Applying Swarm Algorithms on Unsupervised Color Image Classification

Khalid Mohammed Saffer **Samahir Khalil Ismail**
Collage of Science University of Diyala *Collage of Science University of Baghdad*
kha2005ms@yahoo.com

Abstract

Clustering algorithms are used as a solution for different image problem, such as pre-processing, segmentation, feature extraction, dimensionality reduction, data visualization, and final classification. The one that proposed in this paper is the classification, so this work takes two clustering algorithm (k-mean & pso) then applying it on an unsupervised color images (satellite image, medicine image, natural image).

Keywords:- Particle Swarm Optimization(PSO), K-mean ,Unsupervised Image, Compactness, Image Classification.

الخلاصة

خوارزميات التجميع (Clustering algorithms) تستخدم لحل مختلف مشاكل الصور مثل التصنيف قبل المعالجة (pre-processing) ، التجزئة (segmentation) ، استخراج او استخلاص الميزات (feature extraction)، الحد من الابعاد (dimensionality reduction)، اعادة عرض البيانات(data visualization) واخيرا التصنيف(classification) . في هذا البحث تم التطرق لموضوع التصنيف حيث تم اخذ اثنين من خوارزميات التجميع (K-mean & PSO) وطبقت على مجموعة من الصور الملونة غير خاضعة للرقابة (Unsupervised) في مختلف المجالات منها صور اقمار صناعية وصور طبية وصور للطبيعة .

الكلمات المفتاحية:صورة غير خاضعة للرقابة ، تصنيف الصور

1- Introduction

Image clustering is the process of identifying groups of similar image primitives. These image primitives can be pixels, regions, line elements and so on, depending on the problem encountered. Many basic image processing techniques such as quantization, segmentation and coarsening can be viewed as different instances of the clustering problem [Puzicha, Hofmann and .Buhmann,.*at al* 2000].

There are two main approaches to image classification:

Supervised and unsupervised. In the supervised approach, the number and the numerical characteristics (e.g. mean and variance) of the classes in the image are known in advance (by the analyst) and used in the training step which is followed by the classification step. There are several popular supervised algorithms such as the minimum-distance-to-mean, parallelepiped and the Gaussian maximum likelihood classifiers. In the unsupervised approach the classes are unknown and the approach starts by partitioning the image data into groups (or clusters), according to a similarity measure, which can be compared with reference to data by an analyst [Lillesand and .Kiefer,*at al* 1994].

Therefore, unsupervised classification is also referred to as a clustering problem. In general, the unsupervised approach has several advantages over the supervised approach, namely [Davies,1997]

- For unsupervised approaches, there is no need for an analyst to specify in advance all the classes in the image data set. The clustering algorithm automatically finds distinct classes, which dramatically reduces the work of the analyst.
- The characteristics of the objects being classified can vary with time; the unsupervised approach is an excellent way to monitor these changes.
- Some characteristics of objects may not be known in advance. The unsupervised approach automatically flags these characteristics.

The main objective of using a clustering algorithm is to combine data into small groups such that the data in each group possess similar characteristics though the data clusters are distinct from each other. The most widely used methods to classify the structure of data is the K-means algorithm. It is an unsupervised clustering technique which has a strong inclination towards the local minima while finding an optimal solution. Therefore, the distribution of initial cluster centers drastically decides the process of clustering. Therefore, the determination of the good initial parameters is a challenging problem and hence the clustering algorithms necessitate more number of experimentation to decide the input parameters for the optimal or suboptimal clustering results [Murugesan and Palaniswami.2012].

2- Objectives

The objective of this paper is to classify unsupervised color images of the satellite image and medical Images by Appling K-mean and PSO algorithms. These two approaches determine the "optimum" number of clusters by Run the algorithms repetitively using different input values and Select the partitioning of data resulting in the best validity measure. One criterion that has been widely considered sufficient in measuring the quality of partitioning a data set into a number of clusters is:

Compactness: samples in one cluster should be similar to Each other and different from samples in other clusters.

.3- K-Means Clustering Algorithm

The K-means algorithm is an unsupervised clustering algorithm which partitions a set of data, usually termed dataset into a certain number of clusters. Minimization of a performance index is the primary basis of K-means Algorithm, which is defined as the sum of the squared distances from all points in a cluster domain in the cluster center. Initially K random cluster centers are chosen. Then, each sample in the sample space is assigned to a cluster based on the minimum distance to the center of the cluster. Finally the cluster centers are updated to the average of the values in each cluster. This is repeated until cluster centers no longer change. Steps in the K-means algorithm are [K.M. Murugesan and S. Palaniswami.2012]:

Step 1: Initialize K initial cluster centers randomly.

Step 2: For each pixel, calculate the distance to the cluster centers and assign the pixel to a cluster which has the minimum distance to its center.

Step 3: Calculate the average of the pixel values in each cluster and take them as new cluster centers.

Step 4: Repeat steps 2 and 3 until new cluster centers converge to the previous ones.

