-A graph portioning method, EM algorithm and K-Means clustering. An important component of clustering measure is the distance measure ,objects belong to same cluster if they are close in distance ,this is called distance based clustering .another is conceptual clustering when objects belong to same cluster in a given concept. Expectation Maximization (EM) is one of the most common algorithms used for density estimation of data points in an unsupervised setting. In EM, alternating steps of Expectation (E) and Maximization (M) are performed

iteratively till the results obtained. [6]Normalized cuts method image segmentation can also be viewed as an optimal partitioning of a graph. The image is presented as a weighted undirected graph G = (V, E). This image graph can be partitioned into two sub graphs. K-Means clustering is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance. The algorithm assumes data form a vector space and tries to find natural clustering in them.

Possible applications of clustering are in many field like in library, earthquake study and other fields.[5]

2. REVIEW WORK

Fuzzy C-Means clustering algorithm was developed by Dunn in 1973 and improved by Bezdek in 1981.In Fuzzy C-means we have to assign each data point corresponding to each cluster center on the basis of the distance from the data point to cluster center.

3. Contrast image enhancement techniques:

Contrast is the difference in brightness value of two surfaces. Histogram enhancement techniques are of mainly two types CLAHE and Histogram Equalization. CLAHE is contrast limit adaptive histogram equalization and use for contrast enhancement and is also used for Histogram Equalization. CLAHE is adaptive histogram equalization image enhancement. Contrast enhancement is of two type linear contrast enhancement and non-linear contrast enhancement.

Linear contrast enhancement: Linear contrast enhancement tells us about linear enhancement of contrast of an image. Linear contrast enhancement is of mainly three types Max-Min Linear stretch enhancement, Percentage Linear contrast enhancement and Piecewise Linear contrast enhancement. A gray scale image lies between 0 to 255 pixels. Let an image lies between 44 to 59 pixels. Then Minimum intensity values are 44 and Maximum intensity value is 59.

Max-Min Linear contrast enhancement-Max-Min Linear contrast enhancement tells us about Minimum contrast range and Maximum contrast range of an image, Let an image pixels lies between 43 to 48 then the Minimum range of image will be 0 to 43 and Maximum range will be from 48 to 255.In gray scale image there are total 256 no. of pixels because Grayscale image lies between 0 to 256 number of total intensity values is the total number of individual intensity numbers. If an image lies between 0 to 256 intensity values then total number of intensity values is 256.

Percentage Linear contrast stretch: Percentage Linear contrast stretch tells us about linear contrast stretch from mean of the histogram. A standard deviation is used in Percentage Linear contrast stretch.

Piecewise Linear contrast stretch: Percentage Linear contrast stretch enhancement is about enhancement of certain points of an image, If an image lies between 43 to 48 then Piecewise Linear contrast stretch is enhancement of image in certain points like from 43 to 45 and 46 to 47.

Nonlinear contrast image enhancement: In Nonlinear contrast image enhancement there many types, in which CLAHE and Histogram equalization [8].CLAHE and Histogram Equalization are enhancement techniques used in the contrast enhancement of the image. CLAHE and Histogram Equalization are used for contrast enhancement of certain limits of pixels. On-Linear contrast enhancements give many results for single input so it became hard to obtain the image between many images. Histogram equalization is one of Non-Linear contrast limit enhancement in which certain number of histogram peaks is enhanced. In Histogram Equalization Image is redistributed in the Pixel values for each user specified ranges like 32, 64.44.Histogram Equalization Enhances the Peak values of the Image. Histogram Equalization enhances the image by enhancing darker side of Images to brighter and by enhancing the brighter side of Image to brighter side of image.

Histogram: Histogram tells us the graph plot of contrast of an Image.

Contrast: Contrast is the Image Intensity values of an Image. Contrast decides the brighter and darker part of an Image.

Adaptive Histogram equalization: CLAHE is Contrast Limit Adaptive Histogram Equalization. CLAHE is used for image enhancement overcoming the problem of Histogram Equalization. CLAHE divides the image into rectangular domains linked with each other. CLAHE decides the uniformity of the image. Image stretch is also used for contrast enhancement of the image. Image blurring and image with addition with noise are also used for image enhancement. Image adjust is also Contrast enhancement technique in which we get enhanced image[7]. Image resize is used to resize the image. These all image enhancement techniques are used in the gray image enhancement.

