SPECTRAL UNMIXING USING MODIFIED PICTURE FUZZY CLUSTERING METHOD

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF SCIENCE
IN
MATHEMATICS

Submitted by

VATSAL JAIN (23/MSCMAT/56)

Under the supervision of

MS. TRASHA GUPTA

DEPARTMENT OF APPLIED MATHEMATICS
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

MAY, 2025

DEPARTMENT OF APPLIED MATHEMATICS
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **VATSAL JAIN**, Roll No – **23/MSCMAT/56** student of M.Sc. **Mathematics**, hereby declare that the project Dissertation titled "**Spectral Unmixing using Modified Picture Fuzzy Clustering Method**" which is submitted by me to the **Department of Applied Mathematics**, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of degree of Master of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi   Vatsal Jain

Date: 21.05.25

DEPARTMENT OF MECHANICAL ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**Spectral Unmixing using Modified Picture Fuzzy Clustering Method**" which is submitted by **Vatsal Jain**, Roll No's – **23/MSCMAT/56**, **Department of Applied Mathematics** , Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Maaster of Science, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi **Ms. Trasha Gupta**

Date: 21.05.2025 **SUPERVISOR**

DEPARTMENT OF MECHANICAL ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I wish to express my sincerest gratitude to Ms. **Trasha Gupta** for her continuous guidance and mentorship that she provided me during the project. She showed me the path to achieve my targets by explaining all the tasks to be done and explained to me the importance of this project as well as its industrial relevance. She was always ready to help me and clear my doubts regarding any hurdles in this project. Without her constant support and motivation, this project would not have been successful.

Place: Delhi Vatsal Jain

Date: 21.05.2025

Abstract

Unsupervised clustering methods are used in spectral unmixing of multispectral images since to estimate the endmembers. But their performance is hampered due to inadequate data and imprecise initialization of endmembers. In this study we propose modified picture fuzzy c-means (PFCM) technique to initialize the endmembers.

Experimental assessments carried out on multi-spectral datasets reveal that our approach consistently surpasses current leading techniques. Performance indicators such as the BVI, UI, and IRCM validate its enhanced capabilities in endmember separability, accuracy of abundance estimation, and spatial consistency. These findings underscore the efficacy and adaptability of the proposed method across various multi-spectral imaging contexts. Incorporating the heuristic consistently enhances performance in various datasets. Significantly, a revised version of PFCM that includes our heuristic achieving a BVI of 0.004, UI of 0.999 and IRCM of 0.466 on the multispectral image dataset. Tests performed on both the Iris dataset and a multispectral image reveal enhancements in the BVI, UI, and IRCM metrics, as well as faster convergence rates compared to traditional implementations.

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Spectral Imaging

Spectral imaging is an advanced technique that acquires image data over a broad spectrum of wavelengths within the electromagnetic spectrum. In contrast to conventional photography, which captures only three primary colors (red, green, and blue), spectral imaging gathers numerous narrow spectral bands, ranging from dozens to hundreds. This capability enables the precise identification and analysis of materials, surfaces, or biological tissues based on their distinct spectral signatures. Spectral imaging has an extensive application in fields such as remote sensing, agriculture, environmental monitoring, medical diagnostics, and industrial inspection. Through the examination of spectral data, we can identify subtle variations in material composition, assess crop health, or detect pollutants in aquatic environments.

Multispectral imaging (MSI) is a specialized form of spectral imaging. Originally created for military reconnaissance, MSI captures and analyzes data from specific wavelengths within the electromagnetic spectrum, uncovering details that are not perceivable by the human eye. This technique ensures that wavelengths are isolated without overlap, enabling users to examine the effects of distinct subsets of wavelengths. The benefits of multispectral imaging, which set it apart from RGB imaging, include monitoring changes, detecting concealed patterns, achieving ultra-high resolution, and providing detailed compositional information. Its ability to track subtle, often unnoticed changes over time holds considerable importance in agriculture and environmental monitoring. Additionally, the ultra-high resolution of multispectral images offers users precise and comprehensive data, which is especially advantageous in fields that demand high accuracy, such as studies of vegetation. The capability of MSI to detect light beyond the visible spectrum and deliver valuable insights is extensively applied in weather forecasting and environmental monitoring, highlighting its practical significance.

Multi-spectral imaging is a significant component of remote sensing technology. These images possess exceptionally high spectral resolution. They capture data across numerous narrow and adjacent electromagnetic bands that span a wide range of the electro-magnetic spectrum, typically from visible to infrared regions. However, due to the distance of sensor from Earth's surface and various technical factors, the spatial resolution of most of these images is often quite coarse, meaning that a single pixel can represent a large area on the Earth's surface.

