Color Image Clustering using Hybrid Approach based on CIELab Color Space

Vaneet Kumar
Lecturer
QIF-GOI, Mohali,
Punjab Technical University

Anu Bala
Assistant Professor
SBBSIET, Hoshiarpur,
Punjab Technical University

Seema Bansal
Assistant Professor
CCET, Chandigarh
Punjab University

ABSTRACT

In contrast with the problem of phantom colors which fails the slow spatial variation along the edges results in the poor clustering of pixels from the color image science.[1] In this paper color clustering results are improved using an hybrid approach where we are combining the benefits of bilateral filter with an improved ant based clustering algorithm. Bilateral filter is a non iterative, local and simple method for smoothening and edge preserving of an image. It combines the colors based on both their vicinity in the domain and range. in our approach ants are created dynamically on the spatial grid. Ants pick its respected data item using CMC as pheromone. As CMC similarity measure is best suitable for CIELab color space to quantifies the perceptual visual differences. The initial knowledge of the number of clusters is not required during the clustering process. As MSE is the global quality measure we applied here using Euclidian distance to evaluate the performance of proposed technique that decay with dissimilarity in clusters.

Keywords:- ACO, Bilateral Filtering, CMC, MSE, CIELab

1. INTRODUCTION

Segmentation is the one of the most important and challenging low-level problems in color image analysis. It consists in determining the regions of the image which extract the meaningful objects in the observed scene.[3] Image segmentation can be viewed as a clustering problem, which aims to partition the image into clusters such that the pixels within a cluster are as homogenous as possible where as the clusters among each other are as heterogeneous as possible with respect to a similarity measure (5) Pixel clustering is performed by computing the color differences using CMC distance measure to meet the perceptual similarity. As the intensities of pixels which change drastically on the boundaries of object in the image results in poor and complex segmentation [16]. So to improve the results of segmentation even on the edges of object we have used hybrid approach where we are combining the benefits of bilateral filter with an improved ant based clustering algorithm.

In this paper, an improved ant-based clustering is proposed. Here, the proposed algorithm doesn’t have any assumptions about the population of ants. And the proposed method will automatically calculate the number of ants required for clustering. With this modification a hybrid ant-based clustering is presented and compared with ant clustering[8].

2. CLUSTER ANALYSIS

In the process of cluster analysis the data sets are partitioned into distinct groups that each group is characterized by unique feature and the union of no two clusters is homogeneous. In color image clustering any property of pixels can be taken such as intensity, color etc. here in our work we are taking color of the pixel as its property. Clustering is often a complex problem color images because of ambiguous boundaries between classes.

2.1 Ant Based Clustring

Image segmentation based on ant clustering was originally introduced for tasks in robotics by Deneubourg et al.[11]. Lumer and Faieta modified the algorithm to be applicable to numerical data analysis, and it has subsequently been used for data-mining, graph partitioning [5]. In literature,clustering with ants is performed by placing data on 2dimensional grid where In a course of processing ants can die and can reproduce.[11] Each ant moves around this grid picking and dropping the data items. The decision to pick up or drop an item is random but is influenced by the data items in the ant’s immediate neighborhood, thus causing similar items to be more likely placed together on the grid [10].

3. BILATERAL FILTERING

Anisotropic diffusion is an iterative approach used the solution of partial differential equations for smoothening and preserving edges of color and gray scale images. In contrast bilateral filtering has recently been proposed as a non iterative approach for smoothening and preserving edges. But unlike anisotropic diffusion, bilateral filtering does not involve the solution of partial differential equations and can be implemented in a single iteration. Despite the difference in implementation, both methods are designed to prevent averaging across edges while smoothing an image.

The idea underlying bilateral filtering is to filter in the range of an image while traditional filters do in its domain. In particular, bilateral filters can be applied to color images just as easily as they are applied to black-and-white ones.[16] The CIELab color space [2] is the space of colors with a perceptually meaningful measure of color similarity, in which CMC distance measure correlate strongly with human color discrimination performance. Thus, if we use this metric in our bilateral filter, Only perceptually similar colors are averaged together, and only perceptually visible edges are preserved [14].

