data, while SMKFCM works well with overlapping data.

Experiments in the dissertation have shown that the proposed methods can overcome some disadvantages and produce higher accuracy in most cases than several other methods. They still have some limitations, such as:

- In principle, the proposed methods can work with any dimensional image data, but in fact, it has not been applied to hyperspectral image data. Applications for hyperspectral image often requires a massive amount of calculations, which is only feasible when a parallel computing model or high-performance computing based on graphics processing units (GPUs) is employed.

- The parameters of the algorithms established in the above experiments may not be useful on other data sets. This is due to the fact that surface objects are continually changing in shape, size, and color. Image data of the same object in different periods may be different.

Future works

Although the proposed methods in the dissertation can overcome disadvantages and give better results than several previous approaches. Most algorithms still face difficulty working with large data and multidimensional data. The author believes that further research in this direction can succeed in speeding up calculations and optimizing parameters for algorithms, reducing data dimensions and learning based on deep learning.

- Speed up the calculation: With the explosion of information and data, most algorithms have difficulty facing "big data". Several approaches, including parallel processing, high-performance computing based on GPU architecture, are suggested for this research direction.

- Dimensional reduction: RS image data is often characterized by many dimensions and large capacity, especially hyperspectral RS image; the number of dimensions can be up to hundreds or more. Therefore, reducing the size to eliminate unnecessary attributes (features) will help the algorithms work more effectively.

- Deep learning: For supervised classification problem, it requires a large amount of labeled data for training. While traditional learning algorithms are ineffective, deep learning can solve this problem well. Therefore, this might be a good research direction for the remote sensing image analysis problem for now and in the future.

Abstract

Remote sensing images have been widely used in many fields thanks to their outstanding advantages such as large coverage area, short update time and diverse spectrum. On the other hand, this data is subject to a number of drawbacks, including: a high number of dimensions, numerous nonlinearities, as well as a high level of noise and outlier data, which pose serious challenges in practical applications.

The dissertation develops a number of fuzzy clustering techniques applied to the remote sensing image analysis problem. The proposed methods are based on the type-1 fuzzy clustering and interval type-2 fuzzy clustering. Learning techniques and labeled data are used to overcome some disadvantages of existing methods. The problem of classification and detection of land-cover changes from remote sensing image data is applied to prove the effectiveness of the proposed methods.

Preamble

Problem statement and motivations

RS image data with many advantages have been applied in many different applications. The strong development of satellite technology has led to a large amount of RS image data that needs to be processed. Besides, It also faces many challenges, such as "big data", multi-dimensions and exists many uncertainties and vaguenesses.

For the problem of land-cover mapping, because of the urbanization process, the objects on the surface are constantly changing. Traditional methods of creating land-cover maps did not meet the time and money requirements, which leads to the need for improvement, proposing more modern and powerful new techniques.

It can be seen that the study of RS image analysis problem is essential and has a great significance in terms of academic as well as practical. These are great motivations to help me choose the topic ***"Fuzzy clustering techniques for remote sensing image analysis"*** for my dissertation.

The dissertation contents will focus on developing robust clustering algorithms based on the fuzzy set including the type-1 fuzzy clustering, interval type-2 fuzzy clustering; Combined with some learning methods and labeled data to overcome some drawbacks of the previous method. With the advantage of uncertain data processing, fuzzy clustering is a good choice for RS image analysis problems. Moreover, the approach to semi-supervised learning method is a solution suitable for problems with very little labeled data. The issue of selecting the optimal parameters can be solved by using optimization techniques.

Objectives and scopes

The main objective of the dissertation is to research and develop some robust fuzzy-based methods to classify and detect the land cover changes from RS image data.

The research scope of the dissertation includes the type-1, interval type-2 fuzzy clustering, and several learning methods include the semi-supervised method, kernel technique, and particle swarm optimization (PSO). The problem of classification and detection of land-cover changes from RS image is applied to prove the effectiveness of the proposed method.

Contributions of the thesis

Most of the work described in this dissertation was conducted at the Military Technical Academy in Vietnam. The dissertation has following main contributions:

1. The dissertation proposes two unsupervised fuzzy c-means clustering algorithm (FCM), including DFCM and IFCM. DFCM algorithm proposes using density information for selecting initial centroids for FCM algorithm. IFCM algorithm proposes to using the spectral clustering and spatial information as a preprocessing step to map the original data space to a new space based on the main components. The proposed methods can improve the accuracy of the algorithm compared to the original algorithm.

2. The dissertation develops three semi-supervised fuzzy c-means clustering algorithms, including SMKFCM, SFCM-PSO and GIT2SPFCM-PSO. SMKFCM proposes the multiple-kernel technique to make data better separated; moreover, it uses labelled data to adjust the focus during clustering with the hope that the algorithm runs more stable. SFCM-PSO is a hybrid algorithm between semi-supervised method and PSO optimization technique. GIT2SPFCM-PSO is a hybrid clustering algorithm developed by the semi-supervised possibilistic fuzzy c-means clustering based on interval type-2 fuzzy set with the parameters optimized by PSO technique. By using PSO technique for finding the optimal parameters. The proposed methods achieve better accuracy than existing methods.

Conclusions and future works

Conclusions

The dissertation has presented several robust classification models to overcome the disadvantages of current methods and apply these models to land cover classification of RS image data. The proposed method can be applied to many types of RS images (radar, optics) and spatial resolutions ($10m$, $30m$). In this dissertation, some main contributions can be summarized as follows:

The dissertation proposes two unsupervised fuzzy clustering algorithm which extended from FCM including DFCM [Pub7] and IFCM [Pub1], [Pub3]. DFCM algorithm proposes to use density information to select initial centroids for FCM algorithm. IFCM algorithm proposed the use of spectral clustering as a preprocessing step to map the original data space to a new space based on the main components.

The dissertation also develops three semi-supervised fuzzy clustering algorithms including SMKFCM [Pub8], SFCM-PSO [Pub2] and GIT2SPFCM-PSO [Pub9] which integrate the semi-supervised fuzzy clustering method [Pub4], [Pub5], [Pub6] and PSO technique. SMKFCM algorithm proposes the multiple-kernel technique to improve data separation. Moreover, the proposed method uses labeled data to adjust the focus during clustering; so the algorithm to run with greater stability. For algorithms SFCM-PSO and GIT2SPFCM-PSO, PSO technique is used for finding the optimal parameters.

The proposed algorithms all produce higher accuracy than the original algorithms. From the experimental results of the algorithms proposed in Chapter 2 and Chapter 3, some recommendations are provided as follows:

- When all data is unlabeled, DFCM and IFCM algorithms should be used. The land-cover classification results by IFCM algorithm provide better accuracy than DFCM algorithm, while DFCM algorithm has smaller computational complexity than IFCM algorithm.

