

Aggregation Pheromone Density Based Image Segmentation

Susmita Ghosh¹, Megha Kothari¹, and Ashish Ghosh^{2,*}

¹ Department of Computer Science and Engineering
Jadavpur University, Kolkata 700032, India

² Machine Intelligence Unit and Center for Soft Computing Research
Indian Statistical Institute
203 B. T. Road, Kolkata 700108, India
ash@isical.ac.in

Abstract. Ants, bees and other social insects deposit pheromone (a type of chemical) in order to communicate between the members of their community. Pheromone that causes clumping or clustering behavior in a species and brings individuals into a closer proximity is called aggregation pheromone. This paper presents a novel method for image segmentation considering the aggregation behavior of ants. Image segmentation is viewed as a clustering problem which aims to partition a given set of pixels into a number of homogenous clusters/segments. At each location of data point representing a pixel an ant is placed; and the ants are allowed to move in the search space to find out the points with higher pheromone density. The movement of an ant is governed by the amount of pheromone deposited at different points of the search space. More the deposited pheromone, more is the aggregation of ants. This leads to the formation of homogenous groups of data. The proposed algorithm is evaluated on a number of images using different cluster validity measures. Results are compared with those obtained using *average linkage* and *k-means* clustering algorithms and are found to be better.

1 Introduction

Image segmentation plays a vital role in image processing and computer vision problems. Number of image segmentation techniques exist in the literature, which can be grouped into several categories such as edge based, region based, histogram thresholding and clustering [1] based. Several image segmentation techniques exist [2,3,4] where clustering is used to group the points in the characteristics feature space into segments. Features that are commonly used for image segmentation by clustering not only include the gray values, but also include textural features defined on local neighborhood [5].

Numerous abilities of ants have inspired researches for designing various clustering techniques [6,7]. Several species of ants cluster their corpses into “cemeteries” in an effort to clean up their nests. Experimental works illustrate that ants

* Corresponding author.

group corpses, which are initially randomly distributed in space, into clusters within a few hours. It seems that some feedback mechanism (using local density or similarity of data items) determines the probability that an ant will pick up or drop a corpse. Such behavior is used as a model to design several algorithms for clustering data [6,7,8,9,10,11]. Besides nest cleaning, many functions of aggregation behavior have been observed in ants and ant like agents [12,13,14]. These include foraging-site marking and mating, finding shelter and defense. Tsutsui *et al.* [15,16] used aggregation pheromone systems for continuous function optimization where aggregation pheromone density is defined by a density function in the search space.

Inspired by the aforementioned aggregation behavior found in ants and ant like agents a clustering algorithm was proposed in [17]. In this paper the aggregation behavior of ants is used to perform image segmentation, viewing image segmentation as a clustering problem. The aim is to partition the image into clusters/segments such that pixels within a cluster are homogeneous, whereas pixels of different clusters are heterogeneous.

2 Aggregation Pheromone Based Image Segmentation

As mentioned in the introduction, aggregation pheromone brings individuals into closer proximity. This group formation nature of aggregation pheromone is being used as the basic idea of the proposed algorithm. Here each ant represents one data. The ants move with an aim to create homogenous groups of data. The amount of movement of an ant towards a point is governed by the intensity of aggregation pheromone deposited by all other ants at that point. This gradual movement of ants in due course of time will result in formation of groups or clusters. The proposed technique has two parts. In the first part, clusters are formed based on ants' property of depositing aggregation pheromone. The number of clusters thus formed might be more than the desired number. So, to obtain the desired number of clusters, in the second part, agglomerative *average linkage* clustering algorithm is applied on these already formed clusters.

2.1 Formation of Clusters

While performing image segmentation for a given image we group similar pixels together to form a set of coherent image regions. Similarity of the pixels can be measured based on intensity, color, texture and consistency of location of different pixels. Individual features or combination of them can be used to represent an image pixel. For each image pixel we associate a feature vector \mathbf{x} . Clustering is then performed on these set of feature vectors so as to group them. Finally, clustering result is mapped back to the original spatial domain to have segmented image.