The similarity function s operates in the range of the image function f as shown in Eq. (2), while the closeness function c operates in the domain of f as shown in Eq.(1).

$$k_d^{-1}(x)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi)c(\xi,x)d\xi \qquad \text{Eq. (1)}$$

where $c(\xi, x)$ measures the geometric closeness between the neighborhood center x

$$k_r^{-1}(x)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi)s(f(\xi),f(x))d\xi \qquad \text{Eq. (2)}$$

$s(f(\xi), f(x))$ measures the photometric similarity between the pixel at the neighborhood center x.

The appropriate method of bilateral filter is to combine domain (σ_d) and range (σ_r) filtering, there by enforcing both geometric and photometric locality. Combined filtering can be described as follows:

$$H(x) = k^{-1}(x) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi)c(\xi,x)s\big(f(\xi), f(x)\big)d\xi \qquad \text{Eq. (3)}$$

With normalization

$$k(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(\xi,x)s\big(f(\xi), f(x)\big)d\xi \qquad \text{Eq. (4)}$$

Given Eq. (3) replaces the pixel value at x with an average of similar and nearby pixel values. The large value of σ_d in the domain blurs more, ie. combines values from more distant pixels. Also, if an image is scaled up or down, σ_d must be adjusted accordingly in order to obtain equivalent results. Similarly, the photometric spread σ_r in the image range is set to achieve the desired amount of combination of pixel values. Pixels with values much closer to each other than σ_r are mixed together. If the image is amplified or attenuated, σ_r must be adjusted accordingly in order to leave the results unchanged.

4. COLOR MODEL

A color model is an abstract mathematical model describing the way colors can be represented as tuples of numbers. When this model is associated with a precise description of how the components are to be interpreted , the resulting set of colors is called color space. One approach to specifying color is to use the CIE chromaticity diagram. There are several models used to describe the tristimulus color scheme: RGB, CMY, YIQ, and HSI. Each model was derived for specific purposes and has certain advantages over the others. Converting between the different models is generally done by a relatively simple mapping. Some color models allow us to determine very effectively whether two colors are "similar" just by examining these numbers. CIELab color model is very useful for finding "similarity" of two color values as human perceives.[3][9][5]

4.1 CIELab Color Space

CIE L*a*b* (CIELAB) is the most complete color space specified by the International Commission on Illumination. It describes all the colors visible to the human eye and was created to serve as a device independent model.The three coordinates of CIELAB represent the lightness of the color (L* = 0 yields black and L* = 100 indicates diffuse white; specular white may be higher), The Luminance ranges from 0 to 100, the A component ranges from -120 to +120 (from green to red/ magenta) and the B component ranges from -120 to +120 (from blue to yellow*]*) [6] [7]. An important characteristic of the CIELab system is that the distance calculated between two colors is directly proportional to the difference between the two colors as perceived by the human eye.

4.2 CMC Distance Measure

CMC is a modification of CIELAB developed by the Color Measurement Committee of the Society of Dyers and Colorists.[1] Color difference calculated using CMC method are believed to correlate better with visual assessment than color differences calculated using other instrumental systems [4]. It has two parameters: lightness (l) and chroma (c).

5. LITERATURE SURVEY

Several partitioning methods are provided in the literature, they can be classified as Hard or Fuzzy algorithms [13].In Hard clustering algorithms, the pixel is assigned to one cluster. Fuzzy algorithms can assign pixels to multiple clusters. Salima Ouadfel et.al proposed an efficient ant based clustering approach which does not require any prior knowledge of the number of clusters [13]. Results conducting on synthetic and real images shows that the algorithm was able to extract the number of clusters with good clustering quality and state results are better than k-means. Efficiency have been checked by rand index and rosenberg's measure. Liqiang Liu et.al indicates that dynamic programming algorithm and genetic algorithm could not solve complex multi-stage decision problems very well because these problems are so complex. R. Laptik, D. Navakauskas. have investigate Ant colony optimization application for two-dimensional electrophoresis gel image. Obtained results show that model is suitable for overlapped protein spots segmentation. Segmentation accuracy was about 67% improve over simple threshold function segmentation. Clustering with swarm based algorithm has recently been shown to produce good results in a wide variety of real world applications[11] ACO algorithm for the segmentation of brain MR images can effectively segments the fine details[12]. R. Kaur proposed BFO technique for Color image quantization and have considered the LAB color model as compared to RGB color model. The LAB color model is considered to be device independent. Color difference calculated using CMC method are believed to correlate better with visual assessment than color differences calculated using other instrumental systems[1].Tomasi et al. (1998) proposed bilateral filtering for gray and color images. Obtained results shows that image blurs with the large value of domain parameter and little effect on quality of image with the varying range parameter. Also it does not produce phantom color along edges in the color images and it reduces phantom colors if they appear in the image [3].