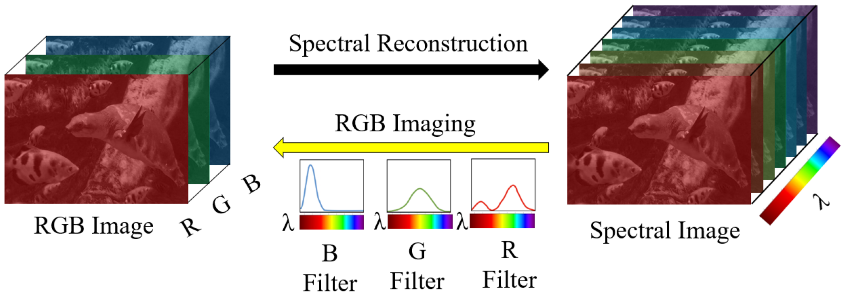

Figure 1.1: Spectral reconstruction from an RGB image

1.2 Spectral Unmixing

Spectral unmixing is a technique used in remote sensing to analyze multispectral images where each pixel may represent a mixture of multiple materials.The objective of spectral unmixing is to break down these mixed pixels into a set of pure spectral signatures, referred to as endmembers, and to assess their respective proportions, known as abundances. This technique improves the precision of land cover classification and material identification, rendering it an essential tool for environmental monitoring, agriculture, and urban research. Clustering methods are typically divided into two categories: hard clustering and soft clustering. In hard clustering, each data point is exclusively allocated to one cluster, ensuring no overlap among clusters. A well-known example of this method is the K-Means algorithm, which groups observations according to their proximity to the nearest cluster centroid. Conversely, soft clustering permits data points to be part of multiple clusters at the same time, with varying levels of membership. The FCM algorithm illustrates this method by assigning a membership value to each data point for every cluster, indicating the extent of association. Clustering is considered as a useful tool in the processes of pattern recognition and knowledge discovery from a database; thus being applied to various crucial applications [1]. Nonetheless, the clustering quality of FCM is not high since this algorithm is deployed on the basis of the traditional fuzzy sets, which have some limitations in the membership representation, the determination of hesitancy and the vagueness of prototype parameters. Among various clustering techniques, FCM and its extensions, such as Intutionistic Fuzzy C-Means (IFCM), Picture Fuzzy C-Means (PFCM), have gained popularity due to their ability to handle overlapping and complex data structure. However, the effectiveness of PFCM heavily depends on the initialization of cluster centers, as poor initialization can lead to slow convergence and suboptimal clustering results. Traditional methods, such as random initialization, often fail to provide a robust starting point, leading to inconsistencies in cluster formation.

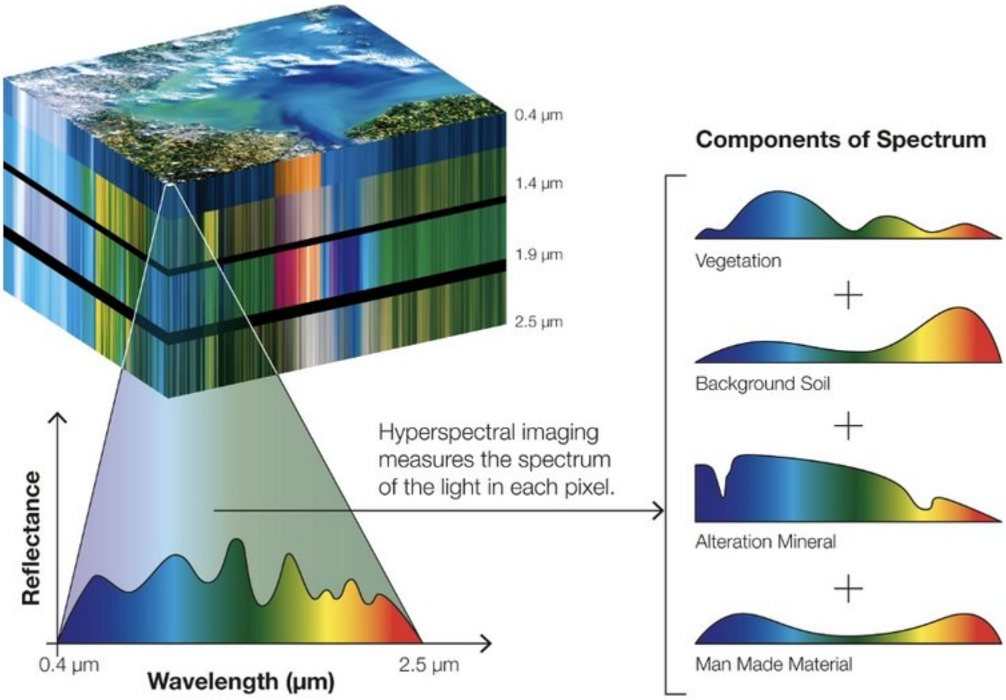

Figure 1.2: Spectral unmixing of various components (endmembers)

1.3 Modified PFCM

Picture Fuzzy C-Means (PFCM) is an advanced version of the traditional Fuzzy C-Means (FCM) algorithm, which utilizes picture fuzzy set theory to more effectively address uncertainty, hesitancy, and indeterminacy present in data. Unlike FCM, which focuses solely on membership degrees, PFCM introduces three distinct parameters for each data point: the degree of membership, non-membership, and hesitation. This triadic framework enables PFCM to more precisely represent the ambiguity found in complex real-world datasets, particularly in situations where cluster assignment decisions are not straightforward. Consequently, PFCM improves clustering performance in fields such as medical image analysis, pattern recognition, and decision support systems, where ambiguous and conflicting information frequently occurs. The algorithm optimizes an objective function that integrates all three components, resulting in more robust and informative clustering results compared to earlier models.

Game theory, a mathematical framework for evaluating strategic interactions, has been widely used for optimization and decision-making problems. The Shapley Value is a fundamental concept in cooperative game theory that provides a fair way to distribute the total gains (or costs) among players in a coalition based on their individual contributions. In this paper, we use game theoretic approach to improve the initialization of PFCM. We achieve a more stable and effective initial prototype by transforming cluster selection into a strategic game. Here, data points compete as rational players for optimal cluster allocations. This game-theoretic technique assures that the initial cluster centers are appropriately located, improving the overall performance of PFCM.

The primary contribution of this research lies in modifying PFCM with a game theory heuristic to optimize the initialization of cluster centers. Once the initial cluster centers have been computed using the heuristic, the usual PFCM procedure refines the clusters. This strategy not only enhances the stability and accuracy of PFCM, but it also addresses concerns associated with inadequate initialization.