Consider a data set of n patterns $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n)$ and a population of n -ants $(A_1, A_2, A_3, \dots, A_n)$ where an ant A_i represents the data pattern \mathbf{x}_i . Each individual ant emits aggregation pheromone in its neighborhood. The intensity

of aggregation pheromone emitted by an individual A , at \mathbf{x} decreases with its distance from it. Thus the pheromone intensity at a point closer to \mathbf{x} is more than those at other points that are farther from it. To achieve this, the pheromone intensity emitted by A is chosen to be a Gaussian function. The pheromone intensity deposited at \mathbf{x}' by an ant A (located at \mathbf{x}) is given by

$$\Delta\tau'(A, \mathbf{x}') = \exp^{-\frac{d(\mathbf{x}, \mathbf{x}')^2}{2\delta^2}}. \quad (1)$$

The total aggregation pheromone density deposited by the entire population of n ants at \mathbf{x}' is then given by

$$\Delta\tau(\mathbf{x}') = \sum_{i=1}^n \exp^{-\frac{d(\mathbf{x}_i, \mathbf{x}')^2}{2\delta^2}}, \quad (2)$$

where, δ denotes the spread of Gaussian function.

Now, an ant A' which is initially at location \mathbf{x}' moves to the new location \mathbf{x}'' (computed using Eq. 3) if the total aggregation pheromone density at \mathbf{x}'' is greater than that of \mathbf{x}' . The movement of an ant is governed by the amount of pheromone deposited at different points in the search space. It is defined as

$$\mathbf{x}'' = \mathbf{x}' + \eta \cdot \frac{Next(A')}{n}, \quad (3)$$

where,

$$Next(A') = \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}') \cdot \exp^{-\frac{d(\mathbf{x}_i, \mathbf{x}')^2}{2\delta^2}}, \quad (4)$$

with, η as a step size. This process of finding a new location continues until an ant meets a location where the total aggregation pheromone density is more than its neighboring points. Once the ant A_i finds out such a point \mathbf{x}'_i with greater pheromone density, then that point is assumed to be a new potential cluster center, say \mathbf{Z}_j ($j = 1, 2, \dots, C$, C being number of clusters); and the data point with which the ant was associated earlier (*i.e.*, \mathbf{x}_i) is now assigned to the cluster so formed with center \mathbf{Z}_j . Also the data points that are within a distance $\delta/2$ from \mathbf{Z}_j are assigned to the newly formed cluster. On the other hand, if the distance between \mathbf{x}'_i and the existing cluster center \mathbf{Z}_j is less than $\delta/2$ and the ratio of their densities is greater than *threshold_density* (a predefined parameter), then the data point \mathbf{x}_i is allocated to the already existing cluster centering at \mathbf{Z}_j . Higher value of density ratio shows that the two points are of nearly similar density and hence should belong to the same cluster. The proposed algorithm for formation of clusters (APC) is given below using pseudo codes.

```

begin
Initialize  $\sigma$ , threshold_density,  $\eta$ 
 $C = 0$ 
for  $i = 1$  to  $n$  do
if the data pattern  $\mathbf{x}_i$  is not already assigned to any cluster

```

```

Compute  $\Delta\tau(\mathbf{x}_i)$  using Eq. 2
label 1: Compute new location  $\mathbf{x}'_i$  using Eq. 3
Compute  $\Delta\tau(\mathbf{x}'_i)$ 
if ( $\Delta\tau(\mathbf{x}'_i) > \Delta\tau(\mathbf{x}_i)$ )
Update the location of ant  $A_i$  to  $\mathbf{x}'_i$  and goto label 1
else continue
if ( $C == 0$ ) //If no cluster exists
Consider  $\mathbf{x}'_i$  as cluster center  $\mathbf{Z}_1$  and increase  $C$  by one
else
for  $j = 1$  to  $C$ 
if ( $\min(\Delta\tau(\mathbf{x}'_i), \Delta\tau(\mathbf{Z}_j)) / \max(\Delta\tau(\mathbf{x}'_i), \Delta\tau(\mathbf{Z}_j)) > threshold\_density$  and
 $d(\mathbf{x}'_i, \mathbf{Z}_j) < \delta/2$ )
Assign  $\mathbf{x}'_i$  to  $\mathbf{Z}_j$ 
else
Assign  $\mathbf{x}'_i$  as a new cluster center say,  $\mathbf{Z}_{C+1}$  and increase  $C$  by one
Assign all the data points that are at a distance of  $\delta/2$  from  $\mathbf{x}'_i$  to
the newly formed cluster  $\mathbf{Z}_{C+1}$ 
end of else
end of for
end of else
end of if (if the data pattern  $\mathbf{x}_i \dots$ )
end of for
end