- When very little data is labeled, SMKFCM, SFCM-PSO, and GIT2SPFCM-PSO algorithms should be used. GIT2SPFCM-PSO algorithms give the highest accuracy, while SFCM-PSO is suitable for large data cases because they have lower computational complexity than GIT2SPFCM-PSO and SMKFCM algorithms. The GIT2SPFCM-PSO algorithm can work well with highly uncertain

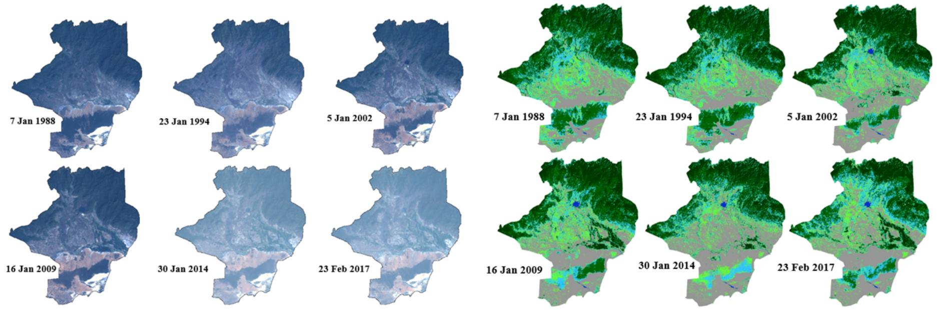

Figure 3.1: RGB image and classification results: Bac Binh district, Binh Thuan province, Vietnam

Table 3.5: Land-cover classification results by the Erdas software, DFCM, IFCM, SMKFCM, SFCM-PSO, and GIT2SPFCM-PSO

Class	***Erdas***	***DFCM***	***IFCM***	***SMKFCM***	***SFCM-PSO***	***GIT2SPFCM-PSO***
1	94.32(%)	98.08(%)	98.11(%)	99.19(%)	98.45(%)	**99.43(%)**
2	94.25(%)	97.35(%)	96.42(%)	**98.88(%)**	97.65(%)	98.63(%)
3	92.33(%)	95.76(%)	97.29(%)	97.60(%)	99.32(%)	**99.45(%)**
4	96.16(%)	96.88(%)	96.34(%)	97.98(%)	97.78(%)	**99.25(%)**
5	93.91(%)	97.21(%)	95.81(%)	**99.14(%)**	98.49(%)	98.76(%)
6	91.79(%)	94.29(%)	97.89(%)	98.47(%)	96.23(%)	**99.05(%)**
Total	93.55(%)	95.83(%)	96.66(%)	98.64(%)	97.71(%)	**99.13(%)**

can be seen that, GIT2SPFCM-PSO algorithm gives the highest accuracy with over 99%, followed by algorithms SMKFCM and SFCM-PSO. The unsupervised algorithms IFCM and DFCM gave worse results than the semi-supervised algorithms. However, they still give classification results with higher accuracy than the Erdas software.

3.6 Summary

This chapter presents three semi-supervised fuzzy clustering algorithms including SMKFCM, SFCM-PSO and GIT2SPFCM-PSO. The classification result on satellite images shows that the proposed methods can be for higher accuracy than some previous algorithms.

The proposed methods can be applied to many types of RS images (radar, optics) and spatial resolutions ($10m$, $30m$). Most of the experiments are used to the problem of the land cover classification of RS images. Although some limitations exist, the proposed methods can provide significantly better classification results than some recent other classification methods.

Organization of the thesis

The dissertation is organized into three chapters and two sections, as follows:

Introduction: This section introduces the general issues of the dissertation. The content presented in this section includes the urgency of the research topic, motivations, objectives and scopes, contributions, scientific and practical meanings and organization of the dissertation.

Chapter 1 discusses the main issues and foundational theories used in the dissertation's studies. In this chapter, an overview of the research and some of the related works is introduced. Several reviews and comparisons of advantages and disadvantages are also given for previous studies.

Chapter 2 introduces two unsupervised fuzzy clustering algorithms, including the density-based fuzzy c-means clustering (DFCM) and the improved fuzzy c-means clustering (IFCM).

Chapter 3 presents three semi-supervised fuzzy clustering algorithms, including the semi-supervised multiple kernel fuzzy c-means clustering (SMKFCM), semi-supervised fuzzy c-means clustering and the particle swarm optimization technique (SFCM-PSO), the interval type-2 semi-supervised possibilistic fuzzy c-means clustering and the particle swarm optimization technique (GIT2SPFCM-PSO).

Conclusions: Summary of dissertation contents, achieved issues and main contributions of the dissertation, some limitations and future research directions.

Chapter 1

Background and related works

This chapter presents the basic knowledge used in the dissertation including fuzzy clustering, interval type-2 fuzzy clustering, and learning techniques. Some methods evaluated the accuracy of the clustering algorithm is also given as a way to demonstrate the effectiveness of the method proposed in the dissertation. This chapter also addresses a number of the previous works with an analysis of their advantages and disadvantages.

1.1 Background concepts

1.1.1 Fuzzy clustering

Definition 1.1. *If X is a set of objects x, a fuzzy set A, $A \subseteq X$ is defined as a set of element pairs of degree as follows:* $\mathrm{A} = \{(\mathrm{x}, \mu_A(\mathrm{x})) \,|\mathrm{x} \in \mathrm{X}\}$

Where $\mu_A(x)$ is a membership function for the fuzzy set A. MF maps each element $x \in X$ to the interval $[0,1]$.

a. Fuzzy c-means clustering

One of the widely used fuzzy set applications is FCM algorithm. This algorithm allows each data element can belong to many different clusters according to different membership grades.

FCM algorithm model is to optimize the objective function:

$$\min\{J_m(U,V,X) = \sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}^{m}d_{ik}^{2}\} \tag{1.1}$$

Where $U = [\mu_{ik}]_{cxn}$ is a fuzzy MF, $V = (v_1, v_2, ..., v_c)$ is a vector of (unknown) cluster centers, $X = \{\mathrm{x}_k, \mathrm{x}_k \in \mathrm{R}^M, \mathrm{k} = 1,...,\mathrm{n}\}$, $d_{ik} = \|v_i - x_k\|$.

b. Possibilistic fuzzy c-means clustering

Possibilistic c-means algorithm (PCM) is proposed by Krishnapuram and Keller, which was introduced to avoid the sensitivity of FCM algorithm. Instead of using the fuzzy MFs such as FCM, PCM uses possibilistic MFs to represent typicality by τ_{ik}, the typicality matrix as $T = [\tau_{ik}]_{cxn}$.

GIT2SPFCM-PSO). On all three datasets, GIT2SPFCM-PSO algorithm has the highest accuracy with TPR, ACC higher than 98.77%, and FPR less than 0.69%. Next, the SMKFCM algorithm, with TPR, ACC is higher than 97.54% and FPR is less than 0.98%. As can be seen, GIT2SPFCM-PSO algorithm gives the highest accuracy, followed by SMKFCM, SFCM-PSO, IFCM, and DFCM algorithms, respectively. Table 3.3 is the average time of 10 runs by the five proposed algorithms on three datasets. SFCM-PSO algorithm has the lowest computation time, followed by DFCM, SMKFCM, GIT2SPFCM-PSO, and IFCM algorithms, respectively.