This report is structured into seven primary chapters that together offer a thorough overview of the research conducted. Chapter 1, Introduction, outlines the background, motivation, and objectives of the study. Chapter 2, Related Work, provides a review of the existing literature. Chapter 3, Preliminary, introduces the essential concepts and definitions required to comprehend the following content. Chapter 4, Methodology, elaborates on the research approach, which is divided into two phases, Phase I and Phase II, to systematically outline the steps taken. Chapter 5, Experimental setup, describes the experimental design, tools, and datasets used to validate the proposed method. Chapter 6, Results and Discussion, presents the results of the experiments and interprets the findings in relation to the study's objectives. Lastly, Chapter 7, Conclusion and Future Scope, encapsulates the main contributions and outlines potential directions for future work.

Chapter 2

RELATED WORK

The concept of unsupervised clustering in machine learning has evolved over time. The study [2] presents a Unsupervised k-means clustering algorithm that removes the necessity for initialization and parameter selection, while also autonomously identifying the optimal number of clusters, thereby enhancing both the efficiency and accuracy of clustering. Key unsupervised machine learning methods like k-means, hierarchical clustering, and PCA are investigated in this work [1] with an eye toward their application in high-dimensional data analysis thereby enhancing healthcare research and individualized patient care. Entropy-based fuzzy c-means techniques have been used in many disciplines to handle uncertain data and enhance clustering accuracy. The study [3] introduces an entropy-based variant of the Fuzzy C- Ordered Means algorithm aimed at minimizing the sensitivity to the fuzzification parameter m, thereby enhancing clustering precision, resilience to noise, and efficiency of convergence. The study [4] introduces an entropy-driven fuzzy c-means clustering algorithm designed for the segmentation of noisy 3D brain MRI images, tackling issues of uncertainty and intensity inhomogeneity through the application of both global and locally constrained membership functions. The study [5] introduces an effective entropy-driven fuzzy clustering approach that autonomously determines cluster centers and develops a rule-based fuzzy model, thereby minimizing time complexity and streamlining parameter selection. The study [6] offers a fuzzy clustering-based method for data stream clustering and idea drift detection based on entropy, which provides more precise correlations and higher sensitivity than traditional hard clustering. The research [7] proposes a robust unsupervised spectral unmixing approach, termed as Gradient Descent Maximum Entropy (GDME), which employs the maximum entropy principle to increase abundance estimation accuracy in noisy or complicated hyperspectral data. The paper [8]introduces the unsupervised fuzzy partition-optimal number of classes (UFP-ONC) algorithm, which combines fuzzy k-means and fuzzy maximum likelihood estimation to conduct fuzzy clustering without making previous assumptions about cluster counts, resulting in better classification accuracy.

Particularly for picture segmentation and disease identification, many medical researchers have effectively used fuzzy c-means clustering. The work [9] presents the Fuzzy Joint Points (FJP) algorithm, a hierarchical fuzzy clustering technique addressing outliers over spatial data applications, noise robustness, and hidden cluster structures detection.A variety of studies have performed comparative analyses of different unsupervised clustering algorithms to assess their effectiveness across a range of scenarios. The paper [10] compares Fuzzy C-Means with Subtractive Clustering algorithms in terms of clustering validity, modeling capabilities, and performance, with a focus on subtractive clustering's accuracy, consistency, and absence of training requirements. The research [11] presents a theoretical framework for comparing clustering algorithms based on functional aspects, including fuzzy membership types, learning modes, and structural qualities, in order to unify clustering system evaluation. The research [12] examines various fuzzy clustering methods used to build fuzzy classification models, concentrating on computational efficiency and accuracy across benchmark datasets and a real-world bankruptcy prediction scenario. The study [13] compares Fuzzy C-Means and Entropy-Based Fuzzy Clustering on numerous datasets, measuring cluster quality and computation time, and visualizing with self-organizing maps for easier interpretation.

Various algorithms are used in multispectral unmixing. The research [14] analyzes the mathematical properties of picture fuzzy clustering, proving its convergence and examining its loss function behavior. The research [15] proposes an entropy-based spatial fuzzy c-means technique for multispectral image unmixing, improving noise reduction, classification, and efficiency. Using geographical information, the research [16] builds a spectral unmixing technique to improve proportion estimates and investigate spatial spectral enhancements of end-member identification. Using an optimization method based on ADMM, the research [17] presents a linear mixing model for multispectral unmixing that explicitly accounts for spatial and spectral variability of the endmembers, thus lowering the estimate errors and increasing the accuracy. The research [18] enhances fuzzy C-means clustering by using picture fuzzy sets, addressing limitations in the representation and uncertainty of membership. The proposed FC-PFS algorithm shows improved clustering accuracy over existing methods in real-world datasets. The study [19] improves Linear Spectral Mixture Analysis (LSMA) by incorporating a Fully Constrained Least Squares (FCLS) approach, which guarantees precise material quantification in multispectral images while effectively managing nonnegativity and sum-to-one constraints. This work by [20] improves cluster stability and accuracy in land cover categorization by means of a game theory-based Shapley value strategy, hence strengthening FCM clustering for remote sensing image segmentation.

Chapter 3

PRELIMINARY

In many spectral unmixing algorithms clustering has been used rather than linear unmixing methods to tackle the cases where we don't know about the endmembers involved. Clustering is an unsupervised method of machine learning which groups various data points together on the basis of their properties (here pixel values). K-means and FCM are two typical clustering algorithms used in machine learning. However, it is observed that FCM performs better as compared to K means.

Further, focusing on FCM, which has been extended Intutionistic FCM, Picture FCM etc.

