Multilevel Quantum Inspired Fractional Order Ant Colony Optimization for Automatic Clustering of Hyperspectral Images

Siddhartha Bhattacharyya
Department of Computer Science & Engineering
CHRIST (Deemed to be University),
Bangalore, Karnataka, India
dr.siddhartha.bhattacharyya@gmail.com

Tulika Dutta
Department of Computer Science & Engineering
University Institute of Technology,
Burdwan University, West Bengal, India
munai.tulika@gmail.com

Sandip Dey
Department of Computer Science
Sukanta Mahavidyalaya
Dhupguri, Jalpaiguri, West Bengal, India
dr.ssandip.dey@gmail.com

***Abstract*—Hyperspectral images contain a wide variety of information, varying from relatively large regions to smaller man-made buildings, roads and others. Automatic clustering of various regions in such images is a tedious task. A multilevel quantum inspired fractional order ant colony optimization algorithm is proposed in this paper for automatic clustering of hyperspectral images. Application of fractional order pheromone updation technique in the proposed algorithm produces more accurate results. Moreover, the quantum inspired version of the algorithm produces results faster than its classical counterpart. A new band fusion technique, applying principal component analysis and adaptive subspace decomposition, is successfully proposed for the pre-processing of hyperspectral images. Score Function is used as the fitness function and K-Harmonic Means is used to determine the clusters. The proposed algorithm is implemented on the Xuzhou HYSPEX dataset and compared with classical Ant Colony Optimization and fractional order Ant Colony Optimization algorithms. Furthermore, the performance of each method is validated by peak signal-to-noise ratio which clearly indicates better segmentation in the proposed algorithm. The Kruskal-Wallis test is also conducted along with box plot, which establishes that the proposed algorithm performs better when compared with other algorithms.**

***Index Terms*—Hyperspectral Image Segmentation, Peak signal-to-noise ratio, Fractional Order Ant Colony Optimization, Qutrits, Kruskal-Wallis test**

I. Introduction

A Hyperspectral Image (HSI) is a data cube of images collected over narrow spectral channels. The spectral channels are commonly referred to as *bands*, which may vary from 10 to above 400 depending upon the sensors used to capture them. Massive information content is the prime feature of a HSI. HSIs are rich in information and are extensively used in agriculture, food processing, biomedical imaging [1] and others. As the images are collected over contiguous spectral channels, highly correlated and redundant data is a common problem for HSIs. This curse of dimensionality is called the Hughes phenomenon [2]. To reduce the burden of dealing with highly correlated bands, various dimensionality reduction techniques are proposed in literature. These techniques are classified as feature extraction [3] and feature selection techniques [4]. Dimensionality reduction is done keeping in mind that subsequent procedures like classification, segmentation and other processes are not compromised.

The process of dividing an image into relevant parts based on common similar properties is called image segmentation. Similarity can be based on color, texture or intensity. Clustering is a popular method used for image segmentation. Generally clustering is referred to as segregating unlabeled data into groups of objects with similarity. These homogenous groups are called *"clusters"*. Mathematically, clustering can be explained as

$$\begin{aligned} &I = \bigcup_{t=1}^{p} Clust_t \\ &\text{where, } Clust \subset I, Clust = \{Clust_1, Clust_2, \ldots, Clust_p\} \\ &\text{and } Clust_i \cap Clust_j = \{\} \text{ for } i \neq j \end{aligned} \tag{1}$$

Here, *I* is the data provided for segmentation, *Clust* are the individual clusters which are non-intersecting in nature. Finding the number of clusters for different problems remains a very challenging task. Automatic cluster detection has attracted the attention of a lot of researchers lately. K-means [5], Fuzzy C Means [6] are few well known clustering methods. A Cluster Validity Index (CVI) is used to determine the quality of clustering. This is determined using two properties - the *compactness* within the cluster elements and the *separation* between different cluster elements. The most popularly used CVIs are Calinski - Harabasz Index [7], Xie-Beni Index [8], I-Index [9] to name a few.

Exhaustive search methods are often adopted for problems where no efficient solution finding method is known. Metaheuristics are often implemented for finding solutions to exhaustive search problems. Nature inspired metaheuristics have recently gained a lot of popularity. The effectiveness of the nature inspired algorithms in finding near optimal solutions in lesser time has made them popular among the researchers. Metaheuristics are very effective, but as per *No*

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

Free Lunch [10], no metaheuristic is best suited for solving all problems. Few well known nature inspired metaheuristic algorithms are Genetic Algorithm (GA) [11], Ant Colony Optimization (ACO) [12], Particle Swarm Optimization (PSO) [13], Grey Wolf Optimizer [14] and others.

In recent years, researchers have shown considerable interest in quantum computing as traditional computing models are reaching their limits. By applying the quantum phenomena like superposition and entanglement, quantum computers can exponentially accelerate the rate of information processing [15]. Quantum inspired metaheuristics have thus become a wide area of research, as they provide better results to exhaustive search problems. A few bi-level quantum algorithms can be found in [15][16].

The objectives of the proposed work are presented below.

- To reduce the computational burden of dealing with numerous bands of a hyperspectral image, a band amalgamation technique is introduced. This is done by combining Principal Component Analysis (PCA) and Adaptive Subspace Decomposition (ASD) techniques to select three bands from the hyperspectral image cube.
- Fractional order calculus is capable of preserving the memory of past events. The pheromones of artificial ants are updated in the proposed algorithm by applying the fractional order derivative. This helps the algorithm to converge faster and achieve near optimal values.
- The superposed form of *qutrits* used in the proposed algorithm, helps to achieve optimal results in lesser time. This also helps to explore the search space exponentially when compared to its traditional counterpart.
- The determination of optimal number of clusters for hyperspectral images is a major problem. The foremost objective of this paper is to develop a multilevel quantum inspired fractional order Ant Colony Optimization (MQi-FOACO) algorithm to optimize the number of clusters of HSI datasets automatically.