Table 3.3: Implementation time (s) of the proposed algorithms

Algorithm	Hanoi area	Quy Hop area	Vinh Phuc area
DFCM	235.465	196.254	188.264
IFCM	623.982	558.785	719.482
SMKFCM	285.522	209.095	276.541
SFCM-PSO	115.981	147.472	132.753
GIT2SPFCM-PSO	398.164	318.498	387.907

3.5.2 Landcover change detection

In this section, RS image in Bac Binh district, Binh Thuan province from 1988 to 2017 is used to assess the land cover change including Landsat-5 TM, Landsat-7 ETM+ and Landsat-8. Figure 3.1 shows the classification results according to 6 land covers by years; it can be seen a significant change in the land cover distribution. The land cover classification result using GIT2SPFCM-PSO algorithm into six classes by percentage (%) is shown in Table 3.4. Table 3.5 shows the accuracy of the proposed algorithms based on the labeled data. It

Table 3.4: Land cover classification results using GIT2SPFCM-PSO

Class	***1988***	***1994***	***2002***	***2009***	***2014***	***2017***
1	0.06(%)	0.09(%)	0.33(%)	0.42(%)	0.38(%)	0.32(%)
2	23.94(%)	24.69(%)	28.53(%)	28.22(%)	32.19(%)	37.51(%)
3	14.00(%)	13.21(%)	14.34(%)	14.02(%)	16.00(%)	14.10(%)
4	17.15(%)	15.42(%)	14.23(%)	13.72(%)	16.25(%)	14.80(%)
5	28.40(%)	26.98(%)	21.60(%)	24.40(%)	21.92(%)	17.62(%)
6	16.44(%)	19.59(%)	20.94(%)	19.18(%)	13.80(%)	15.61(%)

3.5 Experiments

For a multi-spectral image with d bands, each pixel will be characterized by d components on d gray bands which described as follows $X = [\mathrm{x}_1, \mathrm{x}_2, ... \mathrm{x}_n]$ with $x_i = (b_{i1}, b_{i2}, ..., b_{id})$. Experimental algorithms include SFCM, GSPFCM, SPFCM-W, SPFCM-SS, SMKFCM, SIIT2FCM, SFCM-PSO, GIT2SPFCM, and GIT2SPFCM-PSO.

3.5.1 Landcover classification

The dissertation tested the landcover classification on three types of satellite images, including images Landsat-7 ETM+ (Hanoi), Landsat-8 (Quy Hop), and Sentinel-2A (Vinh Phuc). In Table 3.1, the parameter values are achieved by GIT2SPFCM-PSO algorithm implementation. Table 3.2 shows the accuracy of the proposed algorithms calculated by the TPR, FPR, and ACC indicators compared with Erdas software on three experimental datasets. We can see that the accuracy of the landcover classification result when using Erdas software is the lowest compared to the five proposed algorithms on all data sets. The accuracy of two unsupervised algorithms (DFCM and IFCM) is lower than that of three semi-supervised algorithms (SMKFCM, SFCM-PSO, and

Table 3.1: Parameters achieved by GIT2SPFCM-PSO algorithm

Dataset	$\boldsymbol{m}$	m_1	m_2	η	η_1	η_2	$\boldsymbol{a}$	$\boldsymbol{b}$
Hanoi	2.21364	1.36534	3.26513	2.19874	1.47635	3.07366	0.52752	0.52463
Quy Hop	2.2876	1.4764	3.4565	2.1876	1.3768	3.3764	0.3764	0.3798
Vinh Phuc	2.2653	1.4762	3.0984	2.1987	1.6872	2.9875	0.7653	0.7759

Table 3.2: The accuracy of the proposed algorithms on three datasets

Dataset	Hanoi area (%)			Quy Hop area (%)			Vinh Phuc area (%)		
Algorithm	***TPR***	***FPR***	***ACC***	***TPR***	***FPR***	***ACC***	***TPR***	***FPR***	***ACC***
Erdas	89.52	1.36	89.18	91.61	1.09	90.99	88.76	0.87	89.14
DFCM	92.67	1.21	92.67	91.18	0.87	91.13	92.09	1.02	92.01
IFCM	93.71	1.09	93.39	93.26	1.12	93.25	94.64	0.68	94.32
SMKFCM	98.23	0.87	98.11	97.58	0.98	97.54	98.45	0.75	98.42
SFCM-PSO	95.48	0.99	95.21	96.83	0.79	96.84	95.83	0.89	95.78
GIT2SPFCM-PSO	99.08	0.58	99.02	98.97	0.52	98.77	99.15	0.69	99.13

PFCM model is the constrained optimization problem:

$$J_{m,\eta}(U,T,V,X,\gamma) = \sum_{i=1}^{c}\sum_{k=1}^{n}(a\mu_{ik}^{m} + b\tau_{ik}^{\eta})d_{ik}^{2} + \sum_{i=1}^{c}\gamma_i\sum_{k=1}^{n}(1-\tau_{ik})^{\eta} \tag{1.2}$$

Subject to the constraints:

$$m,\eta > 1; a,b > 0; 0 \le \mu_{ik},\tau_{ik} \le 1; \sum_{i=1}^{c}\mu_{ik} = 1; \sum_{k=1}^{n}\tau_{ik} = 1; 1 \le i \le c; 1 \le k \le n \tag{1.3}$$

1.1.2 Interval type-2 fuzzy c-means clustering

Definition 1.2. *A T2FS, denoted $\tilde{A}$, is characterized by a type-2 MF $\mu_{\tilde{A}}(x,u)$ where $x \in X$ and $u \in J_x \subseteq [0,1]$, i. e. ,*

$$\tilde{A} = \{((x,u), \mu_{\tilde{A}}(x,u)) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \tag{1.4}$$

or

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x} \mu_{\tilde{A}}(x,u))/(x,u), J_x \subseteq [0,1] \tag{1.5}$$

T2FSs are called an IT2FSs if the secondary MF $f_{x'}(u) = 1 \; \forall u \in J_x$ i. e. an IT2FS is defined as follows:

Definition 1.3. *An IT2FS $\tilde{A}$ is characterized by an interval type-2 MF $\mu_{\tilde{A}}(x,u) = 1$ where $x \in X$ and $u \in J_x \subseteq [0,1]$, i. e. ,*

$$\tilde{A} = \{((x,u), 1) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \tag{1.6}$$

IT2FCM is an extension of FCM algorithm by using two fuzziness parameters m_1, m_2 to make FOU, corresponding to upper and lower values of fuzzy clustering. The use of fuzzifiers gives different objective functions to be minimized as follows:

$$J_{m_1}(U,V,X) = \sum_{k=1}^{N}\sum_{i=1}^{C} u_{ik}^{m_1} d_{ik}^{2} \quad and \quad J_{m_2}(U,V,X) = \sum_{k=1}^{N}\sum_{i=1}^{C} u_{ik}^{m_2} d_{ik}^{2} \tag{1.7}$$

1.1.3 Some learning methods

This section covers some of the learning techniques used in the dissertation that can help improve the accuracy of clustering algorithms, including the semi-supervised learning method, kernel technique, spectral clustering, and particle swarm optimization.

1.1.4 Evaluation methods

There are two commonly used methods, including the internal evaluation and external evaluation. In this dissertation, both approaches are used to evaluate the quality of cluster results.

1.2 Related works

This section covers an overview of fuzzy clustering and type-2 fuzzy clustering. Some limitations of the previous methods are also mentioned and solutions to overcome these disadvantages.

1.3 Framework of remote sensing image analysis problem

Figure 1.5 shows the general framework of the proposed algorithms in the dissertation.