3.1 Fuzzy-c-Means

The Objective function for FCM is

$$J = \sum_{k=1}^{N} \sum_{j=1}^{C} \mu_{kj}^{m} ||X_k - V_j||^2 \tag{3.1}$$

subject to contraints

$$\begin{cases} 0 \leq \mu_{kj} \leq 1 \\ \sum_{j=1}^{C} \mu_{kj} = 1 \end{cases} \tag{3.2}$$

Solving the above optimization problem using Lagrange's Multiplier method we get the cluster prototype,

$$V_j = \frac{\sum_{k=1}^{N} \mu_{kj}^{m} X_k}{\sum_{k=1}^{N} \mu_{kj}^{m}} \qquad (j = 1, 2, ..., C) \tag{3.3}$$

$$\mu_{kj} = \frac{1}{\sum_{i=1}^{C} \left(\frac{||X_k - V_j||}{||X_k - V_i||} \right)^{\frac{2}{m-1}}} \tag{3.4}$$

$$k = 1, 2, 3....., N \qquad j = 1, 2, 3, ..., C$$

3.2 Intutionistic Fuzzy-c-Means

Intuitionistic Fuzzy C-Means (IFCM) is a sophisticated clustering technique that enhances the conventional Fuzzy C-Means (FCM) by integrating the notion of intuitionistic fuzzy sets, which consist of two elements: membership and hesitation degrees. This added dimension of information aids in more effectively addressing uncertainty and ambiguity within data, particularly in intricate or noisy contexts such as remote sensing imagery. In IFCM, each data point is assigned not only a degree of membership to a cluster but also a corresponding hesitation margin, which indicates the uncertainty of the assignment. This characteristic renders IFCM more resilient and efficient in situations where data categories overlap or are vaguely defined, resulting in superior clustering outcomes compared to traditional FCM. The Objective function for IFCM is

$$J = \sum_{k=1}^{N} \sum_{j=1}^{C} \mu_{kj}^{m} ||X_k - V_j||^2 + \sum_{j=1}^{C} \pi_j^* e^{1-\pi_j^*} \tag{3.5}$$

where

$$\pi_j^* = \frac{1}{N} \sum_{k=1}^{N} \pi_{kj} \tag{3.6}$$

subject to contraints (3.2)

Solving the above optimization problem using Lagrange's Multiplier method we get the cluster prototype as in equation (3.3), (3.4)

Further, the hesitation degree is calculated by

$$\pi_{kj} = 1 - \mu_{kj} - (1 - \mu_{kj})^{\alpha})^{\frac{1}{\alpha}} \tag{3.7}$$

and π_{kj} is used to update the membership degree as follows:

$$\mu_{kj} = \mu_{kj} + \pi_{kj} \tag{3.8}$$

$$k = 1, 2, 3....., N \qquad j = 1, 2, 3, ..., C$$

Chapter 4

METHODOLOGY

In this paper, we have proposed a 2-phase method to overcome the limitations of PFCM by using game-theoretic heuristic. Our proposed method consists of 2 phases. Phase I initializes the endmember, abundance and neutality matrices. Phase II uses PFCM technique to find the final abundance and endmember matrices

4.1 Phase-I

Here we define a game theory heuristic to initialize the cluster centers. In cooperative game theory, the Shapley Value is a basic idea used to equitably divide the overall payoff among participants depending on their contributions to several coalitions.It ensures that each player receives their fair share of the overall value created by cooperation. In multispectral unmixing, the Shapley value aids in evaluating the contribution of endmembers and enhancing pixel-wise interpretability, which is beneficial for initializing the cluster centers (i.e., endmembers).

$$\phi_i = \frac{1}{2}\sum_{j=1}^{N} sim(X_i, X_j) \qquad (i = 1, 2, ..., N) \tag{4.1}$$

where N is the no. of data points.

The similarity function acts as a quantitative indicator of the resemblance between data instances or features, and is essential in assessing the degree of similarity a sample has with others in the dataset. X denotes the data matrix with dimensions N $\times$ C. Here X_i and X_j ($\in$ X)represent a row corresponding to each data point from the data matrix X .

$$sim(X_i, X_j) = 1 - \frac{||X_i - X_j||}{max(||X_i - X_j||)} \tag{4.2}$$

where $\|.\|$ is computed using Euclidean distance.

By employing the endmembers derived from the Shapley value-based method, we set up the abundance matrix in alignment with the Fuzzy C-Means (FCM) algorithm. The abundance matrix is initialized using equation 3.4 The initial neutrality matrix η is obtained by the formula

$$\eta_{kj} = \begin{cases} 0.5 & , X_j \in \mathbb{C} \text{ and } k = j \\ 0 & , X_j \in \mathbb{C} \text{ and } k \neq j \\ \frac{1-|X_j - X_k|}{2} & , X_j \notin \mathbb{C} \text{ and } X_k \in \mathbb{C} \end{cases} \tag{4.3}$$

4.2 Phase-II

Here we elaborate the Phase II PFCM which helps in forming unsupervised clusters so as to unmix the multispectral image using the cluster centers initialised in Phase I.

The variable V signifies the endmembers, μ refers to the abundance (or membership) matrix and η indicates the neutrality matrix, which illustrates the extent to which a specific data point (or pixel) is not associated with a particular cluster. ξ represents the hesitancy matrix, reflecting the uncertainty regarding a data point (or pixel) belonging to a specific cluster. The Objective function for PFCM is

$$J = \sum_{k=1}^{N}\sum_{j=1}^{C}(\mu_{kj}(2-\xi_{kj}))^m ||X_k - V_j||^2 + \sum_{k=1}^{N}\sum_{j=1}^{C}\eta_{kj}(\log \eta_{kj} + \xi_{kj}) \tag{4.4}$$

subject to constraints:

$$\mu_{kj} + \xi_{kj} + \eta_{kj} \leq 1$$

$$\sum_{j=1}^{C}\mu_{kj}(2-\xi_{kj}) = 1$$

$$\sum_{j=1}^{C}\left(\eta_{kj} + \frac{\xi_{kj}}{C}\right) = 1$$

$$k = 1, 2, 3....., C \qquad j = 1, 2, 3, ..., N$$

We obtain optimal solutions of the system by solving above optimization problem using Lagrange's Multiplier method.