The layout of the paper is as follows, Section II contains a brief survey of the related works. Important background concepts are explained briefly in Section III. In Section IV, the proposed methodology is discussed in details. Experimental results and analysis are presented in Section V. Section VI draws the conclusion of the paper.

II. Literature Survey

HSIs are obtained over a long distance with the aid of satellite or airborne sensors [17]. Sensors used for capturing HSIs are capable of capturing large number of spectral channels simultaneously. The images are captured over the same target area. Thus dealing with a hyperspectral image is like processing hundreds of 2D images simultaneously. An abundance of information is provided by the narrow spectral channels. This abundance of data, in spite of being beneficial, is cursed with increased computational cost, decreased efficiency and redundancy [2]. Extensive research is carried out on dimensionality reduction techniques in HSI analysis. They are classified as feature extraction and feature selection techniques. Principal Component Analysis (PCA) [3] is a well known feature extraction technique. Feature selection techniques can be further classified as supervised and unsupervised techniques. Information Gain Approach [18], Adaptive Subspace Decomposition [4] are few well known feature selection techniques.

The inspiration behind the invention of the Ant Colony Optimization algorithm is the pheromone trail laying behavior of biological ants. Ant System (AS) [19] is one of the earliest proposed algorithms, which is based on the behavior of ants. Later on, Dorigo *et al.* [12] proposed the ACO algorithm, which proved to be far more efficient than AS.

Fractional Order (FO) differential equations are widely used as an effective tool in scientific research nowadays [20]. They provide far more accurate results due to their long term memory [21]. In [22], FO is introduced in metaheuristic approaches for the first time. Fractional order PSO (FOPSO) is further studied in [23] by Couceiro *et al*. FOPSO is further implemented on gray scale images, multi-spectral images and HSIs [24][25], medical image segmentation [26] and others. Fractional order PID controller using ACO is implemented in [27].

Sir Richard Feynman for the first time suggested the concept of quantum computing [28]. Faster execution speed and capability of exploring the search space with fewer individuals in an efficient manner, have paved the way for the development of quantum inspired metaheuristics [29]. A Quantum Inspired Evolutionary Algorithm is proposed in [30]. In [31], the authors proposed the quantum inspired versions of GA, PSO and ACO by employing the concepts of quantum rotation gates and orthogonality principle. A quantum inspired ACO for automatic clustering is proposed in [32]. Tkachuk introduced a *qutrit* based GA in [33] with a unique Quantum Disaster Operation.

In image segmentation, an image is subdivided or partitioned into homogenous and separate regions. Image segmentation can be broadly classified into texture analysis methods, histogram based thresholding methods, clustering and region based split and merge methods [34]. Clustering can be further classified into fuzzy clustering (K-Means [5]) and hard clustering (fuzzy c-means [6]). The above mentioned algorithms are popular due to their simplicity. Though they are very famous and widely used, need for a prior knowledge of the clusters makes them very vulnerable. To overcome the shortcomings of the K-means, the K Harmonic Means (KHM) algorithm is proposed in [35]. The KHM algorithm is further studied and combined with different validity indices in [36].

Cluster validation is a technique employed to evaluate the quality of clustering. CVI usually measures the compactness within the objects of a cluster and the separation between different clusters [32]. Calinski - Harabasz Index [7], Xie-Beni Index [8], I-Index [9] are few widely used indices.

III. Important Concepts

Few important background concepts related to the proposed work are presented below.

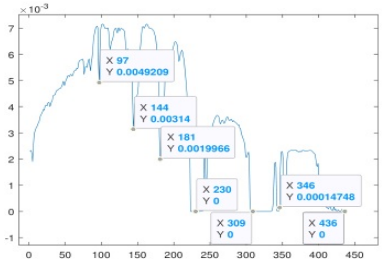

Fig. 1. Correlation Coefficient Graph of Bands

A. Hyperspectral Image Band Reduction

A hyperspectral image data cube contains a lot of information captured from narrow spectral channels. The information content is thus mostly redundant in nature. To overcome this problem, different techniques are used to reduce the number of bands for efficient segmentation. Few techniques are discussed below.

1) Adaptive Subspace Decomposition Technique (ASD): This is a feature selection technique. In ASD, it is assumed that the bands with similar information should be grouped together [4]. Let an image be of $N \times M$ spatial dimensions and S number of spectral bands. Then the correlation coefficient (C_{ij}) between bands i and j is expressed as follows

$$C_{ij} = \frac{\sum_{k=1}^{N\times M} \left(Sp_{ik} - \bar{Sp}_i\right)\left(Sp_{jk} - \bar{Sp}_j\right)}{\sqrt{\sum_{k=1}^{N\times M} \left(Sp_{ik} - \bar{Sp}_i\right)^2 \sum_{k=1}^{N\times M} \left(Sp_{jk} - \bar{Sp}_j\right)^2}} \tag{2}$$

Here, each image is represented using Sp, consisting of ($N \times M$) pixels as $Sp_i = [Sp_{i1}, Sp_{i2}, Sp_{i3}, ..., Sp_{i(N\times M)}]$. Sp_{ik} represents the k^{th} pixel in the i^{th} spectral band. $\bar{Sp}_i$ and $\bar{Sp}_j$ represent the average of the i^{th} and j^{th} bands. In Fig. 1, we see the number of groups in which the spectral bands can be divided into by plotting the C_{ij} values for the Xuzhou HYSPEX dataset [37][38]. Every group contains redundant information.