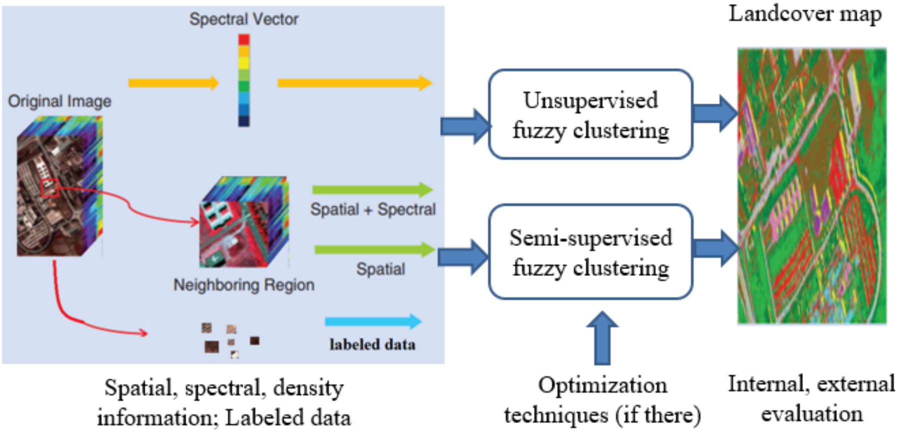
Spectral Vector
Original Image
Spatial + Spectral
Spatial
Neighboring Region
labeled data
Unsupervised fuzzy clustering
Semi-supervised fuzzy clustering
Landcover map
Spatial, spectral, density information; Labeled data
Optimization techniques (if there)
Internal, external evaluation

Figure 1.1: Framework of remote sensing image analysis problem

1.4 Chapter summary

Chapter 1 has introduced an overview of research issues, related background theories, and reviewing previous work related to the dissertation. Several commonly used methods to evaluate the accuracy of RS image classification results are also introduced. In the next chapter, the dissertation will present some improvements of FCM algorithm.

For each iteration of PSO algorithm, p_i and vel_i are updated as follows:

$$\begin{aligned} vel_i^{(t+1)} &= \omega * vel_i^{(t)} + c_1 * r_1 * (pBest_i^{(t)} - p_i^{(t)}) + c_2 * r_2 * (gBest^{(t)} - p_i^{(t)}) \\ p_i^{(t+1)} &= p_i^{(t)} + vel_i^{(t+1)} \end{aligned} \tag{3.30}$$

The hybrid algorithm between GIT2SPFCM and PSO is considered to minimum the objective function following:

$$F_{m_1,\eta_1,m_2,\eta_2}(U,T,V,X,\gamma) = \frac{F_{m_1,\eta_1}(U,T,V,X,\gamma) + F_{m_2,\eta_2}(U,T,V,X,\gamma)}{\min\limits_{i,j=1,\dots,c;i\neq j} \|v_i - v_j\|^2} \tag{3.31}$$

Steps to implement GIT2SPFCM-PSO algorithm show in 3.4. The computational complexity of GIT2SPFCM-PSO algorithm is $O(nd^2c^2T_{max})$.

Algorithm 3.4 GIT2SPFCM-PSO algorithm

```
**Input**: Dataset $X = \{x_k, x_k \in R^d, k = 1,...,n\}$, the labeled dataset $X^* = \{P_{is}, P_{is} \in R^d; s << n; i = 1,...,c\}$, the number of clusters $c(1 < c < n)$, fuzzifier parameters, $\epsilon$, and $T_{max}$, $t = 0$, $c_1, c_2, r_1, r_2, \omega$.
**Step 1**: Compute $V^* = [v_i^*]$ by Equation 3.15, $U^* = [\mu_{ik}^*]$ by Equation 3.16, $T^* = [\tau_{ik}^*]$ by Equation 3.17.
**Step 2**: Initialization
2.1 Initialize $V^{(0)} = [v_i^{(0)}], V^{(0)} \in R^{dxc}$ by using FCM algorithm.
2.2 Initialize the location particles $P^{(0)} = (p_1^{(0)}, p_2^{(0)}, ..., p_{c*d}^{(0)}, p_{c*d+1}^{(0)}, ..., p_{c*d+8}^{(0)})$ and the random values $m, m_1, m_2, \eta, \eta_1, \eta_2, a, b$.
2.3 Create the random velocity of particles: $vel_1^{(0)}, vel_2^{(0)}, ..., vel_{c*d}^{(0)}, vel_{c*d+1}^{(0)}, ..., vel_{c*d+8}^{(0)}$.
2.4 Compute $U^{(0)}$ by Equations 3.20, 3.21, 3.22.
2.5 Compute $T^{(0)}$ by Equations 3.23, 3.24, 3.25, 3.26.
2.6 Compute $F_{m_1,\eta_1,m_2,\eta_2}^{(0)}$ by Equation 3.31.
2.7 Let $pBest_i^{(0)} = p_i^{(0)}$, $gBest^{(0)}$ by Equation 3.29.
**Step 3**: $t = t + 1$
3.1 For each particle $i$.
+ Compute $vel_i^{(t+1)} = \omega * vel_i^{(t)} + c_1 * r_1 * (pBest_i^{(t)} - p_i^{(t)}) + c_2 * r_2 * (gBest^{(t)} - p_i^{(t)})$
+ Compute $p_i^{(t+1)} = p_i^{(t)} + vel_i^{(t+1)}$.
+ Compute $F_{m_1,\eta_1,m_2,\eta_2}^{(t)}$ by Equation 3.31.
+ Update $pBest_i^{(t)}$ by Equation 3.28.
+ Update $V^{(t)} = [v_i^{(t)}]$ and $m, m_1, m_2, \eta, \eta_1, \eta_2, a, b$ (if change).
3.2 Find the global best solution $gBest^{(t)}$ by Equation 3.29.
3.3 Update $U^{(t)}$ by Equations 3.20, 3.21, 3.22.
3.4 Update $T^{(t)}$ by Equations 3.23, 3.24, 3.25, 3.26.
3.5 **IF** $t > T_{max}$ **THEN** go to Output **ELSE** go to step 3.
**Output**: $V^{(t)}$, $U^{(t)}$, $T^{(t)}$, $m, m_1, m_2, \eta, \eta_1, \eta_2, a, b$. Defuzzification: Assign data $x_k$ to the $i^{th}$ cluster if $u_{ik} \geq u_{jk}, j = 1, ..., c; j \neq c$.
```

Algorithm 3.3 GIT2SPFCM algorithm

```
Input: Dataset $X = \{\mathrm{x}_k,\mathrm{x}_k \in \mathrm{R}^d,\ \mathrm{k}=1,...,\mathrm{n}\}$, the labeled data set $X^* = \{\mathrm{P_{is}},\mathrm{P_{is}} \in \mathrm{R}^d; \mathrm{s} << \mathrm{n}; \mathrm{i}=1,...,\mathrm{c}\}$, the number of clusters $c(1 < c < n)$, fuzzifier parameters $m_1, m_2, m, \eta_1, \eta_2, \eta$, and $T_{\max}$, $t=0$.
Step 1: Compute $V^* = [\mathrm{v}_i^*]$ by Equation 3.15, $U^* = [\mu_{ik}^*]$ by Equation 3.16, $T^* = [\tau_{ik}^*]$ by Equation 3.17.
Step 2: Initialize $V^{(t)} = [v_i^{(t)}], V^{(t)} \in R^{dxc}$ by choosing randomly from the input dataset $X$.
Step 3: Compute $U^{(t)}$ by Equations 3.20, 3.21, 3.22, ??, ??, ??.
Step 4: Compute $T^{(t)}$ by Equations 3.23, 3.24, 3.25, 3.26.
Step 5: $t = t+1$
5.1 Compute the centroids $v^R$ and $v^L$ use Equation 3.18.
5.2 Update the centroid matrix $V^{(t)} = [v_1^{(t)}, v_2^{(t)}, ..., v_C^{(t)}]$.
5.3 Update $U^{(t)}$ by Equations 3.20, 3.21, 3.22.
5.4 Update $T^{(t)}$ by Equations 3.23, 3.24, 3.25, 3.26.
5.5 Assign data $x_k$ to the $i^{th}$ cluster if $u_{ik} \geq u_{jk}, j = 1, ..., c; j \neq c$.
5.6 IF $max(\left\|U^{(t+1)} - U^{(t)}\right\| + \left\|T^{(t+1)} - T^{(t)}\right\|) \leq \varepsilon$ or $t > T_{max}$ THEN stop and go to Output
ELSE go to Step 5.
Output: The membership matrix $U$, $T$ and the centroid matrix $V$. Defuzzification: Assign the data pattern $x_k$ to the $i^{th}$ cluster if $u_{ik} \geq u_{jk}, j = 1, ..., c; j \neq c$.
```