From the literature we have concluded that image segmentation based on standard ACO is not efficient in no. of iterations on pixels and clusters are not fine. So we propose a hybrid algorithm which combines the properties of bilateral filtering and ant based clustering algorithm.

6. PROPOSED ALGORITHM

AntClust algorithm present in the literature survey is time consuming because of the redundancy in pixel accessibility. Due to sharp variations on the boundaries of objects in color images, objects are not clearly distributed to clusters by standard ant based clustering. In our proposed work we have introduced a number of slight modifications that reduce the repetitions by improved antClust algorithm and improve the quality of clustering by bilateral filter. In this research work, the fitness function is taken as CMC distance to find out the distance between two food sources i.e. colors. The CMC distance formula shows convincing results on its property to better characterize perceptual color similarity.

HybridAntClust

1. Take an image and convert it to a Lab image.
2. Initialize the cluster for the all data items with 0.
3. Initialize the cluster index with 1.

   Introduce an ant

For each pixels in the image.

Define a *Extract local region* **imin,imax,jmin,jmax** to form a *grid* based on given parameters

4. Calculate *bilateral filter response* by combining domain and range filtering. It replaces the pixel value at x with an average of similar and nearby pixel values
5. Initialize the ant by choosing a data item and assign the ant to its colony.
6. Assign the current cluster index
7. for each data item **ant_move** do
8. for each colony **ant** do
9. calculate similarity measure ‘**S’**
10. Take CMC Distance as threshold measure of similarity ‘**T’**
11. If S<T
12. Add the data item with the current cluster and increase the colony space
13. Move to the next neighbor.
14. Stop with current ant
15. Else if *check for generating new ant*
16. Introduce a new ant by increasing the ant index to form a new colony
17. Add data item to the colony of new ant.
18. Else
19. Continue
20. End if
21. End-for
22. End for

7. EXPERIMENTS AND RESULTS

Experiments are performed on two different images to perform the color pixel clustering. Four parameters are used to evaluate the results based on the properties of color science and spatial filtering. First and the most important parameter is the window size that define the locality in which spectral domain (σ_d) and range (σ_r) parameters performs filtering on various image pixels. The variation in the value of (σ_d) put little impact on the clustering as shown in the table (4.1).Because the no. of distant image locations that combines during Gaussian filter is directly proposional to (σ_d) where as the intensity component of the color effects more on the comparison between a sample and standard pixel value. To perform clustering on edges of object where the intensity changes sharply and more rapidly than other part of the image, the photometric spread (σ_r) in the image range is set to calculate the preferable amount of combination of pixel values. It is observed from the Table 4.1 and Table 4.2 that better clustering on edges of object in the image is achieved with varying values of (σ_r).Small change in (σ_r)give the drastic change in clustering which effect the clustering in bitter way also.

The proposed algorithm automatically generates and maintained the population of ants without prior knowledge. The no. of clusters varies with the CMC similarity measure taken as threshold value. The selected threshold value is chosen based on the heuristic approach which depends on the no. of colors in the image and the variation in the range of CIELab ellipsoid. As shown in the Table 4.1 MSE is least in the onion image at threshold value 15 and in lion image taking the threshold value 10 give better results. For MSE calculation, we have used Euclidian distance as a color measure.

For the verification of work, results of AntsClust algorithm have been compared with proposed algorithm.as shown in the Table 4.3 our proposed algorithm gives better results and also improves the standard algorithm by reducing number of iterations.