Once the bands are grouped, Shannon's Entropy is used to find the information content in each band as [39]

$$E\left(Sp_i\right) = -\int_{Sp_i} p\left(Sp_i\right) \log p\left(Sp_i\right) dSp_i \tag{3}$$

The information content in each group is calculated using the following formula.

$$I = \frac{1}{Gr}\left(\sum_{i=1}^{Gr} E\left(Sp_i\right)\right) \tag{4}$$

where Gr is the number of bands in a group. In Fig. 1, the value of Gr in first group is 97. From each group the bands with the highest I value are chosen. This helps in choosing the most informative bands and thus reduces the redundant information.

2) Principal Component Analysis (PCA): The main aim of PCA is to convert data into lower dimensions [3]. In the process, it tries to retain as much relevant information as possible. A new set of variables, orthogonal to each other, called the principal components are obtained. The first component usually contains the largest variation which, however, keeps on decreasing with consecutive components [40].

B. Basic Principles of Quantum Computing

The basic unit of a quantum computer is a *qubit*. It exists in *ground* and *excited* states of $|0\rangle$ and $|1\rangle$. Dirac's bra and ket notations are popularly used to denote these states [41]. The main difference between a qubit and a bit is that it exhibits the properties of superposition and entanglement. Qubits are usually represented using the following column vector.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5}$$

Quantum superposition means that the qubit can exist in a linear combination of both $|0\rangle$ and $|1\rangle$ states. If a superposition state is denoted by ψ, then the superposition equation will be given as

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle \tag{6}$$

Here, α_0 and α_1, are the probabilities of finding the quantum system in the states $|0\rangle$ and $|1\rangle$, respectively. From the basic theory of probability, it can be stated that the certainty of a possible event is always 1 and that of an impossible event's occurrence is 0. As it is certain that the qubit will always be found in either $|0\rangle$ or $|1\rangle$ state, so the normalized form of a qubit state can be mathematically represented as

$$\alpha_0^2 + \alpha_1^2 = 1 \tag{7}$$

Quantum systems are not bound by two states like traditional bits. They can exist in $|0\rangle$, $|1\rangle$, $|2\rangle, \cdots, |n\rangle$ number of states. The basic unit of a multivalued quantum bit is called a *qudit*. The simplest of this many valued quantum logic is the ternary quantum logic. The basic unit is called the *qutrit*. The vector representation of a *qutrit* is given by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{8}$$

The superposition of a *qutrit* can be explained with the following equation.

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle \tag{9}$$

with a normalization constraint of

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 = 1 \tag{10}$$

A classical system with n bits can represent 2^n different numbers and only one out of these numbers can be stored at any point of time. In comparison to this, a *qubit* can exist in all of the 2^n states simultaneously in superposed form. Similarly, a *qutrit* and *qudit*, can simultaneously exist in 3^n states and n^n states, respectively. Hence, in quantum domain, switching from bi-valued to multi-valued logic, increases the search capability in an exponential manner.

C. Ant Colony Optimization

Initially the ACO algorithm was developed to deal with the Travelling Salesman Problem [19]. The inspiration for the ACO algorithm is the pheromone trail laying behavior of real ants. In a simple ant colony, every ant while travelling to the food source lays pheromone in its path, which evaporates eventually. As time passes, more and more ants follow the path with more pheromone concentration.
The steps of the ACO algorithm [12], are represented by Algorithm 1. In Algorithm 1, the pheromone matrix is updated

Result: $fit_{(Global\_Best)}$
Initialize : Pheromone Matrix(τ) and Artificial Ant Population(pop) ;
for *t in 1,2,...,Iter* **do**
  fit_t=Fitness of all ants at t^{th} iteration;
  Update Pheromone Matrix (τ);
  Update ant positions using pheromone matrix (pop);
  if $fit_{(t\_B)} > fit_{(G\_B)}$ **then**
    $fit_{(G\_B)} = fit_{(t\_B)}$;
  else
  end
end

Algorithm 1: Ant Colony Optimization

using the following equation.

$$\tau_{t+1} = \rho_1 \tau_t + pop_t(1 - \rho_1) \tag{11}$$

The fitness of the individual ants is depicted by fit_t for the t^{th} iteration, $fit_{(t\_B)}$ being the best fitness of t^{th} iteration and $fit_{(G\_B)}$ being the global best fitness of all iterations.

D. Fractional order derivatives

The idea of fractional calculus first emerged from a conversation between L'Hopital and Liebniz in 1695 [21]. The n^{th} derivative of the linear function, $f(x) = x$, $\frac{D^n x}{Dx^n}$ was used by Liebniz, which leads to the question that if $n = \frac{1}{2}$, what will be the result. This led to a number of famous mathematicians like Euler, Fourier, Laplace working on it. The most famous derivatives in fractional calculus are Riemann-Liouville and Grunwald-Letnikov.
The idea of fractional differentiation is defined in terms of the Grunwald-Letnikov derivative [22], for fractional coefficient α and a general signal $x(t)$ as

$$D^{\alpha}[x(t)] = \lim_{h \to 0} \left[\frac{1}{h^{\alpha}} \sum_{k=0}^{+\infty} \frac{(-1)^k \Gamma(\alpha + 1) x(t - kh)}{\Gamma(k+1)\Gamma(\alpha - k + 1)} \right] \tag{12}$$

Here, it is seen that the fractional order derivatives contain infinite terms and preserve memory of all the past events. This is not found in integer order derivatives.
When implemented in discrete time, the above equation is approximated as follows

$$D^{\alpha}[x(t)] = \frac{1}{T^{\alpha}} \sum_{k=0}^{r} \frac{(-1)^k \Gamma(\alpha + 1) x(t - kT)}{\Gamma(k+1)\Gamma(\alpha - k + 1)} \tag{13}$$

Equation (13) contains T as the sampling period and r as the truncation order.