3.4.2 Hybrid method of GIT2SPFCM and PSO

With satellite image data has M spectrum bands ($d = 3$ with RGB color image), the number of clusters is c, so the total number of particles initialized is $d * c + 8$ (see 3.27).

$$\underbrace{v_{11}, v_{12}, ...v_{1d}}_{V_1} \quad \underbrace{v_{21}, v_{22}, ...v_{2d}}_{V_2} \quad \cdots \quad \underbrace{v_{c1}, v_{c2}, ...v_{cd}}_{V_c} \quad \underbrace{m, m_1, m_2, \eta, \eta_1, \eta_2, a, b}_{\text{parameters}} \tag{3.27}$$

where $v_i = [\mathrm{v_{ij}}]$ is cluster centroids $(i = 1, ..., c; j = 1, ..., d)$ and $m, m_1, m_2, \eta, \eta_1, \eta_2$ are fuzzy and possibilistic parameters, and a, b are user-defined parameters.

Let $P = (p_1, p_2, ..., p_{c*d}, p_{c*d+1}, ..., p_{c*d+8})$ be the set of all particles position. Each particle will include the following information: p_i the current position of i^{th} particle; vel_i the current velocity of i^{th} particle; $pBest_i$ the personal best position of i^{th} particle. With the objective function F, then the personal best position of a particle at the time t is updated as:

$$pBest_i^{(t+1)} = \begin{cases} pBest_i^{(t)} & \text{if} \quad F(p_i^{(t+1)}) \geq F(pBest_i^{(t)}) \\ p_i^{(t+1)} & \text{if} \quad F(p_i^{(t+1)}) < F(pBest_i^{(t)}) \end{cases} \tag{3.28}$$

The best position of the population is denoted by g_{Best}:

$$\begin{aligned} gBest^{(t)} &= \{pBest_i^{(t)} | F(pBest_i^{(t)}) \\ &= \min\{F(pBest_1^{(t)}), F(pBest_2^{(t)}), ..., F(pBest_{c*d+4}^{(t)})\}\} \end{aligned} \tag{3.29}$$

Chapter 2

Fuzzy c-means clustering algorithm using density and spatial information

2.1 Introduction

This chapter will bring you to two unsupervised-methods improved from FCM algorithm. Algorithms are introduced including DFCM and IFCM.

2.2 Density fuzzy c-mean clustering

The key to determine a pixel belonging to a particular area is based on the similarity in spectral values. This measurement is calculated through a distance function in the color space d_{ik} between the pattern x_k and the centroid v_i. In that, the centroid will be in the samples that the density surrounding the sample data is large.

For the first step, the mean pattern $\bar{x}_j$ is computed by the following equation:

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{\mathrm{ij}}, j = \overline{1, d} \tag{2.1}$$

And standard deviation s_i:

$$s_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{\mathrm{ij}} - \bar{x}_i)^2}, j = \overline{1, d} \tag{2.2}$$

In which, $i = \overline{1, d}$, $X = \{x_1, x_2, ..., x_n\}, x_k \in R^d, k = \overline{1, n}$. Considering the surround of each data point is the m-dimensional box with a radius defined by the standard deviation is $r = \min\limits_{1<j<d}(s_j)$. Compute density D_i of pattern x_i:

$$D_i = \sum_{j=1}^{n} T(r - |x_j - x_i|) = \sum_{j=1}^{n} T(\Delta r); T(\Delta r) = \begin{cases} 1 & \Delta r \geq 0 \\ 0 & \Delta r < 0 \end{cases} \tag{2.3}$$

Call V is a set of pixels in order of density from high to low. Find pixel satisfying the condition: $D_i{}^* = \max\limits_{1 \leq j \leq d}(\mathrm{D}_i)$.

Put x_i into the result set V according to the following equations: $V = V \cup x_i$ and $X = X \backslash x_i$. If $X = \emptyset$ given a set of candidate points V , else back to finding D_i .

The initial centroids can be initialized by choosing in V according to the density of candidates. DFCM algorithm will have a computational complexity of $O(ndcT_{max})$.

Algorithm 2.1 Density-based fuzzy clustering algorithm (DFCM)

```
Input: Data set $X$ with $n$ data sample $X = \{x_1, x_2, ..., x_n\}, x_k \in R^d, k = \overline{1,n}$, the number of clusters is $C$, stop condition $\epsilon$.
Output: Set of result clusters
Step 1. Calculate sample expectations and standard deviations by Equation 2.1 and 2.2, the radius of the sphere $r = \min_{1<i<d} (s_i)$ in the m-dimensional space.
Step 2. Density calculation $D_i$ by Equation 2.3.
Step 3. Find $x_i$ by $D_i{}^* = \max_{1 \le i \le n} (\mathrm{D}_i)$, and assign $x_i$ to result set by $V = V \cup x_i$ and $X = X \backslash x_i$.
Step 4. Calculate $Y = \{\mathrm{x}_j, \mathrm{r} - |x_i - x_j| \geq 0\}$ and set $X = X \backslash Y$. If $X = \emptyset$ the go to Step 5, else go to Step 1.
Step 5. Given set of centroids $V = \{v_1, v_2, ..., v_C\}$.
Step 6. Use the fuzzy clustering algorithm to cluster with the initial centroids just found.
```

2.3 Spatial-spectral fuzzy c-mean clustering

The shape and structure of the cluster also have a certain influence on clustering results. To determine the degree of influence of the neighboring pixels for the center pixels, a local information measure M_i is defined on the basis of the distance $\|x_i - x_j\|$ and the attraction distance r_{ij}:

$$M_i = \sum_{j=1}^{P} \left(\|x_i - x_j\| \, r_{ij}\right)^{-1} / \sum_{j=1}^{P} r_{ij}^{-1} \tag{2.4}$$

Consider the local nxn mask and for sliding the mask on the image. Calculating the spatial information of the center pixel x_i based on the location of the center pixel x_i with the pixels x_j in the mask and the distance in color space $\|x_i - x_j\|$.