Image enhancement with homomorphic filters: Image enhancement with Homomorphic filters is use for enhancing the image in both High frequency range and Low Frequency range. in Homomorphic filter. There are many other image enhancement techniques which are used for RGB color images. Filter passes some certain frequency of values and stops some frequency values. In Homomorphic filters are the filters enhance the High frequency images in the High Frequency range and Low frequency images in the low frequency range.

4. PROCESS FLOW GRAPH

Clustering Based Methods

There are various clustering based methods out of which popular methods are reviewed here sequentially.

A) K-Means clustering: The main idea in clustering is to define K new centroids one for each cluster. The centroids should be placed in cunning way because different location causes different results. It will be better to place them as far as possible. The next step is to take each point belonging to the data set and associate it to the nearest centroid.When there is no point first step is completed and now recalculates K new centroids. It is from previous results. Now repeating the previous step associate the data to the nearest centroid this process will stop until centroids do not

move further. Finally in the process K new centroids move their position we repeat this until the centroids do not move further.

This algorithm aims at minimizing the following objective function, in this case a squared error function. The objective function is:

$$J = \sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)} - c_j\right\|^2 \quad (6)$$

Where $\|x_i^{(j)} - c_j\|^2$ is the distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j . $\|x_i^{(j)}\text{-}c_j\|^2$ Is the indicator of the distance between n data points and their respective cluster centers? It can be proved that the procedure will always terminate, the K-Means algorithm does not necessarily find the most optimal configuration corresponding to the global objective function minimum.

The algorithm is also for randomly selected clusters. The K-Means algorithm is used multiple points to reduce this effect. Clustering deals with scalability, interpretability, usability, dealing with different type of attributes, discovering clusters with proper dimensions, ability to deal with noise, minimum requirements to determine the input parameters and high dimensionality.

K-Means as shown in Figure 1 is Exclusive clustering algorithm in which data belongs to single centroid.

Fuzzy C-Means clustering is the overlapping clustering. In Fuzzy clustering single data is connected to each cluster. In overlapping clustering every data is connected to every cluster. Fuzzy C-Means clustering is most efficient than any other clustering like K-Means clustering method.[11]

Below the Process Flow graph of K-Means clustering is drawn.

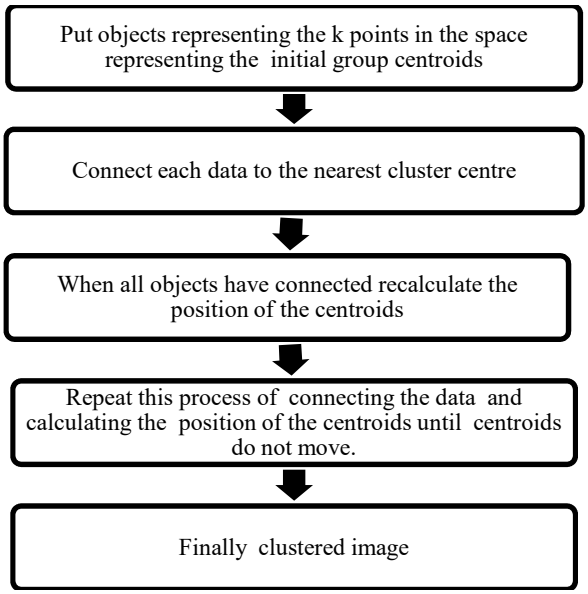

Figure 1- Process flow graph of K-Means clustering algorithm

Fuzzy C-Means clustering algorithm is based on the minimization of the following objective function:

$J_M = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^m \|x_i - c_j\|^2$, 1≤m<∞

Where m is fuzziness exponent and any real number greater than 1,u_{ij} is the degree of membership of x_i in cluster j,x_i is the ith of d-dimensional measured data, c_j is the d-dimension center of the cluster and $\|*\|$ is any norm expressing the similarity between any measured data and the center.[9]

Fuzzy portioning is carried out through an iterative optimization of the objective function shown above should be in the above range. With the update of membership u_{ij} and the cluster center c_j by

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \quad (1)$$

This equation can be further elaborated as;

$$= \frac{1}{\left(\frac{\|x_i - c_j\|}{\|x_i - c_1\|}\right)^{\frac{2}{(m-1)}} + \left(\frac{\|x_i - c_j\|}{\|x_i - c_2\|}\right)^{\frac{2}{(m-1)}} + \cdots + \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{(m-1)}}} \quad (2)$$

Where cluster center c_j is defined as;

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m} \quad (3)$$

Where $\|x_i - c_j\|$ is the distance from i to current cluster centre j, $\|x_i - c_k\|$ is the distance from I to other cluster centers k. Iteration will stop when

$$max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \varepsilon,$$

Where ε is a termination criterion between 0 and 1, whereas k is the iteration steps. This procedure converges to a local minimum or a saddle point of J_m.