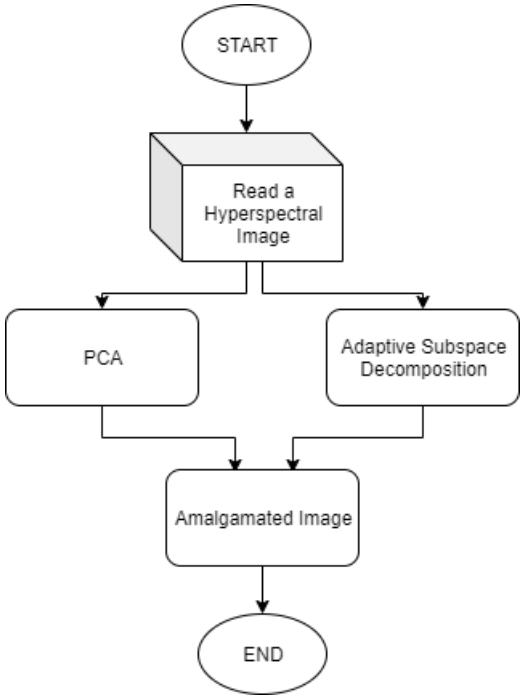

Fig. 2. Hyperspectral Image Fusion Technique

IV. Proposed Methodology

The proposed methodology can be divided into three broad divisions, (i) Hyperspectral image band amalgamation, (ii) the fractional order approach for implementing ant colony optimization and (iii) the multilevel quantum inspired ant colony optimization algorithm for automatic clustering of HSIs.

A. Band Amalgamation Technique

A new approach is implemented using PCA and ASD methods inspired from the approach in [42]. Equation (4) is used to select two bands with the highest values. These bands are fused with the first principal component to obtain a fused image. It is an amalgamation of both feature extraction and feature selection techniques. The fused image thus obtained for segmentation purpose carries a lot of relevant information. As only three bands are used for the purpose of segmentation, the complexity of dealing with huge number of band decreases. The pictorial representation of the proposed technique is given in Fig. 2.

B. ACO with fractional pheromone updation

A new method for updation of the pheromone matrix is introduced by applying the Grunwald - Letnikov definition of fractional order derivatives. The main advantage of this method is that it leads to smooth variation and a long memory effect [22]. In ACO, only the last pheromone trail is used for updation whereas when fractional order pheromone updation is done, the pheromone trail can be chosen to be impacted by any number of previous pheromone trails. This memory preserving property of the fractional order system produces optimized

results, as it is a better tool to describe the irreversible and chaotic behavior of the ant's trajectory path.
Equation (11) can be rearranged as

$$\tau_{t+1} - \tau_t = (1-\rho_1)(pop_t - \tau_t) \tag{14}$$

After taking the Grunwald-Letnikov derivative of the left side, we get.

$$D^{\alpha}[\tau_{t+1}] = (1-\rho_1)(pop_t - \tau_t) \tag{15}$$

Equation (13) contains the discrete time implementation of $D^{\alpha}[\tau_{t+1}]$. When this is substituted in the above equation, the following equation is obtained.

$$\frac{1}{T^{\alpha}} \sum_{k=0}^{r} \frac{(-1)^k \Gamma(\alpha+1) x(t-kT)}{\Gamma(k+1)\Gamma(\alpha-k+1)} = (1-\rho_1)(pop_t - \tau_t) \tag{16}$$

Here, taking the truncation order (r) as 4 [23], leads to

$$\begin{aligned} &\tau_{t+1} - \alpha\tau_t - \tfrac{1}{2}\alpha\tau_{t-1} - \tfrac{1}{6}\alpha(1-\alpha)\tau_{t-2} \\ &-\tfrac{1}{24}\alpha(1-\alpha)(2-\alpha)\tau_{t-3} = (1-\rho_1)(pop_t - \tau_t) \end{aligned} \tag{17}$$

or,

$$\begin{aligned} &\tau_{t+1} = \alpha\tau_t + \tfrac{1}{2}\alpha\tau_{t-1} + \tfrac{1}{6}\alpha(1-\alpha)\tau_{t-2} \\ &+\tfrac{1}{24}\alpha(1-\alpha)(2-\alpha)\tau_{t-3} + (1-\rho_1)(pop_t - \tau_t) \end{aligned} \tag{18}$$

Equation (18) is then used for the updation of the pheromone matrix. Only first four r values are taken, as for $r > 4$ the algorithm presents similar results with increased computational costs. Equation (14) is the discrete version of the derivative of order $\alpha = 1$ of (15). The fractional order derivative used in the updation of pheromone matrix, leads to faster convergence of the system.

C. MQi-FOACO

In the proposed algorithm, a *qutrit* inspired population is considered. As the initial ant population does not contain any memory, so the quantum population is initialized with equal values. The quantum orthogonality principle is maintained in the process as

$$|qpop\rangle = \frac{1}{\sqrt{3}}|0\rangle + \frac{1}{\sqrt{3}}|1\rangle + \frac{1}{\sqrt{3}}|2\rangle \tag{19}$$

Equation (19) is used from [33] for the quantum population initialization and its classical interpretation. The main objective of the algorithm is to optimize the number of clusters automatically. For this purpose, random number of 0's are introduced in each row of the classical interpretation matrix. The number of non-zero values is considered as the number of clusters. In every step, the value of quantum population, for which the classical value is obtained as non-zero, is considered as cluster center. The quantum population is updated based on the pheromone values. The distance of each pixel (P_i) to the cluster centers is calculated using the following equation [36].

$$mem\left(\mathrm{cl}_j/P_i\right) = \frac{\|P_i - \mathrm{cl}_j\|^{-s-2}}{\sum_{j=1}^{k} \|P_i - \mathrm{cl}_j\|^{-s-2}}, \quad mem\left(\mathrm{cl}_j/P_i\right) \in [0,1] \tag{20}$$

Pixel (P_i) is assigned to the cluster which gives the highest value for the above equation.