Set $r = max(r_{ij})_{\forall i,j}$ is the radius of the largest circle in which pixels that affect the central pixel. Next, without loss of generality, we standardized similar measurements on the following equation:

$$\overline{M}_i = \frac{M_i - \min{(M_i)_{\forall i}}}{\max(M_i)_{\forall i} - \min{(M_i)_{\forall i}}} \tag{2.5}$$

From above description, a new similarity measure is defined as follows:

$$s_{\mathrm{ij}} = \exp\left(-\frac{d^2(\mathrm{x}_i, \mathrm{x}_j)}{r^2}\right) \tag{2.6}$$

$J_{m_1,\eta_1}(U,T,V,X,\gamma)$ and $J_{m_2,\eta_2}(U,T,V,X,\gamma)$ may minimize if only:

$$\mu_{ik}^{(1)} = \begin{cases} \mu_{ik}^* + \dfrac{(1 - \sum\limits_{i=1}^{c} \mu_{ik}^*)\left[1/D_{ik}^2\right]^{1/(m_1-1)}}{\sum\limits_{i=1}^{c} \left[1/D_{ik}^2\right]^{1/(m_1-1)}} & if \dfrac{1}{\sum\limits_{j=1}^{C} (D_{ik}/D_{jk})} < \frac{1}{c} \\ \mu_{ik}^* + \dfrac{(1 - \sum\limits_{i=1}^{c} \mu_{ik}^*)\left[1/D_{ik}^2\right]^{1/(m_2-1)}}{\sum\limits_{i=1}^{c} \left[1/D_{ik}^2\right]^{1/(m_2-1)}} & otherwise \end{cases} \tag{3.20}$$

$$\mu_{ik}^{(2)} = \begin{cases} \mu_{ik}^* + \dfrac{(1 - \sum\limits_{i=1}^{c} \mu_{ik}^*)\left[1/D_{ik}^2\right]^{1/(m_1-1)}}{\sum\limits_{i=1}^{c} \left[1/D_{ik}^2\right]^{1/(m_1-1)}} & if \dfrac{1}{\sum\limits_{j=1}^{C} (D_{ik}/D_{jk})} \geq \frac{1}{c} \\ \mu_{ik}^* + \dfrac{(1 - \sum\limits_{i=1}^{c} \mu_{ik}^*)\left[1/D_{ik}^2\right]^{1/(m_2-1)}}{\sum\limits_{i=1}^{c} \left[1/D_{ik}^2\right]^{1/(m_2-1)}} & otherwise \end{cases} \tag{3.21}$$

Where

$$\begin{aligned} \bar{\mu}_i(x_k) &= max\{\mu_{ik}^{(1)}, \mu_{ik}^{(2)}\} \\ \underline{\mu}_i(x_k) &= \min\{\mu_{ik}^{(1)}, \mu_{ik}^{(2)}\} \end{aligned} \tag{3.22}$$

$$\tau_{ik}^{(1)} = \begin{cases} \left(\tau_{ik}^* + \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_1-1)}\right) / \left(1 + \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_1-1)}\right) & \tau_{ik} \geq \tau_{ik}^* \\ \left(\tau_{ik}^* - \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_1-1)}\right) / \left(1 - \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_1-1)}\right) & else \end{cases} \tag{3.23}$$

$$\tau_{ik}^{(2)} = \begin{cases} \left(\tau_{ik}^* + \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_2-1)}\right) / \left(1 + \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_2-1)}\right) & \tau_{ik} \geq \tau_{ik}^* \\ \left(\tau_{ik}^* - \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_2-1)}\right) / \left(1 - \left[\gamma_i/bD_{ik}^2\right]^{1/(\eta_2-1)}\right) & else \end{cases} \tag{3.24}$$

Where

$$\begin{aligned} \bar{\tau}_i(x_k) &= max\{\tau_{ik}^{(1)}, \tau_{ik}^{(2)}\} \\ \underline{\tau}_i(x_k) &= \min\{\tau_{ik}^{(1)}, \tau_{ik}^{(2)}\} \end{aligned} \tag{3.25}$$

For possibilistic membership grades:

$$\tau_i(x_k) = (\bar{\tau}_i(x_k) + \underline{\tau}_i(x_k))/2; i = 1, ..., c; k = 1, ..., n \tag{3.26}$$

The implementation steps of GIT2SPFCM algorithm are similar to IT2FCM, details of the steps are as follows: The computational complexity of GIT2SPFCM algorithm is $O(dcnlognT_{max})$.

```
Algorithm 3.2 SFCM-PSO algorithm
Input: Given a set of $n$ samples $X = \{\mathrm{x}_i\}_{i=1}^{n}$, where $A = A_1 \cup A_2 \cup ... \cup A_c$ is the set of labeled data samples, $A_i, i = \overline{1,c}$ is a set of labeled data samples for class $i$, $T_{max}$.
Output: $U = [u_{ik}]$.
Step 1: Initialize swarm
    1.1 Calculation $c$ centroids: $V^* = [\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_c]$ by Equation 3.10.
    1.2 Set the constants: Maximum loop number $T, t = 0, c_1, c_2, \omega, r_1, r_2, \varepsilon$.
    1.3 Create random locations: $v_1^{(0)}; v_2^{(0)}; ...; v_{c*b}^{(0)}$ and $v_{c*b+1}^{(0)}$ $(m^{(0)})$ within the limits from $v_{min}$ to $v_{max}$.
    1.4 Create random velocity: $vt_1^{(0)}; vt_2^{(0)}; ...; vt_{c*b}^{(0)}$ and $vt_{c*b+1}^{(0)}$ $(vt_m^{(0)})$ within the limits from $vt_{min}$ to $vt_{max}$.
    1.5 Calculate the value of $U$ by Equation 3.12.
Step 2: $t = t + 1$
    2.1 $v_i^{(t)} = v_i^{(t)} + vt_i^{(t)}, i = 1, ..., c * b + 1$
    2.2 Update $F$ by Equation 3.14.
    2.3 Update $P_{ibest}$ and $G_{ibest}$.
    2.4 $vt_i^{(t+1)} = \omega * vt_i^{(t)} + c_1 * r_1 * (P_{ibest} - v_i^{(t)}) + c_2 * r_2 * (G_{ibest} - v_i^{(t)}), i = 1, ..., c * b + 1$
    2.5 Update $U$ by Equation 3.12.
    2.6 IF $max(\left\| u_{ik}^{(t+1)} - u_{ik}^{(t)} \right\|) < \varepsilon$ or $(t > T_{max})$ THEN go to step 3 ELSE go to step 2.1.
Step 3: Given $U = [u_{ik}]$ and defuzzification and assign pixels to the cluster: if $u_{ik} > u_{jk}$ for $j = 1; 2; ...; c$ and then $x_k$ is assigned to cluster $i$.
```

The additional possibilistic MF is calculated based on a set of additional centroid V^* by PCM algorithm:

$$\tau_{ik}^* = 1/\left(1 + (b\left\|v_i^* - x_k\right\|)^{1/(\eta-1)}/\gamma_i\right) \tag{3.17}$$

Set $D_{ik}^2 = \|v_i - x_k\|^2 + \delta\|v_i - v_i^*\|^2$. The use of m_1, m_2 and η_1, η_2 gives different objective functions to be minimized as follows:

$$\begin{aligned} J_{m_1,\eta_1} &= \sum_{i=1}^{c}\sum_{k=1}^{n}(a\|\mu_{ik} - \mu_{ik}^*\|^{m_1} + b\|\tau_{ik} - \tau_{ik}^*\|^{\eta_1})D_{ik}^2 + \sum_{i=1}^{c}\gamma_i\sum_{k=1}^{n}(1 - \tau_{ik})^{\eta_1} \\ J_{m_2,\eta_2} &= \sum_{i=1}^{c}\sum_{k=1}^{n}(a\|\mu_{ik} - \mu_{ik}^*\|^{m_2} + b\|\tau_{ik} - \tau_{ik}^*\|^{\eta_2})D_{ik}^2 + \sum_{i=1}^{c}\gamma_i\sum_{k=1}^{n}(1 - \tau_{ik})^{\eta_2} \end{aligned} \tag{3.18}$$

Subject to the constraints:

$$\begin{aligned} &m_1, \eta_1, m_2, \eta_2 > 1; a, b > 0; \delta \geq 0; 0 \leq \mu_{ik}, \tau_{ik} \leq 1; \\ &\sum_{i=1}^{c}\mu_{ik} = 1; \sum_{k=1}^{n}\tau_{ik} = 1; 1 \leq i \leq c; 1 \leq k \leq n \end{aligned} \tag{3.19}$$

Theorem 3.1. *For $X = \{\mathrm{x}_k,\mathrm{x}_k \in \mathrm{R}^M, \mathrm{k} = 1,...,\mathrm{n}\}$, $m, \eta > 1; c \geq 1$, $\delta \geq 0$ and X contains at least c distinct data points. With the constraints 3.19 and Equation 3.18 then*

where s_{ij} showing pairwise similarities between pixels x_i and x_j; $d(x_i, x_j) = \|x_i - x_j\|$ is the Euclidean distance between x_i and x_j; r is the radius of the largest circle in which pixels that affect the central pixel.