```

2.2 Merging of Clusters

Number of clusters formed by the above algorithm may be more than the desired number. In this stage, to obtain the desired number of clusters we apply *average linkage* [18] algorithm. The algorithm starts by merging the two most similar clusters until the desired number of clusters are obtained.

3 Objective Evaluation of Clustering Results

The clustering results obtained by the above described method are quantified using following cluster validity measures.

- *Davies Bouldin Index*: This index is a function of the ratio of the sum of within-cluster scatter to between-cluster separation [18]. The average scatter of order q within the i^{th} cluster ($S_{i, q}$) and the distance between i^{th} and j^{th} clusters ($d_{ij, q}$) are computed as

$$S_{i, q} = \left(\frac{1}{|C_i|} \sum_{x \in C_i} \|x - \mathbf{Z}_i\|^q \right)^{\frac{1}{q}}, \quad (5)$$

$$d_{ij, q} = \|\mathbf{Z}_i - \mathbf{Z}_j\|_q. \quad (6)$$

where \mathbf{x} is the data point belonging to cluster C_i , \mathbf{Z}_i is the centroid of cluster C_i , $q \geq 1$ and $d_{ij, q}$ is the Minkowski distance of order q . Subsequently, we compute the index for the i^{th} cluster as

$$R_{i, q} = \max_{j, i \neq j} \left\{ \frac{S_{i, q} + S_{j, q}}{d_{ij, q}} \right\}. \quad (7)$$

The *Davies-Bouldin (DB)* index for C clusters is defined as

$$DB = \frac{1}{C} \sum_{i=1}^C R_{i, q}. \quad (8)$$

The smaller the DB value, better is the clustering.

- *S_Dbw* : *S_Dbw* index with C number of clusters is based on the clusters' compactness in terms of intra-cluster variance and inter-cluster density [19]. It is defined as

$$S_Dbw(C) = Scat(C) + Den(C), \quad (9)$$

where *Scat(C)* represents the intra-cluster variance and is defined as

$$Scat(C) = \frac{1}{C} \sum_{i=1}^C \|\sigma(\mathbf{Z}_i)\| / \sigma(s); \quad (10)$$

the term $\sigma(s)$ is the variance of the data set and $\sigma(\mathbf{Z}_i)$ is the variance of cluster C_i . Inter-cluster density, *Den(C)*, is defined as

$$Den(C) = \frac{1}{C-1} \sum_{i=1}^C \left(\sum_{i=1, i \neq j}^C \frac{den(\mathbf{u}_{ij})}{\max\{den(\mathbf{Z}_i), den(\mathbf{Z}_j)\}} \right) \quad (11)$$

where \mathbf{Z}_i and \mathbf{Z}_j are centers of clusters C_i and C_j , respectively and \mathbf{u}_{ij} is the mid point of the line segment joining \mathbf{Z}_i and \mathbf{Z}_j . The term *den(u)* is defined as

$$den(\mathbf{u}) = \sum_{i=1}^{n_{ij}} f(\mathbf{x}, \mathbf{u}), \quad (12)$$

where n_{ij} is the total number of data points belong to clusters C_i and C_j ; the function $f(\mathbf{x}, \mathbf{u})$ is defined as

$$f(\mathbf{x}, \mathbf{u}) = \begin{cases} 0, & \text{if } d(\mathbf{x}, \mathbf{u}) > stdev; \\ 1, & \text{otherwise.} \end{cases} \quad (13)$$

where *stdev* is the average standard deviation of C clusters and is defined as

$$stdev = \frac{1}{C} \sqrt{\sum_{i=1}^C \|\sigma(\mathbf{Z}_i)\|} \quad (14)$$

and $d(\mathbf{x}, \mathbf{u})$ is the Euclidean distance between \mathbf{x} and \mathbf{u} .

Lower the value of S_Dbw, better is the clustering.