The Score Function (SF) [43] is used as the fitness function. The definition of the SF index is given by

$$SF = 1 - \frac{1}{e^{e^{B-\mathrm{W}}}} \tag{21}$$

Here,

$$B = \frac{\sum_{i=1}^{k} \|\ \mathrm{cl}_{\ i} - \ \mathrm{cl}_{\ m}\| \cdot n_i}{n \cdot k} \tag{22}$$

and

$$W = \sum_{i=1}^{k} \left(\frac{1}{n_i} \sum_{P_{\mathrm{cl}\ i}} \|P_i - \ \mathrm{cl}_{\ i}\| \right) \tag{23}$$

Here, pixels are denoted by P_i, cl_i are the cluster centers and the centroid of all the clusters is denoted by cl_m. The aggregate number of pixels in the image is n and n_i is the total number of pixels in i^{th} cluster. The total number of clusters is denoted by k.
The steps of the MQi-FOACO are explained with the help of Algorithm 2. In Algorithm 2, the fitness of the ants are calculated using (21). The pheromone matrix at iteration

Result: $Optimal\ Number\ of\ Clusters$
Intialise: Number of Generation - $Iter$,
Size Artificial Ant Population - N,
Length of each ant -L,
Pheromone Matrix with random values
between (0,1) -$\tau_{t+1}^{\alpha_0}$,$\tau_t^{\alpha_0}$,$\tau_{t-1}^{\alpha_0}$,$\tau_{t-2}^{\alpha_0}$,$\tau_{t-3}^{\alpha_0}$,
$\tau_{t+1}^{\alpha_1}$,$\tau_t^{\alpha_1}$,$\tau_{t-1}^{\alpha_1}$,$\tau_{t-2}^{\alpha_1}$,$\tau_{t-3}^{\alpha_1}$,$\tau_{t+1}^{\alpha_2}$,$\tau_t^{\alpha_2}$,$\tau_{t-1}^{\alpha_2}$,$\tau_{t-2}^{\alpha_2}$,$\tau_{t-3}^{\alpha_2}$,
Priory defined number - q;
Intialise quantum popualtion $|qpop\rangle$ using (19)
Classical Interpretation of Quantum Population:
for *i in 1,2,...,N* **do**
  for *j in 1,2,...,L* **do**
    $r = random\ number\ between\ (0,1)$
    if $r < qpop_{i,j}^{\alpha_0 2}$ **then**
      $pop_{i,j} = 0$
    end
    else if $r < qpop_{i,j}^{\alpha_0 2} + qpop_{i,j}^{\alpha_1 2}$ **then**
      $pop_{i,j} = 1$
    end
    else
      $pop_{i,j} = 2$
    end
  end
end
for *t in 1,2,...,Iter* **do**
  fit_t=Fitness of all ants at t^{th} iteration;
  Update Pheromone Matrix (τ);
  $ta1 = max(\tau_t^{\alpha_0})$,$tb1 = max(\tau_t^{\alpha_1})$,$tc1 = max(\tau_t^{\alpha_2})$;
  for *i in 1,2,...,N* **do**
    for *j in 1,2,...,L* **do**
      $r1 = random\ number\ between\ (0,1)$
      if $pop_{i,j} == 0$ **then**
        if $r1 > q$ **then**
          $qpop_{i,j}^{\alpha_0} = ta1$
        end
        else
          $qpop_{i,j}^{\alpha_0} = random\ number(0,1)$
        end
      end
    end
  end
  if $fit_{(t\_Best)} > fit_{(Global\_Best)}$ **then**
    $fit_{(Global\_Best)} = fit_{(t\_best)}$;
  end
  $\tau_t^{\alpha_0} = \tau_{t+1}^{\alpha_0}$, $\tau_{t-1}^{\alpha_0} = \tau_t^{\alpha_0}$, $\tau_{t-2}^{\alpha_0} = \tau_{t-1}^{\alpha_0}$, $\tau_{t-3}^{\alpha_0} = \tau_{t-2}^{\alpha_0}$
  $\tau_t^{\alpha_1} = \tau_{t+1}^{\alpha_1}$, $\tau_{t-1}^{\alpha_1} = \tau_t^{\alpha_1}$, $\tau_{t-2}^{\alpha_1} = \tau_{t-1}^{\alpha_1}$, $\tau_{t-3}^{\alpha_1} = \tau_{t-2}^{\alpha_1}$
  $\tau_t^{\alpha_2} = \tau_{t+1}^{\alpha_2}$, $\tau_{t-1}^{\alpha_2} = \tau_t^{\alpha_2}$, $\tau_{t-2}^{\alpha_2} = \tau_{t-1}^{\alpha_2}$, $\tau_{t-3}^{\alpha_2} = \tau_{t-2}^{\alpha_2}$
end

Algorithm 2: MQi-FOACO

t is updated using (18). The quantum population updation procedure with respect to the pheromone matrix is given only for the classical value of 0, similarly for the classical population values, of 1 and 2, the corresponding quantum population is updated. The other two quantum population