With degree matrix D, it is built by adding local spatial information of each pixel, the degree of each pixel, d_i, is computed with:

$$d_i = \overline{M}_i * \sum_j s(i,j) \tag{2.7}$$

From the above description, the new Laplacian matrix L_{new}, is constructed using the new similarity matrix S and new degree matrix D:

$$L_{new} = D^{-1/2}SD^{-1/2} \tag{2.8}$$

Figure 2.1 is a diagram of the implementation steps of IFCM algorithm. The main steps of the proposed method are given in Algorithm 2.2. The computational complexity of IFCM algorithm is $O(n^3d)$.

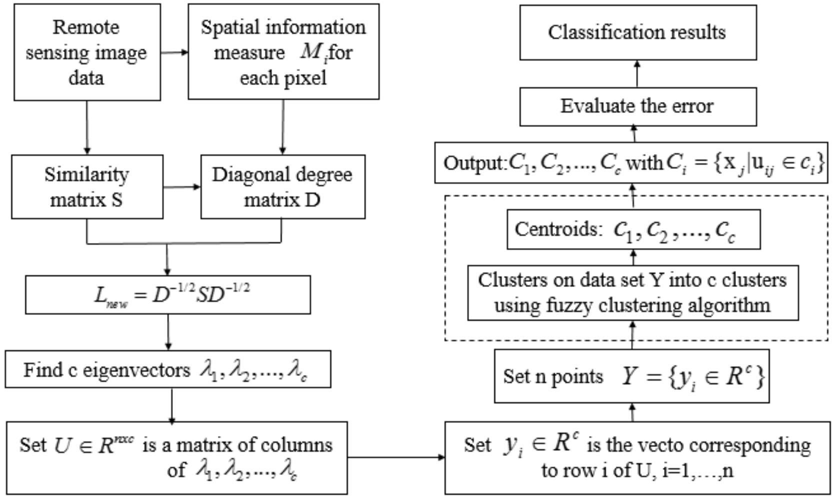
Remote sensing image data
Spatial information measure M_i for each pixel
Similarity matrix S
Diagonal degree matrix D
$L_{new} = D^{-1/2}SD^{-1/2}$
Find c eigenvectors $\lambda_1, \lambda_2, ..., \lambda_c$
Set $U \in R^{nxc}$ is a matrix of columns of $\lambda_1, \lambda_2, ..., \lambda_c$
Set $y_i \in R^c$ is the vecto corresponding to row i of U, i=1,…,n
Set n points $Y = \{y_i \in R^c\}$
Clusters on data set Y into c clusters using fuzzy clustering algorithm
Centroids: $C_1, C_2, ..., C_c$
Output:$C_1, C_2, ..., C_c$ with $C_i = \{x_j | u_{ij} \in c_i\}$
Evaluate the error
Classification results

Figure 2.1: Diagram of the implementation steps of IFCM algorithm

2.4 Experiments

2.4.1 SAR image segmentation

To testing the proposed algorithm, the SAR image is used to classify the oil spill on the sea. Oil stain classification results are shown in Figure 2.2.

```
Algorithm 2.2 Improved fuzzy c-means algorithm (IFCM)
Input: Matrix size used to calculate local spatial information, number of clusters $c$, $\epsilon$, $T_{max}$, $t = 0$.
Output: Clustering results $C_1, C_2, ..., C_c$ with $C_i = \{\mathrm{x_j}|\mathrm{u}_{ij} \in \mathrm{c}_i\}$.
Step 1. Calculate local information measure $M_i$ by Equation 2.5.
Step 2. Calculate a new similarity matrix $S$ by Equation 2.6.
Step 3. Calculate a diagonal degree matrix $D$ by Equation 2.7.
Step 4. Calculate a new matrix $L_{new}$ by Equation 2.8.
Step 5. Find the $c$ eigenvectors $\{e_1, e_2, ..., e_c\}$ of $L_{new}$, associated with the $c$ highest eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_c\}$ and define the c dimensional space $Y = (y_i)_{i=1,...,n} \in R^c$.
Step 6. Running fuzzy clustering algorithm on new space
```

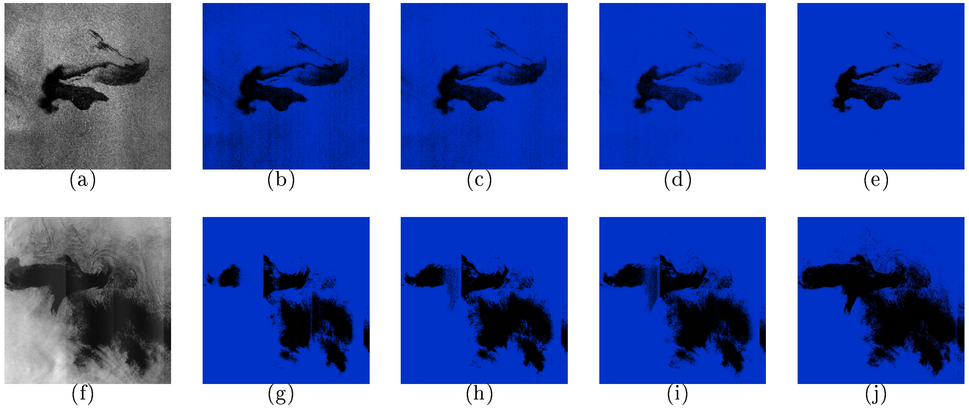

Figure 2.2: Oil spill classification results from the Envisat ASAR image in Gulf of Mexico on 26 April 2010 and 29 April 2010

Table 2.1: Indicators for evaluating oil stain classification results

Dataset	26 April 2010				29 April 2010			
Index	***FCM***	***ISC***	***DFCM***	***IFCM***	***FCM***	***ISC***	***DFCM***	***IFCM***
MSE	0.1871	0.1212	0.1189	**0.0986**	0.1761	0.1082	0.0864	**0.0082**
IQI	0.4595	0.7851	0.8876	**0.8968**	0.4862	0.6823	0.8635	**0.9447**
DI	0.0186	0.0561	0.0604	**0.0659**	0.0372	0.0598	0.0749	**0.0872**
CSI	1.1872	0.8725	0.7628	**0.6521**	1.5786	0.8873	0.7786	**0.5619**
SSE	32.7884	17.4663	16.4726	**15.3742**	15.6455	**8.4629**	8.4871	8.4631