$$\xi_{kj} = 1 - (\mu_{kj} + \eta_{kj}) - (1 - (\mu_{kj} + \eta_{kj})^{\alpha})^{\frac{1}{\alpha}} \tag{4.5}$$

$$\mu_{kj} = \frac{1}{\sum_{i=1}^{C}(2-\xi_{kj})\left(\frac{||X_k - V_j||}{||X_k - V_i||}\right)^{\frac{2}{m-1}}} \tag{4.6}$$

where $\|.\|$ is computed using Euclidean distance.

$$\eta_{kj} = \frac{e^{-\xi_{kj}}}{\sum_{i=1}^{C} e^{-\xi_{ki}}}\left(1 - \frac{1}{C}\sum_{i=1}^{C}\xi_{kj}\right) \tag{4.7}$$

$$V_j = \frac{\sum_{k=1}^{N}(\mu_{kj}(2-\xi_k j))^m X_k}{\sum_{k=1}^{N}(\mu_{kj}(2-\xi_k j))^m} \tag{4.8}$$

$$k = 1, 2, 3....., C \qquad j = 1, 2, 3, ..., N$$

The algorithm for 2-phase method is defined in Algorithm 1

Algorithm 1 PFCM with Game Theory

Input: A threshold value δ and a data matrix X of order n x m.
Output: Final Abundance matrix, U and endmember matrix, C
function GAME THEORY (data , δ)
    C=[ ]
    $i \leftarrow 1$
    while $i \leq n$ **do**
        Calculate Shapley value ϕ_i using equation (4.1)
    end while
    $Q \leftarrow X$
    $\phi_{temp} \leftarrow \phi$
    while $Q \neq \emptyset$ **do**
        t= arg max ϕ_i
        $C = C \cup X_t$
        $I = k : sim(X_k, X_t) \geq \delta$ using equation (4.2)
        $P = X_i \in Q : sim(X_i, X_t) \geq \delta$ using equation (4.2)
        $Q = Q \setminus (P \cup X_t)$
        $\phi_{temp} = \phi_{temp} \setminus (\phi_t \cup \phi_k)$
    end while
    return C
end function
function PFCM (data , C,α)
    Initialize U using equation (3.4)
    Initialize N using formula (4.3)
    Initialize Z using equation (4.5)
    Calculate Objective function (O) using equation (4.4)
    Iterate U,N,Z,C,O until convergence of objective function using equations (4.6), (4.7), (4.5), (4.8) and (4.4) respectively
    return U, C
end function

Chapter 5

EXPERIMENT SETUP

In this study, we utilized two distinct datasets. The first is the Iris dataset, a well-established benchmark in pattern recognition and machine learning. It comprises 150 samples, evenly divided among three species of Iris flowers: Iris setosa, Iris versicolor, and Iris virginica. Each sample is characterized by four numerical attributes—sepal length, sepal width, petal length, and petal width—expressed in centimeters. The Iris dataset is frequently employed for evaluating classification algorithms and data visualization methods due to its balanced class distribution and straightforward nature. The second dataset consists of a multispectral image derived from prior research, aimed at capturing reflectance data across various spectral bands, including those outside the visible spectrum. Unlike conventional RGB images, this dataset offers a more comprehensive array of features by integrating information from near-infrared and short-wave infrared bands. Each pixel in this dataset contains a spectral signature, facilitating in-depth analysis of surface characteristics and material composition. This dataset is particularly advantageous for applications such as land cover classification, anomaly detection, and environmental monitoring, where spectral diversity is essential.

We have performed the experiment on a matlab version 2024 on a system having 16GB RAM and i9 processor. Moreover, various indexes like Bensaid validity index (BVI), uniformity index (UI), and Inter-cluster to Intra-cluster Compactness Measure (IRCM) are used to evaluate the quality of clustering results. These indices guarantee that the clustering structure is relevant and efficient by helping to ascertain the degree of separation, compactness, and balance among the clusters.

Using BVI, we can assess the quality of the clusters.It is particularly useful for fuzzy clustering algorithms. A lower BVI value indicates that the clusters are better isolated from each other and more compact (data points within the same cluster are close to each other). Thus, a low BVI indicates that our clustering result is probably good.

$$BVI = \sum_{p=1}^{C} \frac{\sum_{j=1}^{N} \mu_{pj}^{2} (X_j - V_p)^2}{n_p \sum_{k=1}^{C} (V_k - V_p)^2} \tag{5.1}$$

The uniformity of data point distribution across clusters is evaluated by UI. Indicates the degree of equality with which the data set is distributed among the clusters. A more balanced clustering result is shown by a lower UI value, which implies that the data points are distributed more evenly among the clusters. However, depending on the application, a higher UI value could indicate an imbalance in which some clusters are significantly denser or sparser than others, which could be undesirable.

$$UI = 1 - \frac{2}{N} \sum_{p=1}^{C} \frac{\sum_{X_j \in C_p} (X_j - \overline{X_p})^2}{[\max_{X_j \in C_p}(X_j) - \min_{X_j \in C_p}(X_j)]^2} \tag{5.2}$$

IRCM is an index based on ratios that assesses the distance between clusters (inter-cluster) in relation to the compactness within clusters (intra-cluster). A higher IRCM value indicates superior clustering, signifying that clusters are not only densely packed internally but also distinctly separated from one another. This index is straightforward and commonly utilized to measure the efficacy of clustering concerning spatial organization and the clarity of boundaries between various groups.

$$IRCM = \frac{\sum_{p=1}^{C} y_p c_p}{\sum_{p=1}^{C} y_p} \tag{5.3}$$

Chapter 6

RESULTS and DISCUSSION

This section analyzes the effectiveness of different clustering techniques utilized for multispectral unmixing. The findings are presented in Tables (6.1) and (6.2). We evaluated the abundance matrix in relation to the final endmembers by employing validity indices, including the Bensaid Validity Index (BVI), the Unmixing Index (UI), and the Improved Regional Consistency Measure (IRCM). Our heuristic method shows a quicker convergence when applied to FCM, IFCM, and PFCM, respectively. The accompanying graphs (6.1), (6.2) and (6.3) demonstrates a substantial improvement in the quality of the clustering methods following the implementation of our heuristic.