Result: $Updated\ Quantum\ Population$
for *k in 1,2,...,N* **do**
  for *t in 1,2,...,L* **do**
    if $pop_{k,t} = 0$ **then**

$$qpop_{k,t}^{\alpha 1} = \sqrt{\frac{\left(1 - qpop_{k,t}^{\alpha_0\,2}\right)}{2}} - \text{rand}(< 0.1) \quad (24)$$

$$qpop_{k,t}^{\alpha 2} = \sqrt{1 - \left(qpop_{k,t}^{\alpha_0\,2} + qpop_{k,t}^{\alpha 1\,2}\right)} \quad (25)$$

    end
  end
end

Algorithm 3: Quantum Population Updation

values are then updated using Algorithm 3.
To avoid the quantum population from getting stuck in the local optima, a part of the quantum population is reinitialized using (19). The ants with less fitness values are chosen for this process. This method is called the Quantum Disaster Operation [33].

D. Complexity Analysis

The HSI amalgamation technique consists of reading S bands, each of $N \times M$ pixels and performing PCA and ASD with them. The worst case time complexity for image amalgamation technique is $O(N \times M \times S)$ for each method. The total time complexity is therefore $O(2 \times N \times M \times S)$, as both the processes are run simultaneously.
The worst case time complexity for the MQi-FOACO depends on the size of the ant population and the total number of iterations. Since the *qutrit* version is considered, the ant population of $N \times L$ length consists of three orthogonal states and a classical representation state. So to process the population at every iteration, time required is $O(4 \times N \times L)$. If P_i is the total number of pixels in an image, then at every step ($Iter$), every pixel is read to assign the clusters. Then the total time complexity becomes $O(4 \times N \times L \times P_i \times Iter)$.

V. Results

A brief description of the dataset is given in Section V-A. The favorable results, statistical analysis test results along with the segmented images are presented in Section V-B.

A. Dataset

The proposed MQi-FOACO was tested on the Xuzhou HYSPEX dataset [37][38]. The Xuzhou peri-urban site was captured in November, 2014 by a HYSPEX camera. The dataset has a resolution of 0.73 m/pixel. Each image has a spatial resolution of 500×260 pixels and a spectral dimension of 436 bands. The scene can be divided into nine parts according to the classification image. This consists of different crops, structures like buildings, roads and coal fields primarily.

TABLE I
Mean, Standard Deviation (STD) and PSNR values for ACO, FO-ACO and MQi-FOACO

Sr No	Process	Mean	STD
1	ACO	0.7869	0.0054
2	FO-ACO	0.7877	0.0052
3	MQi-FOACO	**0.7931**	**0.0018**

TABLE II
Five favorable outcomes for ACO, FO-ACO and MQi-FOACO

Process	ACO		FO-ACO		MQi-FOACO	
Sr No	Cl No	FV	Cl No	FV	Cl No	FV
1	4	0.783697	5	0.78176	10	0.791614
2	4	0.782617	5	0.787306	10	0.791819
3	4	0.780481	5	0.783496	11	0.792865
4	4	0.786124	6	**0.787309**	10	0.79289
5	4	**0.790547**	6	0.783886	10	**0.793778**

B. Analysis

In (18), the value of α greatly affects the pheromone deposition of the ants. The value of alpha ranges from 0 to 1. A small value of α makes the previous pheromone deposition values negligible. This may result in the population getting stuck in local optima. Simultaneously for a high α value, the convergence of the algorithm takes more iterations. Thus, the determination of the α value plays a crucial role.
The Score Function [43] is used as the CVI for the determination of the optimal number of clusters along with the KHM algorithm [36]. The value of SF index varies between $[0, 1]$, where a value nearer to 1 indicates an optimal number of clusters. The MQi-FOACO is compared with the classical ACO [12] and the fractional order ACO (FOACO) [24][25]. The classification image is divided into 9 parts, as depicted in Fig. 4. It is found that the proposed MQi-FOACO automatically generates 10 or 11 clusters on every occasion. Hence, it can be said that the proposed MQi-FOACO algorithm is found to yield better results than others. In Table I, the means and standard deviations for all the three processes are recorded. Few fitness values (FV) are recorded in Table II. From Table II, we can observe that the number of clusters obtained in MQi-FOACO is almost same to the number of segments mentioned in the classification image.
In Table III, the best time and peak signal-to-noise ratio (PSNR) [44] values are recorded for all the three processes. The PSNR [44] value of each process is computed by comparing the clustered image to the ground truth image to determine the quality of segmentation. PSNR [44] is found to be higher in case of MQi-FOACO.
A statistical analysis test called the Kruskal-Wallis test [45],

TABLE III
Best Time and PSNR values for ACO, FO-ACO and MQi-FOACO

SR No	Process	Best Time	PSNR
1	ACO	2045.994	4.62433
2	FOACO	444.0172	5.542635
3	MQi-FOACO	**29.61429**	**7.100478**

TABLE IV
KRUSKAL WALLIS TEST

Test	p-value	Significance
Kruskal Wallis Test	6.6243e-05	Highly Significant

was performed on the optimal results obtained from ACO, FOACO and MQi-FOACO and is recorded in Table IV. The result is carried out with 1% significance level. This test is applied to find the p value. The p values less than 0.05 represent significant and less than < 0.001 represent highly significant results. The hypothesis that σ values have the same distribution across all three methods, thus stands rejected. This is also the null hypothesis. The Box Plot for the Kruskal-Wallis test is given in Fig. 3. The true color image and the

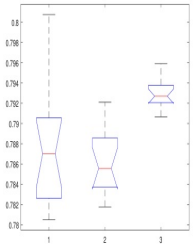

Fig. 3. Box-plot for Kruskal-Wallis Test

classification image for the Xuzhou HYSPEX dataset are given in Fig. 4. Figs. 5 and 6 contain the fused images obtained after using ASD and PCA and the segmented images obtained after applying ACO, FOACO and MQi-FOACO. From the convergence curve shown in Fig. 7 it is observed that the fitness value obtained by the proposed method is the best and it reaches to an optimum result faster than other algorithms. The algorithm was carried out in MATLAB R2019a, on an Intel (R) Core (TM) i7 8700 Processor with Windows 10 environment.