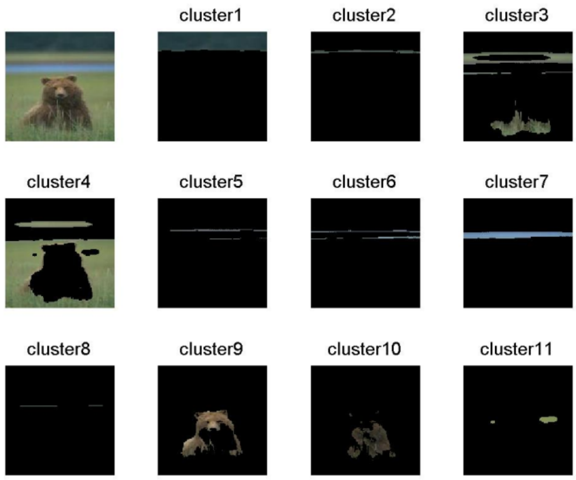

Figure 1: Lion Image and 11 clusters with window size 5, (σ_d) =5, $(\sigma_r) = 0.1$,T=9

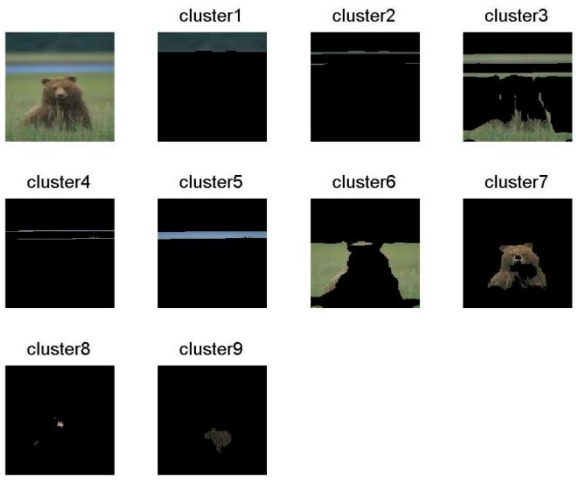

Figure 2: Lion and 9 clusters with window size 5, (σ_d) =5, $(\sigma_r) = 0.1$,T=10

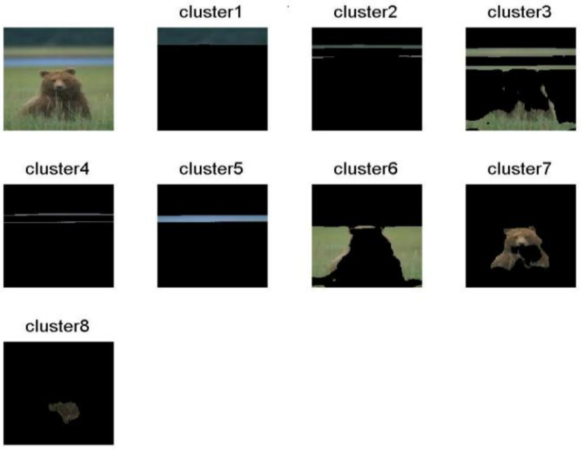

Figure 3: Lion and 8 clusters with window size 5, (σ_d) =5, $(\sigma_r) = 0.2$,T=10

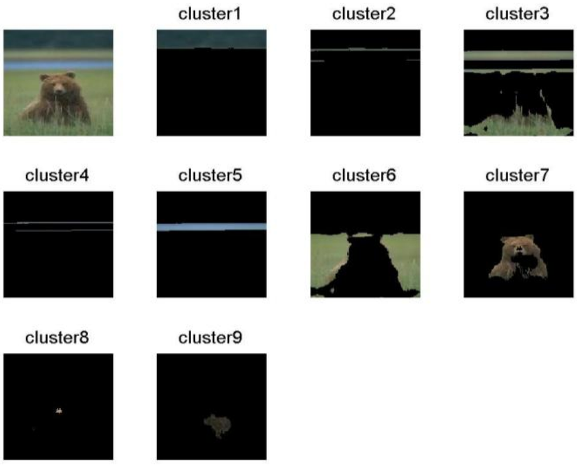

Figure 4: Lion and 9 clusters with window size 5, $(\sigma_d) = 10$, $(\sigma_r) = 0.1$, T=10