The combination of our proposed heuristic with clustering algorithms like FCM, IFCM, and PFCM results in notable enhancements in performance, especially regarding the initialization of endmembers. The heuristic-modified PFCM exhibited the highest performance among all methods tested. When utilized on the Iris dataset and a multispectral image, this setup recorded BVI values of 0.0026, 0.0042, UI values of 0.9988, 0.9998, and IRCM values of 0.2738, 0.4669 respectively. Additionally, it was observed that the heuristic facilitated faster convergence for all three algorithms—FCM, IFCM, and PFCM—highlighting its effectiveness in not only boosting accuracy but also improving computational efficiency.

	FCM	FCM with Game Theory	IFCM	IFCM with Game Theory	PFCM	PFCM with Game Theory
BVI	0.006578	0.004252	0.005761	0.003833	0.003584	0.002619
UI	0.980495	0.990495	0.981495	0.991495	0.990851	0.998851
IRCM	0.214224	0.228403	0.229841	0.236598	0.246198	0.273896
No. of iterations	34	29	27	24	23	18

Table 6.1: Performance metric for IRIS dataset

	FCM	FCM with Game Theory	IFCM	IFCM with Game Theory	PFCM	PFCM with Game Theory
BVI UI IRCM	0.007727 0.979879 0.440357	0.006727 0.986879 0.442523	0.006829 0.989873 0.465912	0.006146 0.989988 0.465913	0.005237 0.999776 0.466129	0.004288 0.999879 0.466913
No. of iterations	32	28	30	27	24	20