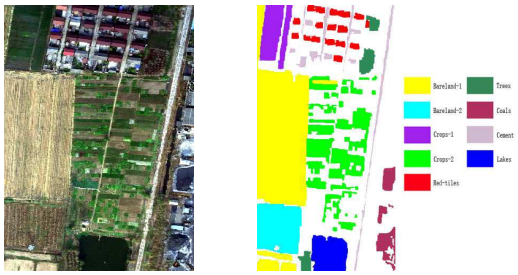

Fig. 4. True Color Image and Classification Image

VI. CONCLUSION

A multilevel quantum inspired FOACO for automatic clustering of HSIs has been proposed in this paper. The fractional order pheromone updation technique produces better results as long term memory of the pheromone trail is preserved. The *qutrit* based application speeds up the process. The

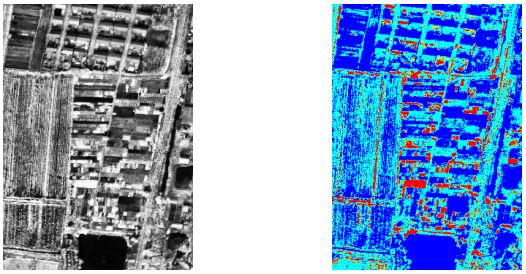

Fig. 5. Fused Image and ACO ($k = 4$)

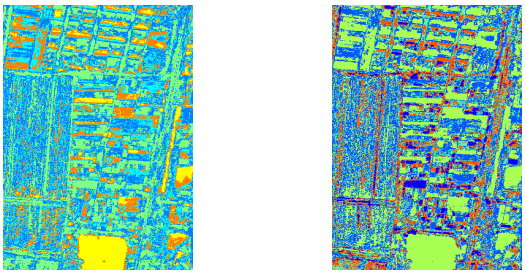

Fig. 6. FOACO ($k = 5$) and MQi-FOACO ($k = 11$)

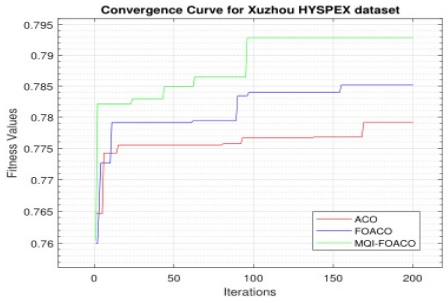

Fig. 7. Convergence Curve for Xuzhou HYSPEX dataset

band fusion technique by using ASD and PCA in the pre-processing stage reduces the computational burden. Moreover, the automatic cluster detection technique comes handy, since knowing the number of clusters for a HSI is very difficult. The superiority of the segmented images obtained by the proposed algorithm has been established by applying PSNR method and comparing it with results obtained from classical ACO and FOACO. A statistical superiority test has also been applied on all the three methods to establish the superiority of the proposed algorithm.

ACKNOWLEDGEMENT

This work was supported by the AICTE sponsored RPS project on Automatic Clustering of Satellite Imagery using Quantum-Inspired Metaheuristics vide F. No. 8-42/RIFD/RPS/Policy-1/2017-18.

REFERENCES

[1] P. J. C. Weijtmans, C. Shan, T. Tan, S. G. B. de Koning, and T. J. M. Ruers, "A dual stream network for tumor detection in hyperspectral images," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, April 2019, pp. 1256–1259.

[2] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Information Theory*, vol. 14, pp. 55–63, 1968.

[3] C. Rodarmel and J. Shan, "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Systems*, vol. 62, 01 2002.

[4] Y. Zhang, M. D. Desai, J. Zhang, and M. Jin, "Adaptive subspace decomposition for hyperspectral data dimensionality reduction," in *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*, vol. 2, Oct 1999, pp. 326–329 vol.2.

[5] J. Macqueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 1967, pp. 281–297.

[6] J. C. Bezdek, R. Ehrlich, and W. Full, "Fcm: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2, pp. 191 – 203, 1984.

[7] T. Caliński and H. JA, "A dendrite method for cluster analysis," *Communications in Statistics - Theory and Methods*, vol. 3, pp. 1–27, 01 1974.

[8] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 841–847, Aug 1991.

[9] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1650–1654, 01 2002.

[10] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, April 1997.

[11] J. H. Holland, *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor : University of Michigan Press, 1975.

[12] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29–41, Feb 1996.

[13] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)*, May 1998, pp. 69–73.

[14] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46 – 61, 2014.

[15] T. Atkinson, A. Karsa, J. Drake, and J. Swan, "Quantum program synthesis: Swarm algorithms and benchmarks," in *Genetic Programming*, Cham, 2019, pp. 19–34.

[16] S. Dey, S. Bhattacharyya, and U. Maulik, "Efficient quantum inspired meta-heuristics for multi-level true colour image thresholding," *Applied Soft Computing*, vol. 56, pp. 472–513, July 2017.

[17] S. Li, J. Qiu, X. Yang, H. Liu, D. Wan, and Y. Zhu, "A novel approach to hyperspectral band selection based on spectral shape similarity analysis and fast branch and bound search," *Engineering Applications of Artificial Intelligence*, vol. 27, pp. 241 – 250, 2014.