Table 1. Comparison of results of the proposed aggregation pheromone based algorithm (APC) with *average linkage* and *k-means*

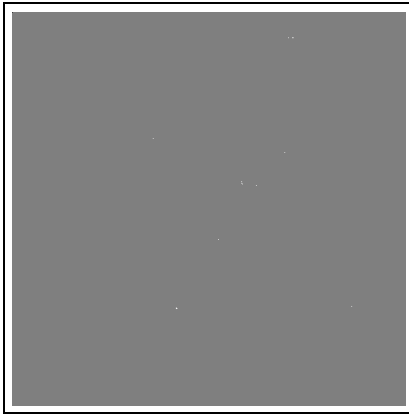
Data set	APC			Average linkage			k-means		
	DB	S_Dbw	Time	DB	S_Dbw	Time	DB	S_Dbw	Time
Tank	0.3026 (0.032)	0.1576 (0.024)	439.2 (4.87)	0.3887 (4.9E-09)	0.2389 (0)	435.4 (11.80)	0.6337 (0.002)	1.0216 (0.056)	470.4 (5.14)
Lena	0.5330 (0.018)	0.5508 (0.029)	506.3 (8.13)	0.4814 (7.0E-09)	0.5560 (1.4E-08)	602.6 (8.26)	0.5556 (0.002)	0.6411 (0.034)	528 (9.84)
Brain	0.4307 (0.078)	0.1567 (0.044)	43.7 (9.91)	0.4797 (0)	0.1798 (0)	134 (7.74)	0.5089 (0.005)	0.2122 (0.003)	53.2 (8.01)



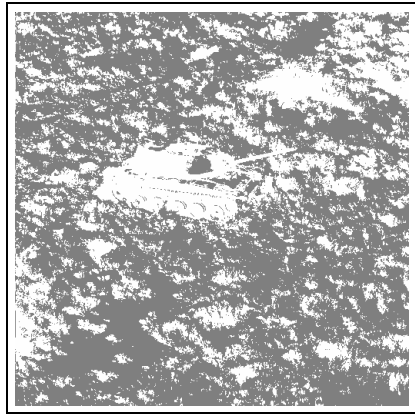
(a)



(b)



(c)



(d)

Fig. 1. a) original image, b) APC segmentation result, c) Average link result and d) k-means result with k=2

4 Experimental Results

4.1 Results

To evaluate the effectiveness of the proposed algorithm, we have considered only gray value of pixels as a feature. Experiments were carried out on three images (Tank, Lena and Brain). Values of η and *threshold_density* were kept to 1 and 0.9, respectively; and different values of δ in the range $[0, 1]$ were considered. Results obtained are validated using two different cluster validity indices as described in Section 3. The results obtained by the proposed *APC* algorithm are compared with those of *average linkage* and *k-means* clustering algorithms. Table 1 gives the mean values (over 10 runs) of different performance indices and their corresponding standard deviations (shown in bracket) for each of the images obtained by *APC*, *average linkage* and *k-means* algorithms. The CPU



Fig. 2. a) original image, b) *APC* segmentation result, c) *Average link* result and d) *k-means* result with $k=3$