The prime motive of the K-means Algorithm is to find the local minima rather than the global minima. Therefore, the algorithm severely sticks to the initial choice of cluster centers and distribution of data. The results become acceptable most of the time when the initial cluster centers are selected relatively far away. It is because; the main clusters in a given data are commonly differentiated in such a manner. If the main clusters in a given data are too close to one another in the sample space, the K-means algorithm fails to identify these clusters. For its improvement the K-means algorithm needs to be enhanced with some optimization technique in order to be less reliant on a given dataset and initial cluster centers.

4- Particle swarm optimization Algorithm (PSO)

Particle swarm optimizers (PSO) are population-based optimization algorithms modeled after the simulation of social behavior of bird flocks [Kennedy, Eberhart,1995], [Kennedy, Eberhart,2001]. In a PSO system, a swarm of individuals (called *particles*) fly through the search space. Each particle represents a candidate solution to the optimization problem. The position of a particle is influenced by the best position visited by itself (i.e. its own experience) and the position of the best particle in its neighborhood (i.e. the experience of neighboring particles).

When the neighborhood of a particle is the entire swarm, the best position in the neighborhood is referred to as the global best particle, and the resulting algorithm is referred to as a *gbest* PSO. When smaller neighborhoods are used, the algorithm is generally referred to as a *lbest* PSO [Y. Shi, R. Eberhart,1998].

The performance of each particle (i.e. how close the particle is from the global optimum) is measured using a fitness function that varies depending on the optimization problem.
Each particle in the swarm is represented by the following characteristics:
xi : The current position of the particle;
vi : The current velocity of the particle;
yi : The personal best position of the particle.
For each iteration of a PSO algorithm, the velocity ***vi*** update step is specified for each dimension j = 1.. *Nd*, where
Nd is the dimension of the problem. Hence, *vij* represents the j^{th} element of the velocity vector of the i^{th} particle. Thus the velocity of particle *i* is updated as using the following equation:

$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1 r_{1,j}(t)(y_{i,j}(t) - x_{i,j}(t)) + c_2 r_{2,j}(t)(\hat{y}_j(t) - x_{i,j}(t)) \qquad (1)$$

Where *w* is the inertia weight [Y. Shi, R. Eberhart,1998], C_1 and C_2 are the acceleration constants and $r_{1j} \in [0,1]$ and $r_{2j} \in [0,1]$.

The position of particle *i*, ***xi***, is then updated using the following equation:

$$x_i(t+1) = x_i(t) + v_i(t+1) \qquad (2)$$

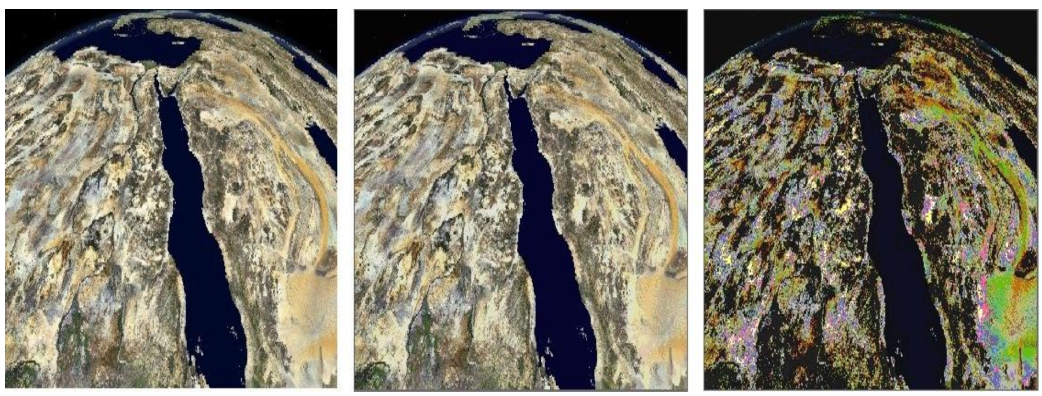
The algorithm can be summarized as follow [Mohd Afizi Mohd Shukran, Yuk Ying Chung, Wei-Chang Yeh, Noorhaniza Wahid and Ahmad Mujahid Ahmad Zaidi,2011]:
1) Initialise: Initialise parameters and population with random position and velocities.
2) Evaluation: Evaluate the fitness value (the desired objective function) for each particle.
3) Find the gbest: If the fitness value of particle i is better than its best fitness value (pbest) in history, then set current fitness value as the new pbest to particle i.
4) Find the gbest: If any pbest is updated and it is better than the current gbest, then set gbest to the current value.
5) Update position: update velocity for each particle by applying equation (1) and (2).