Table 6.2: Performance metric for Multispectral Image

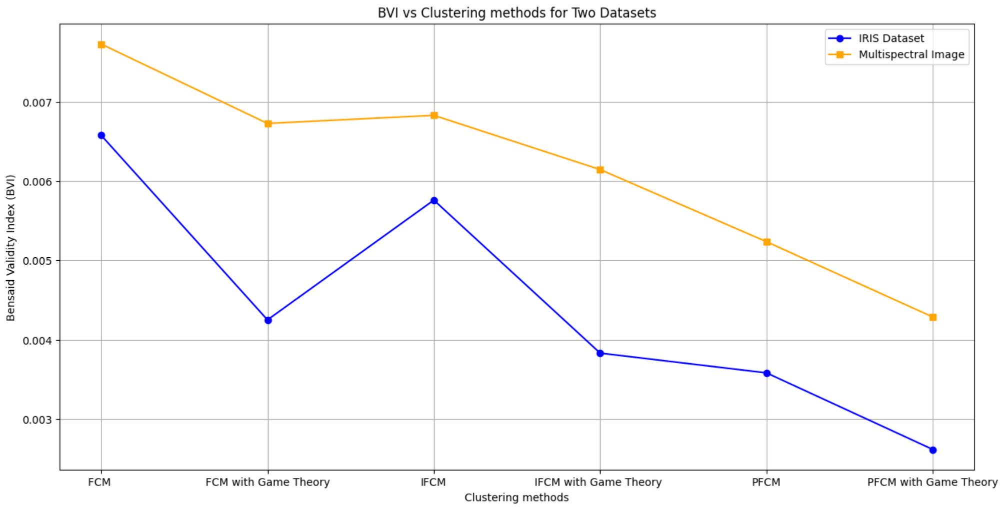

Figure 6.1: Comparision of BVI for Two Datasets

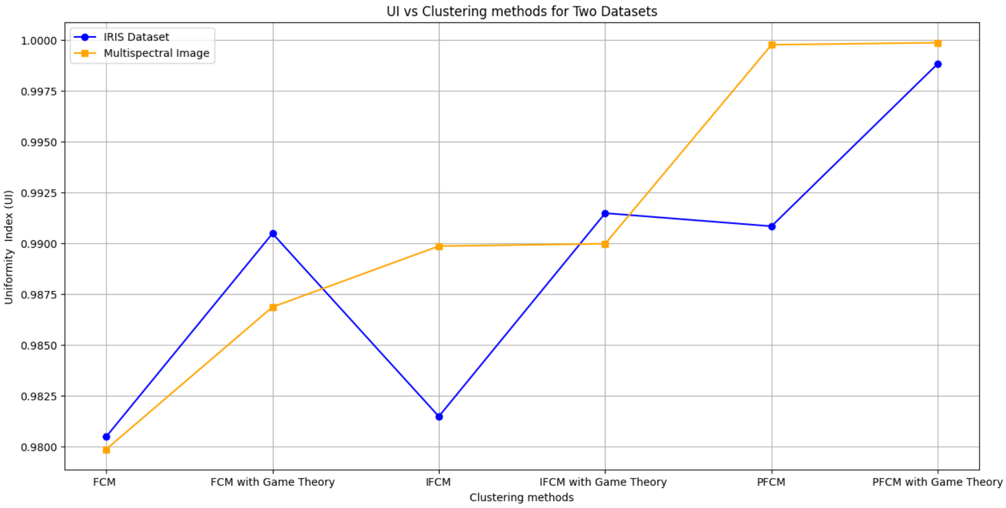

Figure 6.2: Comparision of UI for Two Datasets

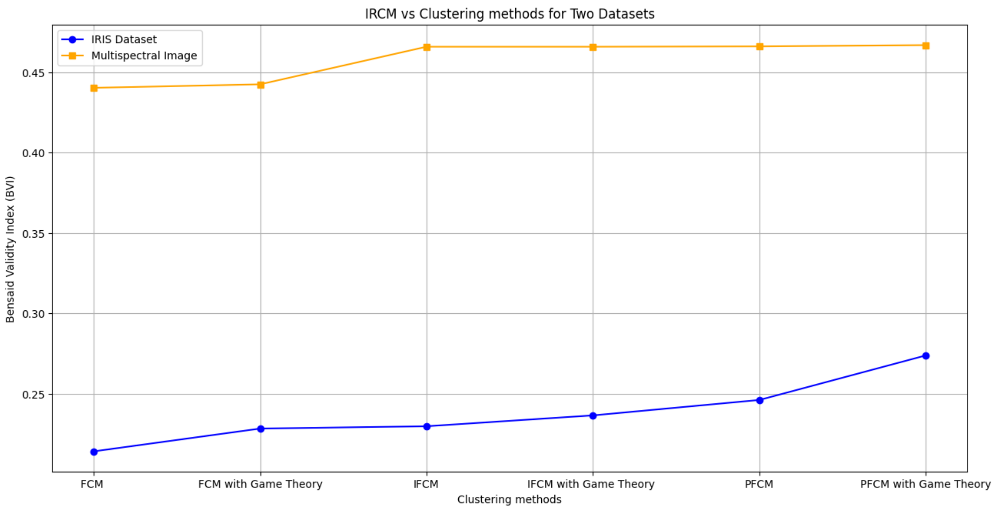

Figure 6.3: Comparision of IRCM for Two Datasets

Chapter 7

CONCLUSION and FUTURE SCOPE

In this research, a new approach for multispectral unmixing was introduced and assessed using established performance metrics. The efficacy of this method was confirmed through various cluster validity indices, including the Bensaid Validity Index (BVI), Uniformity Index (UI), and Inter-cluster to Intra-cluster Compactness Measure (IRCM). The experimental findings reveal that the proposed technique achieves enhanced clustering quality, marked by greater inter-cluster separation, improved intra-cluster compactness, and increased uniformity. These results suggest that the method not only improves unmixing accuracy but also offers a more dependable representation of the underlying spectral signatures, surpassing current methodologies.

The future of multispectral imaging is highly promising across various fields, propelled by advancements in sensor technology, data processing, and artificial intelligence. In agriculture, it is set to transform precision farming by facilitating real-time monitoring of crops, detection of diseases, and prediction of yields through sensors mounted on drones. In the realm of environmental monitoring, multispectral data will play an increasingly vital role in tracking deforestation, urban growth, water quality, and natural disasters, offering enhanced temporal and spatial resolution. The defense and surveillance industries are anticipated to implement lightweight multispectral sensors for reconnaissance and target identification across diverse terrains. In the context of space and planetary exploration, multispectral imaging is being incorporated into next-generation satellites and space missions to investigate Earth-like planets, atmospheric conditions, and mineral compositions. Moreover, with the advancement of machine learning and artificial intelligence, the analysis of multispectral data will become more rapid and precise, facilitating real-time decision-making. Miniaturized multispectral sensors integrated into smartphones, UAVs, and nanosatellites will render the technology more accessible and economical, paving the way for innovations in smart cities, precision forestry, public health monitoring, and disaster response.

Appendix A

ACCEPTANCE and PLAGIARISM REPORT

Acceptance Notification - IEEE ICOCT2025 Inbox ×

Microsoft CMT <noreply@msr-cmt.org> Wed, Apr 23, 5:44 PM

to me

Dear Vatsal Jain

Paper ID / Submission ID : 2306

Title : Spectral Unmixing using Modified Picture Fuzzy Clustering Method

Greeting from IEEE ICOCT2025

We are pleased to inform you that your paper has been accepted for the Presentation as a full paper for the- "IEEE 2025 International Conference on Computing Technologies (ICOCT), Bengaluru , Karnataka, India.

All accepted and presented papers will be submitted to IEEE Xplore for the further publication.

You should finish the registration before deadline, or you will be deemed to withdraw your paper:

Complete the Registration Process (The last date of payment Registration is 27 APRIL 2025)

Payment Links :

For Indian Authors: https://rzp.io/rzp/icoct

For Foreign Authors: https://rzp.io/rzp/ICOCTForeign

Further steps like IEEE PDF xpress and E copyright will be given later once registration is over after the deadline.

Vatsal

Delhi_Technological_University_Thesis_Template__Copy_ (1).pdf

Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:97063434

Submission Date

May 21, 2025, 10:17 PM GMT+5:30

Download Date

May 21, 2025, 10:19 PM GMT+5:30

File Name

Delhi_Technological_University_Thesis_Template__Copy_ (1).pdf

File Size

1.7 MB

27 Pages

5,537 Words

31,697 Characters

15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Small Matches (less than 10 words)

Match Groups

- **28** Not Cited or Quoted 14%
  Matches with neither in-text citation nor quotation marks
- **3** Missing Quotations 1%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

- 12% Internet sources
- 6% Publications
- 10% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **28** Not Cited or Quoted 14%
  Matches with neither in-text citation nor quotation marks
- **3** Missing Quotations 1%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