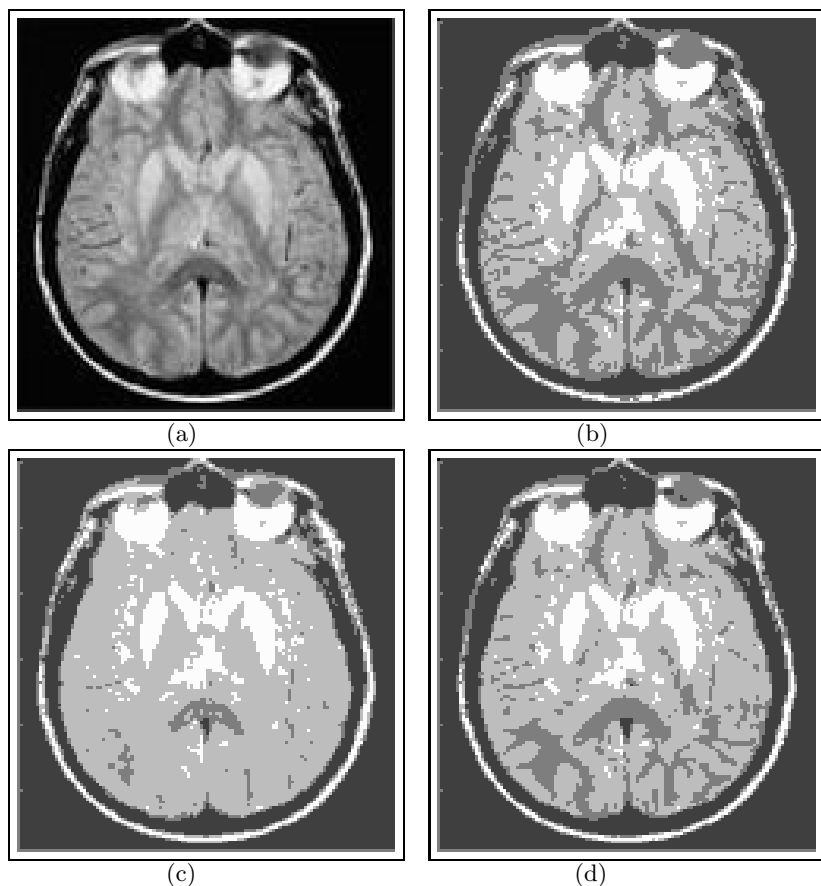


Fig. 3. a) original image, b) *APC* segmentation result, c) *Average link* result and d) *k-means* result with $k=4$

time (in milliseconds) needed for all the algorithms are also given in the table for comparison. Results in bold face indicate the best ones.

Fig. 1 shows the segmentation result obtained on the Tank image by all the three algorithms. As may be seen, *average linkage* algorithm fails completely to segregate the tank (Fig. 1c); *k-means* algorithm segmented out the tank but the amount of false classification was very high; but the proposed *APC* algorithm successfully segmented out the tank with very less false classification (Fig. 1b). In terms of validity measures also the performance of the proposed *APC* algorithm is found to be better (Table 1) than those of other two.

Original Lena image is shown in Fig. 2a and the segmentation results obtained for the Lena image by *APC*, *average linkage* and *k-means* are shown in Figs. 2b-2d. From the results one can see that the background is not clearly separated out in case of *average linkage*, whereas with the *APC* and *k-means* it is well

done. The shade on the face of Lena is well segmented with the proposed *APC* algorithm which is not obtained with that of *k-means*. Also, the top left part of the background is well segmented by the *APC* algorithm in comparison with the other two. When validity measures are taken into consideration it is found that with *DB* index *average linkage* algorithm performs well on Lena image but with *S_Dbw* index *APC* performs better.

Another image that is considered for our experiments is Brain image. The segmentation results obtained by all the three algorithms for this image are shown in Figs. 3b-3d. Original Brain image is shown in Fig. 3a. From the results one can see that the fine structures in the Brain image are very well segmented by the proposed *APC* algorithm; *average linkage* fails to identify these fine structures and *k-means* identifies these fine structures but with higher misclassification. In terms of validity measures also *APC* outperforms the other two.

As evident from the table, the time requirement of the proposed algorithm is mostly less compared to the other two algorithms.

5 Conclusions

In this paper we have proposed a new algorithm for image segmentation based on aggregation pheromone density, which is inspired by the ants' property to accumulate around points with higher pheromone density. Experiments were carried out with three different images to evaluate the performance of the proposed algorithm both qualitatively as well as quantitatively. In this paper segmentation is viewed as a clustering problem and hence for comparative evaluation we have used cluster validity measures; and comparative study is made with two clustering algorithms namely, *average linkage* and *k-means* algorithms. Future study should involve more number of features and object extraction from noisy images.

Acknowledgements

Authors would like to acknowledge the Department of Science and Technology, Government of India and University of Trento, Italy, the sponsors of the India-Trento Program on Advanced Research (ITPAR), under which a project titled “**Advanced Techniques for Remote Sensing Image Processing**” is being carried out at Department of Computer Science and Engineering, Jadavpur University, Kolkata.