The algorithm is composed of the following steps:

1. Initialize U=$[u_{ij}]$ matrix,$U^{(0)}$.
2. At K-step calculate the centers vectors $C^{(k)} = [c_j]$ with $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m} \quad (4)$$

3. Update $U^{(k)}, U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \quad (5)$$

4. If $\|U^{(k+1)} - U^{(k)}\| < \varepsilon$ then stop the process otherwise return to step 2.

Advantages of Fuzzy C-Means clustering:

1. Fuzzy C-Means clustering gives the best result for overlapping data set and comparatively better than K-Means algorithm.[15]
2. In k-Means clustering each data belong to one cluster center, but in Fuzzy C-Means clustering each data belong to each cluster center, so the data connected to every cluster center.[12]

Disadvantages of Fuzzy C-Means Clustering algorithm:

1. We get best result but number of iterations should be high.

2. The information about no. of clusters is apriority.
3. Euclidean distance measures can unequally weight underlying factors.

There is many other segmentation processes like clustering, in which Fuzzy C-Means clustering and K-Means clustering are two type of segmentation. Fuzzy clustering is most widely used than K-means clustering. Images may be two dimensional such as such as photograph, screen display and as well as three dimensional such as statue. They may be captured by optical devices such as mirrors, lenses, telescopes, microscopes and natural object and phenomena such as the human eye or water.[13]

Fuzzy C-Means clustering algorithm as shown in Figure 2, is of many types like Fuzzy C-Means clustering algorithm with using thresholding like CLAHE and Histogram Equalization. CLAHE is contrast limited adaptive histogram equalization. Fuzzy clustering is Fuzzy C-Means clustering algorithm. Below the process flow graph of Fuzzy C-Means clustering algorithm is drawn.

In this objects are in the form of point spaces representing the initial group centroids .The iteration of the process does not stop until the desired value is obtained.[14]

The main problem is distance measure in the Fuzzy C-Means clustering and K-Means clustering algorithm. Euclidean distance is used for calculating the distance value of data points.Mahalanobis distance is also used for distance measurement but Euclidean distance is widely used than Mahalanobis distance.

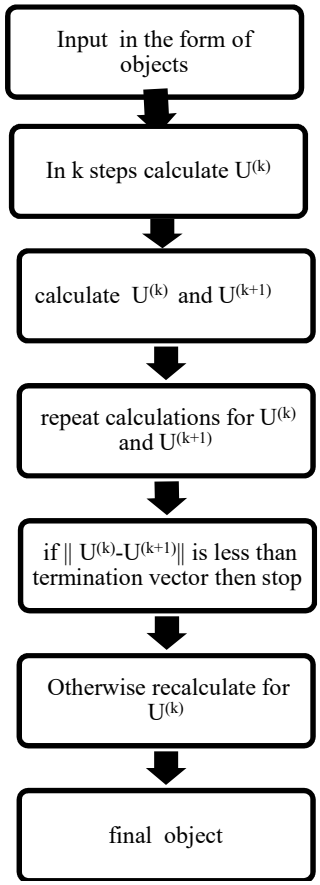

Figure 2 Process flow graph of Fuzzy C-Means clustering algorithm

4. PROBLEM SOLUTION

1. Pre contrast improvement may improve the segmentation performance.
2. Instead of using single fuzzy clustering method use of two images for clustering may improve the performance of the system
3. It is also required to reduce the complexity of fuzzy based methods
4. Using C means may perform better than K means for underwater environment.

Conclusion: Marine image segmented and taken as second input with original image. Fuzzy C-Means clustering gives good outputs. We analyze it and take as an input. Fuzzy C-means clustering give better results than any other clustering.