Window Size	Domain (σ_d)	Range (σ_r)	Treshold value T	MSE	No. of clusters
5	5	0.1	8	2.3761	15
5	5	0.2	8	2.3668	15
5	10	0.1	8	3.0814	15
5	5	0.1	9	2.1548	12
5	5	0.2	9	2.3215	12
5	10	0.1	9	2.3299	12
5	5	0.1	10	1.8993	9
5	5	0.2	10	1.9567	8
5	10	0.1	10	1.4387	9

Table: 4.1 Shows MSE and No.of Clusters of Onion.png image with with varying parameters

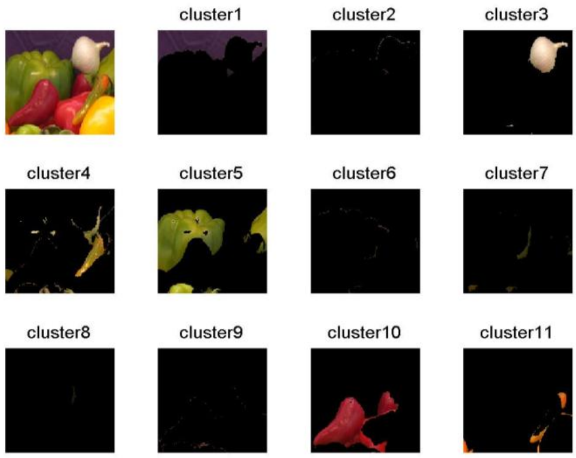

Figure 5: Onion and 16 clusters with window size 5, $(\sigma_d) = 5$, $(\sigma_r) = 0.1$, T=15,rest cluster are on second screen not displayed here.

Figure 6: Onion and 16 clusters with window size 5, $(\sigma_d) = 10$, $(\sigma_r) = 0.1$, T=15,rest cluster are on second screen not displayed here.

Window Size	Domain (σ_d)	Range (σ_r)	Treshold value **T**	MSE	No. of clusters
5	5	0.1	14	4.585	17
5	5	0.2	14	2.217	18
5	10	0.1	14	4.634	17
5	5	0.1	15	1.327	16
5	5	0.2	15	2.47	16
5	10	0.1	15	1.748	18
5	5	0.1	16	2.564	14
5	5	0.2	16	3.745	11
5	10	0.1	16	3.228	13
5	10	0.2	16	3.298	12

Table 4.2: Shows MSE and No.of Clusters of Lion.jpg image with with varying parameters

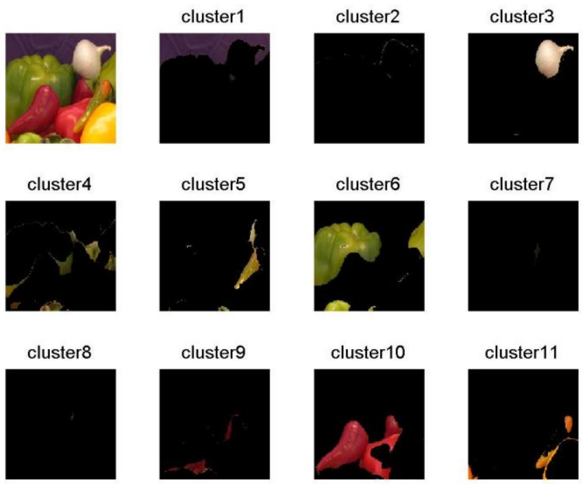

Figure 7: Onion and 16 clusters with window size 5, $(\sigma_d) = 5$, $(\sigma_r) = 0.2$, T=15,rest cluster are on second screen not displayed here.

Figure 8: Onion and 17 clusters with window size 5, $(\sigma_d) = 5$, $(\sigma_r) = 0.1$, T=14,rest cluster are on second screen not displayed here.

Table 4.3: Comparison of standard ant based clustering and Proposed algorithm on Onion.png Image.