References

1. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Second edn. Pearson education (2003)
2. Pal, N., Pal, S.: A review on image segmentation techniques. Pattern Recognition **26(9)** (1993) 1277–1294
3. Kettaf, F.Z., Bi, D., de Beauville, J.P.A.: A comparison study of image segmentation by clustering techniques. In: Proceedings of ICSP'96. (1996) 1280–1283

4. Saatchi, S., Hung, C.C.: Hybridization of the ant colony optimization with the k-means algorithm for clustering. In Kalviainen, H., ed.: SCIA 2005. Volume 3540 of LNCS., Springer-Verlag, Berlin Heidelberg (2005) 511–520
5. Chanda, B., Majumder, D.D.: Digital Image Processing and Image Analysis. Prentice Hall of India, New Delhi (2003)
6. Deneubourg, J.L., Goss, S., Franks, N., Sendova-Franks, A., Detrain, C., Chretien, L.: The dynamics of collective sorting: Robot-like ants and ant-like robots. In Meyer, J.A., Wilson, S.W., eds.: Proceedings of the 1st Conference on Simulation of Adaptive Behavior: From Animals to Animats 1, MIT press/Bradford Books (1991) 356–365
7. Lumer, E.D., Faieta, B.: Diversity and adaptation in populations of clustering ants. In Cliff, D., Husbands, P., Meyer, J.A., Wilson, S.W., eds.: Proceedings of the 3rd International Conference on Simulation of Adaptive Behaviour: From Animals to Animats 3. (1994) 501–508
8. Monmarché, N., Slimane, M., Venturini, G.: On improving clustering in numerical database with artificial ants. In Floreano, D., Nicoud, J.D., Mondala, F., eds.: Advances in Artificial Life, 5th European Conference ECAL'99, Lecture Notes in Artificial Intelligence 1974, Swiss Federal Institute of Technology, Lausanne, Switzerland, Springer-Verlag (1999) 626–635
9. Handl, J., Knowles, J., Dorigo, M.: On the performance of ant-based clustering. In: Proceedings of the 3rd International Conference on Hybrid Intelligent Systems, Design and Application of Hybrid Intelligent Systems, IOS press (2003) 204–213
10. Liu, S., Dou, Z.T., Li, F., Huang, Y.L.: A new ant colony clustering algorithm based on DBSCAN. In: Proceedings of the 3rd International Conference on Machine Learning and Cybernetics, Shanghai (2004) 1491–1496
11. Vizine, A.L., de Castro, L.N., Hruschka, E.R., Gudwin, R.R.: Towards improving clustering ants: an adaptive ant clustering algorithm. *Informatica* **29** (2005) 143–154
12. Bell, W.J.: Chemo-orientation in walking insects. In Bell, W.J., Carde, R.T., eds.: *Chemical Ecology of Insects*. (1984) 93–109
13. Ono, M., Igarashi, T., Ohno, E., Sasaki, M.: Unusual thermal defence by a honeybee against mass attack by hornets. *Nature* **377** (1995) 334–336
14. Sukama, M., Fukami, H.: Aggregation arrestant pheromone of the German cockroach, *Blattella germanica* (L.) (Dictyoptera: Blattellidae): isolation and structure elucidation of blasttellanostanoid-A and B. *Journal of Chemical Ecology* **19** (1993) 2521–2541
15. Tsutsui, S.: Ant colony optimization for continuous domains with aggregation pheromones metaphor. In: Proceedings of the 5th International Conference on Recent Advances in Soft Computing (RASC'04), United Kingdom (2004) 207–212
16. Tsutsui, S., Ghosh, A.: An extension of ant colony optimization for function optimization. In: Proceedings of the 5th Asia Pacific Conference on Simulated Evolution and Learning (SEAL04), Pusan, Korea (2004)
17. Kothari, M., Ghosh, S., Ghosh, A.: Aggregation pheromone density based clustering. In: Proceedings of 9th International conference on Information Technology, Bhubaneswar, India, IEEE Computer Society press (2006)
18. Theodoridis, S., Koutroumbas, K.: *Pattern Recognition*. Second edn. Elsevier Academic Press, Amsterdam (2003)
19. Halkidi, M., Vazirgiannis, M.: Clustering validity assessment: finding the optimal partitioning of a data set. In: Proceedings of ICDM, California, USA (2001)