[18] A. Elmaizi, H. Nhaila, E. Sarhrouni, A. Hammouch, and C. Nacir, "A novel information gain based approach for classification and dimensionality reduction of hyperspectral images," *Procedia Computer Science*, vol. 148, pp. 126 – 134, 2019.

[19] M. Dorigo, "Optimization, learning and natural algorithms," *PhD Thesis, Politecnico di Milano*, 1992. [Online]. Available: https://ci.nii.ac.jp/naid/10000136323/en/

[20] J. T. Machado, A. M. Galhano, and J. J. Trujillo, "Science metrics on fractional calculus development since 1966," *Fractional Calculus and Applied Analysis*, vol. 16, pp. 479–500, 2013.

[21] M. D. Ortigueira and J. A. T. Machado, "What is a fractional derivative?" *Journal of Computational Physics*, vol. 293, pp. 4 – 13, Jul 2015.

[22] E. J. S. Pires, J. A. T. Machado, P. B. D. M. Oliveira, J. B. Cunha, and L. Mendes, "Particle swarm optimization with fractional-order velocity," *Nonlinear Dynamics*, vol. 61, pp. 295–301, Jul 2010.

[23] M. S. Couceiro, R. P. Rocha, N. M. F. Ferreira, and J. A. T. Machado, "Introducing the fractional-order darwinian pso," *Signal, Image and Video Processing*, vol. 6, pp. 343–350, Apr 2012.

[24] P. Ghamisi, M. S. Couceiro, F. M. L. Martins, and J. A. Benediktsson, "Multilevel image segmentation based on fractional-order darwinian particle swarm optimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2382–2394, May 2014.

[25] M. Couceiro, N. Ferreira, and J. T. Machado, "Fractional order darwinian particle swarm optimization," *3th Symposium on Fractional Signals and Systems, FSS*, vol. 2011, pp. 127–136, Jan 2011.

[26] A. Ahilan, G. Manogaran, C. Raja, S. Kadry, S. N. Kumar, C. A. Kumar, T. Jarin, S. Krishnamoorthy, P. M. Kumar, G. C. Babu, N. S. Murugan, and Parthasarathy, "Segmentation by fractional order darwinian particle swarm optimization based multilevel thresholding and improved lossless prediction based compression algorithm for medical images," *IEEE Access*, vol. 7, pp. 89 570–89 580, 2019.

[27] R. Singh, A. Kumar, and R. Sharma, "Fractional order pid control using ant colony optimization," in *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, July 2016, pp. 1–6.

[28] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6, pp. 467–488, Jun 1982.

[29] A. Narayanan and M. Moore, "Quantum-inspired genetic algorithms," in *Proceedings of IEEE International Conference on Evolutionary Computation*, May 1996, pp. 61–66.

[30] K. H. Han and J. H. Kim, "Quantum-inspired evolutionary algorithm for a class of combinatorial optimization," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 6, pp. 580–593, Dec 2002.

[31] S. Dey, I. Saha, S. Bhattacharyya, and U. Maulik, "Multi-level thresholding using quantum inspired meta-heuristics," *Knowledge-Based Systems*, vol. 67, pp. 373–400, Sept 2014.

[32] S. Dey, S. Bhattacharyya, and U. Maullik, *Quantum-Inspired Automatic Clustering Technique Using Ant Colony Optimization Algorithm*. IGI Global, Apr 2018, ch. 002, pp. 27–54.

[33] V. Tkachuk, "Quantum genetic algorithm based on qutrits and its application," *Mathematical Problems in Engineering*, vol. 2018, no. 8614073, 2018.

[34] P. Ghamisi, M. S. Couceiro, J. A. Benediktsson, and N. M. Ferreira, "An efficient method for segmentation of images based on fractional calculus and natural selection," *Expert Systems with Applications*, vol. 39, no. 16, pp. 12 407 – 12 417, 2012.

[35] B. Zhang, "Generalized k-harmonic means boosting in unsupervised learning," Hewlett-Packard Laboratories, Tech. Rep., Oct 2000.

[36] H. Mahi, N. Farhi, and K. Labed, "Remotely sensed data clustering using k-harmonic means algorithm and cluster validity index," in *Computer Science and Its Applications*, Cham, 2015, pp. 105–116.

[37] K. Tan, F. Wu, Q. Du, P. Du, and Y. Chen, "A parallel gaussian-bernoulli restricted boltzmann machine for mining area classification with hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 2, pp. 627–636, Feb 2019.

[38] X. Wang, K. Tan, Q. Du, Y. Chen, and P. Du, "Caps-triplegan: Gan-assisted capsnet for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7232–7245, Sep. 2019.

[39] F. Xie, F. Li, C. Lei, J. Yang, and Y. Zhang, "Unsupervised band selection based on artificial bee colony algorithm for hyperspectral image classification," *Applied Soft Computing*, vol. 75, pp. 428 – 440, 2019.

[40] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectralspatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5771–5782, Sep. 2014.

[41] D. McMahon, *Quantum Computing Explained*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2008.

[42] K. Koonsanit and C. Jaruskulchai, "Band selection for hyperspectral image using principal components analysis and maxima-minima functional," in *Knowledge, Information, and Creativity Support Systems*, Berlin, Heidelberg, 2011, pp. 103–112.

[43] S. Saitta, B. Raphael, and I.F.C.Smith, "A bounded index for cluster validity," in *Machine Learning and Data Mining in Pattern Recognition*, Berlin, Heidelberg, 2007, pp. 174 – 187.

[44] A. Horé and D. Ziou, "Image quality metrics: Psnr vs. ssim," in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 2366–2369.

[45] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.