Onion.png	**AntClust Algorithm**		**Proposed Algorithm**	
Threshold value (T)	14	15	14	15
No. of clusters	23	17	22	16
No. of Iterations	376833	360449	112161	105382
MSE	3.0176	2.9691	2.217	1.748

Table 4.4 Comparison of standard ant based clustering and Proposed algorithm on Lion.jpg Image.

Lion.jpg	**AntClust Algorithm**		**Proposed Algorithm**	
Threshold value (T)	9	10	9	10
No. of clusters	21	18	12	9
No. of Iterations	344065	294912	73143	69443
MSE	2.7454	2.6242	2.1548	1.8993

8. CONCLUSION

In this paper, an hybrid approach based on CIELab color space has been proposed where we have combined the ant based clustering with the bilateral filtering. Testing with only two images is given in this paper but the technique has been tested with images of varying size and format and the results found consistant. In addition from the results it is clear that range filter parameter (σ_r) helps to improve the clustering by preserving edges and domain parameter (σ_d) by smoothing pixels. Experimental results show that, with suitable value of CMC distance and other parameters, the proposed algorithm was able to successively solve the problem of slow speed and improve quality of clusters verified by calculate MSE using Euclidian distance. The parameters used for clustering in our illustrative examples were to some extent arbitrary.

As threshold value remain constant through the entire image in our algorithm. Further researchers also may contribute their research to vary threshold value based on locality of ant and can also try on different color space and similarity measure.

9. REFERENCES

[1] Applictipons Note"CMC" Vol. 8 No. 13

[2] Bhuvaneshwari,K.S.,"Adaptive clustering for image segmentation by updated k-means algorithm" 2011,ncrtes.

[3] Busin,L . and Shi,J. and Vandenbroucke,N. 2009 “Color Space Selection for Color Image Segmentation by Spectral Clustring”,2009 IEEE International Conference on Signal and Image Processing Applications

[4] Foley J. D., Van Dam A., Feiner S. K., Hughes J. F.,Computer Graphics - Principles and Practice, Second Edition, Addison- Wesley Publishing Company, 1990,ISBN 0-201- 12110-7.

[5] Ganesan,P. and Rajini,V. "Segmentation and Edge Detection of Color Images Using CIELAB Color Space and Edge Detectors" 2010 IEEE.

[6] Hunter l,a,b verses CIE (1976) L*a*b* (http://www.hunterlab.com/appnotes/an02_01.pdf) (PDF)

[7] Hunter labs (1996) “ Hunter lab color scale” Insight on color 8 9 (August 1-15, 1996).Reston,VA, USA: Hunter Associates Laboratories

[8] Immaculate M.C. and Kasmir,S.V. "A Modified Ant-based Clustering for Medical Data" 2010 IJCSE Vol. 02, No. 07, 2253-2257.

[9] Kaur,R. and Girdhar,A. Voloume 25-No.7,july 2011 "Color Image Quantization based on Bacteria Foraging Optimization"

[10] Laptik,R. and Navakauskas,D. "Applicationof AntColony OptimizationforImageSegmentation" 2007.

[11] Mohamed,O.A. "Ant-based Clustering Algorithms: A Brief Survey ", International Journal of Computer Theory
and Engineering, Vol. 2, No. 5, October, 2010, 1793-8201.

[12] Myung-Eun Lee1 et al., "Segmentation of Brain MR Images using an Ant Colony Optimization Algorithm", 978-0-7695-3656-9/09 $25.00 © 2009 IEEE.

[13] Ouadfel,S. and Batouche,M. , "An Efficient Ant Algorithm for Swarm-Based Image Clustering", Journal of Computer 3 (3): 162-167, 2007

[14] Pei-Yin,C. "An Efficient Edge-Preserving Algorithm for Removal of Salt-and-Pepper Noise" IEEE Signal processing letter, 2008, 15:833-836.

[15] Thakare,P. et al., "A Study of Image Segmentation and Edge Detection Techniques", International Journal on Computer Science and Engineering (IJCSE), ISSN : 9753397
Vol. 3 No. 2 Feb 2011.

[16] Tomasi,C. and Manduchi,R. "Bilateral Filtering for Gray and Color Images" Proceedings of the 1998 IEEE International Conference.