References:

[1] Marine Images Segmentation Using Adaptive Fuzzy c-Means Algorithm Based on Spatial Neighborhood WANG Shi-long,XU Yu-ru WAN Lei Tang Xu-dong National Key Laboratory of Science and Technology on Autonomous Underwater Vehicle,Harbin Engineering University,Harbin,China wangi5895@yeah.net IEEE

[2]Non linear Image segmentation using fuzzy c means clustering method with thresholding for underwater images Dr.G.Padmavathi,Mr.M.Muthukumar and Mr. Suresh Kumar Thakur.IJCSI International Journal of Computer Science Issues, Vol. 7 Issue 3,No 9,May 2010 ISSN (Online):1694-0784 ISSN(Print):1694-0814

[3]Level set methods for watershed image segmentation Erland Hodneland, Xue-Cheng Tai Joachim Weickert Nickolay V. Bukoreshtliev Arvid Lundervold and Hans-Hermann gerdes

[4] SURVEY OF IMAGE CONTRAST ENHANCEMENT METHODSCHUMMAN LAL CHANDRAKAR1, MONISHA SHARMA2 & S. K. SINGHAI3 Sscet Bhilai ,Sscet Bhilai , Gec Bilaspur, India

[5] AN OVERVIEW ON CLUSTERING METHODS T. Soni Madhulatha Associate Professor, Alluri Institute of Management Sciences, Warangal.

[6] Abnormality Identification from MRI Images Using Curvelets Transform Jasmeen Er. Parvinder Singh CSE, Asra Institute of Engg and Technology Asst. Prof. CSE, Asra Institute of Engg Punjab Technical University, Jalandhar, India and Technology PTU, Jalandhar, India

[7] Various Image Enhancement Techniques- A Critical Review S.S. Bedi1 , Rati Khandelwal2 Assistant Professor, CSIT, MJPRU, Bareilly, India 1 M.Tech IVth Semester, Student, Banasthali Vidyapith, Jaipur, India 2

[8] “Comparison of Histogram Equalization Techniques for Image Enhancement of Grayscale images of Dawn and Dusk” Dinesh Sonker1 , M. P. Parsai2 1, 2

(Department of Electronics and Comm. Engineering, Jabalpur Engineering College, Jabalpur, M.P., 482011, India)

[9] Volume 1, Issue 6, November 2013 International Journal of Advance Research in Computer Science and Management Studies Research Paper Available online at: www.ijarcsms.com Review on determining number of Cluster in K-Means Clustering Trupti M. Kodinariya1 Dr. Prashant R. Makwana 2 Research Scholar in JJT University, Department of Computer Engineering, Jhunjhunu, Rajasthan - India Director - Research GRMECT Research Center Rajkot – India

[10] Image enhancement method for underwater, ground and satellite images using brightness preserving histogram equalization with maximum entropy Singhai, J. ; MANIT Bhopal, Bhopal,Paresh rawat.

[11] Performance Comparison of Fuzzy C Means with Respect to Other Clustering Algorithm Tejwant Singh1 , Mr. Manish Mahajan2 1 Research Scholar, 2 Associate Professor, Dept of IT, Chandigarh Engineering College, Mohali, India.

[12] Implementation of the Fuzzy C-Means Clustering Algorithm in Meteorological Data Yinghua Lu 1 ∗, Tinghuai Ma1, , Changhong Yin2 , Xiaoyu Xie2 , Wei Tian1 and ShuiMing Zhong1 1 School of Computer & Software, Nanjing University of Information Science & Technology, Nanjing 210044; 2Wuhan Meteorological Service, Wuhan 430040 cloveryunyan@163.com

[13] Image Segmentation using k-means clustering,EM and Normalized Cuts Suman Tatiraju Department of EECS University Of California – Irvine Irvine, CA 92612 statiraj@uci.eduAvi MehtaDepartment of EECS University Of California - IrvineIrvine, CA 92612 avim@uci.edu

[14] Determination of Number of Clusters in K-Means Clustering and Application in Colour Image Segmentation

Siddheswar Ray and Rose H. Turi School of Computer Science and Software Engineering Monash University, Wellington Road, Clayton, Victoria, 3168, Australia E-mail: {sid,roset}@csse.monash.edu.au

[15]Color Image Segmentation Using K-Means Clustering and Otsu's Adaptive Thresholding Vijay Jumb, Mandar Sohani, Avinash Shrivas