5- Experiments

After run PSO & K-Mean algorithms on several images that captured from different resource (satellite image, medicine image, natural image).the result show in table(1) below :

Table (1) : Classification results

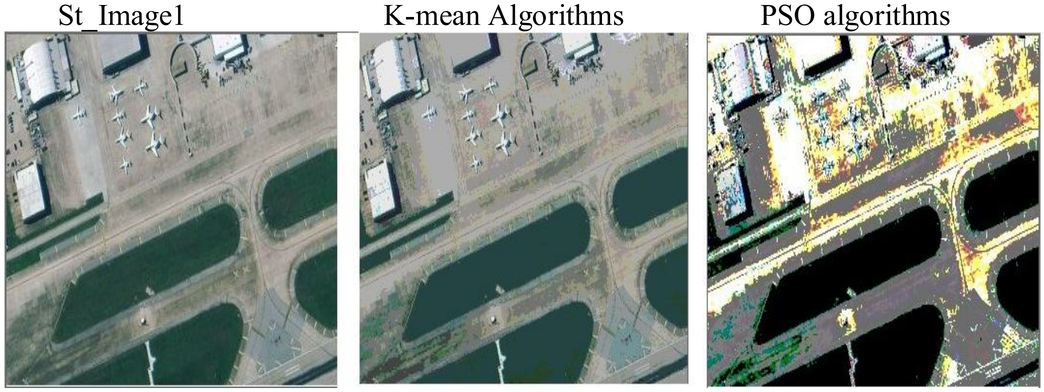
No.	Image Name	Swarm Algorithms	Fitness	
			Start Val.	End Val.
1	St_Image1	K-Mean	15.98309	25.83193
		PSO	20.81855	30.40322
2	St_Image2	K-Mean	19.51335	25.01647
		PSO	14.32291	17.11959
3	Me_Image3	K-Mean	15.06489	20.13685
		PSO	14.32291	20.07587
4	Me_Image4	K-Mean	15.16673	24.80697
		PSO	20.81855	30.40322
5	Na_Image5	K-Mean	16.01869	27.90292
		PSO	14.32291	20.07587
6	Na_Image6	K-Mean	24.94252	39.29664
		PSO	20.81855	39.05147



a- Original Image St_Image1

b- Processing Image with K-mean Algorithms

c- Processing Image with PSO algorithms



d- Original Image St_Image2

e- Processing Image with K-mean Algorithms

f- Processing Image with PSO algorithms

Figure (1): Satellite Image classified by PSO & K-mean algorithms

a- Original Image Me_Image3 | b- Processing Image with K-mean Algorithms | c- Processing Image with PSO algorithms

d- Original Image Me_Image4 | e- Processing Image with K-mean Algorithms | f- Processing Image with PSO algorithms

Figure (2): Medicine Image classified by PSO & K-mean algorithms

a- Original Image Na_Image5 | b- Processing Image with K-mean Algorithms | c- Processing Image with PSO algorithms

d- Original Image Na_Image6 | e- Processing Image with K-mean Algorithms | f- Processing Image with PSO algorithms

Figure (3): Natural Image classified by PSO & K-mean algorithms

6- Conclusions

- Results obtained recorded varies depending on the algorithm used and also depending on the type of the captured image.
- Overall the performance of PSO algorithm was better than the performance of k-mean algorithm because PSO produced a clear Isolate to the colors of the images by the specific areas.
- Fluctuation in fitness values recorded from rise to fall and then rise, but eventually settled on good values.
- If color areas were reduced to the image recorded fluctuation values the results were more stable

References

Davies, E. Machine Vision: 1997. Theory, Algorithms, Practicalities, Academic Press, 2nd Edition,

Kennedy, J. R. Eberhart, 1995. Particle Swarm Optimization, Proceedings of IEEE International Conference on Neural Networks, Perth, Australia, vol. 4, 1942-1948,

Kennedy, J. R. Eberhart, 2001. Swarm Intelligence, Morgan Kaufmann,

Puzicha, J. T. Hofmann and J. M. Buhmann, 2000 Histogram Clustering for Unsupervised Image Segmentation, IEEE Proceedings of the Computer Vision and Pattern Recognition 2, 602-608

Murugesan K.M. and S. Palaniswami. 2012. HYBRID EXPONENTIAL PARTICLE SWARM OPTIMIZATION K-MEANS ALGORITHM FOR EFFICIENT IMAGE SEGMENTATION,

Mohd Afizi Mohd Shukran, Yuk Ying Chung, Wei-Chang Yeh, Noorhaniza Wahid and Ahmad Mujahid Ahmad Zaidi. 2011. Image Classification Technique using Modified Particle Swarm Optimization,

Lillesand T. and R. Kiefer, 1994. Remote Sensing and Image Interpretation, John Wiley & Sons Publishing, N.Y.,

Shi, Y. R. Eberhart, 1998.Parameter Selection in Particle Swarm Optimization, Evolutionary Programming VII: Proceedings of EP 98, 591-600,

Shi, Y. R. Eberhart, 1998. A Modified Particle Swarm Optimizer, Proceedings of the IEEE International Conference on Evolutionary Computation, Piscataway, NJ, 69-73,