12% Internet sources
6% Publications
10% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	dspace.dtu.ac.in:8080	7%
2	Internet	link.springer.com	1%
3	Submitted works	Ventura County Community College District on 2017-04-25	<1%
4	Submitted works	University of Warwick on 2024-05-20	<1%
5	Publication	Srirupa Das, Somdatta Chakravortty. "Efficient entropy-based spatial fuzzy c-mea...	<1%
6	Submitted works	Delhi Technological University on 2025-05-02	<1%
7	Publication	Srimanta Kundu, Ujjwal Maulik, Anirban Mukhopadhyay. "A game theory-based a...	<1%
8	Submitted works	iGroup on 2014-07-04	<1%
9	Publication	Xuetao Tao, Bin Wang, Liming Zhang, Jian Qiu Zhang. "A new endmember extra...	<1%
10	Submitted works	Free University of Bolzano on 2025-01-29	<1%

11	Submitted works	
Delhi Technological University on 2024-05-09		<1%
12	Publication	
Reza Arablouei. "Spectral Unmixing With Perturbed Endmembers", IEEE Transacti...		<1%
13	Internet	
download.excelingtech.co.uk		<1%
14	Internet	
www.dtic.mil		<1%
15	Internet	
www.nature.com		<1%
16	Internet	
www.nipec.n-i.nhs.uk		<1%
17	Publication	
Bo Miao. "Emissions, Pollutants and Environmental Policy in China - Designing a n...		<1%
18	Publication	
Mei, Shaohui, David Feng, Mingyi He, Roger W. Heymann, and Joan Serra-Sagrista...		<1%
19	Submitted works	
University of New South Wales on 2024-11-14		<1%
20	Submitted works	
University of South Alabama on 2007-04-13		<1%
21	Internet	
arxiv.org		<1%
22	Internet	
dalspace.library.dal.ca		<1%
23	Internet	
pdfslide.net		<1%
24	Internet	
www.igi-global.com		<1%

25 Internet

www.mdpi.com <1%

26 Internet

www.researchgate.net <1%

Bibliography

[1] C. Eckhardt, S. Madjarova, R. Williams *et al.*, "Unsupervised machine learning methods and emerging applications in healthcare," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 31, p. 376–381, 2023. [Online]. Available: https://doi.org/10.1007/s00167-022-07233-7

[2] K. P. Sinaga and M.-S. Yang, "Unsupervised k-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80 716–80 727, 2020.

[3] M. Moradi and J. Hamidzadeh, "Entropy-based fuzzy c-ordered-means clustering algorithm," *New Generation Computing*, vol. 41, pp. 739–775, 2023.

[4] M. Ray, N. Mahata, and J. K. Sing, "Uncertainty parameter weighted entropy-based fuzzy c-means algorithm using complemented membership functions for noisy volumetric brain mr image segmentation," *Biomedical Signal Processing and Control*, vol. 85, p. 104925, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809423003580

[5] J. Yao, M. Dash, S. Tan, and H. Liu, "Entropy-based fuzzy clustering and fuzzy modeling," *Fuzzy Sets and Systems*, vol. 113, no. 3, pp. 381–388, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165011498000384

[6] B. Zhang, S. Qin, W. Wang, D. Wang, and L. Xue, "Data stream clustering based on fuzzy c-mean algorithm and entropy theory," *Signal Processing*, vol. 126, pp. 111–116, 2016, signal Processing for Heterogeneous Sensor Networks. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165168415003576

[7] L. Miao, H. Qi, and H. Szu, "A maximum entropy approach to unsupervised mixed-pixel decomposition," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1008–1021, 2007.

[8] I. Gath and A. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 773–780, 1989.

[9] E. N. Nasibov and G. Ulutagay, "A new unsupervised approach for fuzzy clustering," *Fuzzy Sets and Systems*, vol. 158, no. 19,

pp. 2118–2133, 2007, theme: Data Analysis. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165011407001091

[10] M. S. K. M. Bataineh, M. Najia, "A comparison study between various fuzzy clustering algorithms," *Jordan Journal of Mechanical and Industrial Engineering*, vol. 5, no. 4, pp. 335 – 343, 2011, theme: Data Analysis.

[11] A. Baraldi and P. Blonda, "A survey of fuzzy clustering algorithms for pattern recognition. i," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 29, no. 6, pp. 778–785, 1999.

[12] R. J. Almeida and J. M. C. Sousa, "Comparison of fuzzy clustering algorithms for classification," in *2006 International Symposium on Evolving Fuzzy Systems*, 2006, pp. 112–117.

[13] S. Chattopadhyay, D. K. Pratihar, and S. C. De Sarkar, "A comparative study of fuzzy c-means algorithm and entropy-based fuzzy clustering algorithms," *Computing and Informatics*, vol. 30, pp. 701–720, 2011.

[14] L. H. S. PHAM THI MINH PHUONG, PHAM HUY THONG, "Theoretical analysis of picture fuzzy clustering," *Journal of Computer Science and Cybernetic*, vol. V.34, no. 1, pp. 17–31, 2018.

[15] S. Das and S. Chakravortty, "Efficient entropy-based spatial fuzzy c-means method for spectral unmixing of hyperspectral image," *Soft Computing*, vol. 25, no. 1, p. 7379–7397, 2021. [Online]. Available: https://doi.org/10.1007/s00500-021-05697-2

[16] A. Zare, "Spatial-spectral unmixing using fuzzy local information," in *IEEE Int. Geoscience and Remote Sens. Symposium (IGARSS)*, 2011, pp. 1139–1142.

[17] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 525–538, 2016.

[18] P. Thong and L. Son, "Picture fuzzy clustering: A new computational intelligence method," *Soft Computing*, vol. 20, no. 9, pp. 3549–3562, 2016. [Online]. Available: https://doi.org/10.1007/s00500-015-1712-7

[19] D. Heinz and Chein-I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.

[20] S. Kundu, U. Maulik, and A. Mukhopadhyay, "A game theory-based approach to fuzzy clustering for pixel classification in remote sensing imagery," *Soft Computing*, vol. 25, no. 7, pp. 5121–5129, 2021. [Online]. Available: https://doi.org/10.1007/s00500-020-05514-2