Overall, the results classified according to the proposed algorithm for better

computation of derivatives u_{ik} and v_i, we have:

$$u_{\mathrm{ik}} = \left[\frac{1/(d^2(v_i, x_k) + d^2(v_i, v_i^*))}{\sum\limits_{j=1}^{c} [1/(d^2(v_i, x_k) + d^2(v_i, v_i^*))]^{1/(m-1)}} \right]^{1/(m-1)} \tag{3.12}$$

$$v_i = \frac{\sum\limits_{k=1}^{n} u_{ik}^m (v_i^* + x_k)}{2 \sum\limits_{k=1}^{n} u_{ik}^m} \tag{3.13}$$

Subject to $0 < \sum\limits_{k=1}^{n} u_{\mathrm{ik}} < n; 0 \le u_{\mathrm{ik}} \le 1; \sum\limits_{i=1}^{c} u_{\mathrm{ik}} = 1; 1 \le k \le n; 1 \le i \le c.$

In this study, the dissertation proposes a criterion for the minimum distance between cluster centers $\min_{i \neq j}\{d^2(v_i, v_j)\}$. A large value indicates that the clusters are more separated from each other. Therefore, the dissertation proposes an objective function as follows:

$$F = \frac{\sum\limits_{k=1}^{n} \sum\limits_{i=1}^{c} u_{ik}^m [\mathrm{d}^2(\mathrm{v}_i, \mathrm{x}_k) + \mathrm{d}^2(\mathrm{v}_i, \mathrm{v}_i^*)]}{\min_{i \neq j}\{\mathrm{d}^2(\mathrm{v}_i, \mathrm{v}_j)\}} \tag{3.14}$$

Details of implementation steps of SFCM-PSO is presented in algorithm 3.2. The computing complexity of SFCM-PSO algorithm is similar to FCM algorithm.

3.4 Hybrid method of interval type-2 SPFCM and PSO

3.4.1 General Interval type-2 SPFCM

Dataset $X = \{\mathrm{x}_k, \mathrm{x}_k \in \mathrm{R}^d, \mathrm{k} = 1, ..., \mathrm{n}\}$ with $X = X_1 \cup X_2$, $X_1 = [\mathrm{x}_1^*, \mathrm{x}_2^*, ..., \mathrm{x}_L^*]$ is the labeled dataset and $X_2 = [\mathrm{x}_{L+1}, \mathrm{x}_{L+2}, ..., \mathrm{x}_n]$ is the unlabeled dataset ($|X_1| << |X_2|$). Let c be the number of clusters, calculation c centroids $v_1^*, v_2^*, ..., v_c^*$ from the labeled pixel dataset and $V^* = [v_1^*, v_2^*, ..., v_c^*]$ is the set of additional cluster centroids, which is averaged from the labeled data as follows:

$$v_i^* = \sum_{s=1}^{m_i} P_{\mathrm{is}}/N_i \tag{3.15}$$

Where P_{is} is the s^{th} labeled pixel on the i^{th} cluster, N_i is the number of labeled pixels on the i^{th} cluster, $s = 1, 2, ..., N_i$; $i = 1, 2, ..., c$. The additional fuzzy MF is calculated based on a set of additional centroid V^* by FCM algorithm:

$$\mu_{ik}^* = 1/\sum_{z=1}^{c} (\frac{x_k - v_i^*}{x_k - v_z^*})^{2/(m-1)} \tag{3.16}$$

Algorithm 3.1 SMKFCM algorithm

```
Input: Given a set of $n$ patterns $X = \{x_i\}_{i=1}^{n}$, a set of kernel functions $\{K_k\}_{k=1}^{M}$, and the number of clusters $c$, $T_{max}$.
Output: Membership matrix $U = \{u_{ij}\}_{i,j=1}^{n,c}$ and weights $\{\omega_k\}_{k=1}^{M}$ for the kernels. To construct multiple kernels, we use the Gaussian kernel as $K_1$ and Polynomial kernel as $K_2$.
Step 1: Estimating centroids from the labeled data
    1.1 Extracting the labeled patterns from the dataset.
    1.2. Calculating the rudimentary centroids $V^* = [v_i^*], v_i^* \in R^n$ from labeled patterns.
Step 2: Initialization
    2.1 Choose fuzzifier $m, (1 < m)$, error $\epsilon$.
    2.2 Initialize membership matrix $U^{(0)}$.
Step 3: $t = t + 1$
    3.1 Calculate constants $\beta_j$ by Equation 3.7.
    3.2 Update weights $\omega_k$ by Equation 3.5.
    3.2 Calculate the distance in kernel space $d_{ij}$ by Equation 3.8.
    3.3 Update memberships $U^{(t)}$ by Equation 3.4.
    3.4 IF $(|U^{(t)} - U^{t-1}|) < \epsilon$ or $t > T_{max}$ THEN go to step 4 ELSE go to step 3.
Step 4: Report results clustering.
    4.1. Return $(t)$ and $\omega_k$ with $k = 1, 2, ...,$.
    4.2. Assign a pattern to a cluster and report the results of clustering.
```

A_i is the set of pixels that have been labeled for the i^{th} cluster, with $i = 1, ..., c$. Calculation c centroids by the following formula:

$$v_i^* = \sum_{j=1}^{|A_i|} p_j(A_i) / |A_i| \tag{3.10}$$

In which, $|A_i|$ is the number of labeled pixels for the i^{th} cluster, p_j is the j^{th} pixel in A_i.

The objective function J_m of FCM algorithm is changed as follows:

$$J_m = \sum_{k=1}^{n}\sum_{i=1}^{c} u_{ik}^m[\mathrm{d}^2(\mathrm{v}_i,\mathrm{x}_k)+\mathrm{d}^2(\mathrm{v}_i,\mathrm{v}_i^*)], 1 < m < \infty \tag{3.11}$$

With $d(v_i, x_k)$ is the euclidean distance between the pixel x_k and the cluster centroid v_i and $d(v_i, v_i^*)$ is the distance between the calculated cluster centroid and the desired cluster centroid, cluster results are good when this distance is small.

To minimize the objective function J_m, based on the Lagrange method by

results than the algorithms FCM, ISC and DFCM (Table 2.1).

2.4.2 *Landcover classification*

The proposed method tests on the Landsat 7-ETM+ image taken at Lam Dong province. Based on the value of the clustering quality index (on the Table 2.2), most of the cases showed IFCM algorithm for clustering results better than algorithms DFCM, ISC and FCM.

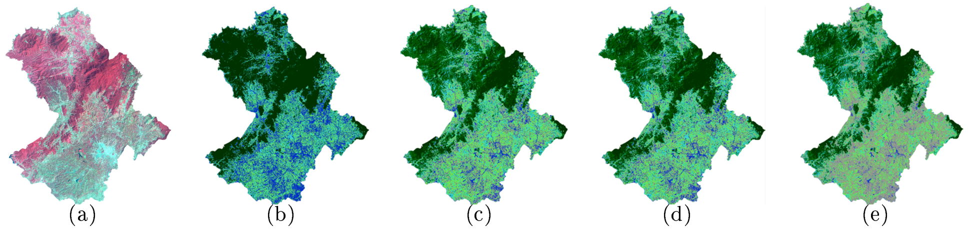

Figure 2.3: Color image and classification results of Lamdong area

Table 2.2: Indicators for evaluating classification results of Lamdong area

Index	FCM	ISC	DFCM	IFCM
MSE	0.1763	0.1075	0.0982	**0.0918**
IQI	0.5623	0.6732	0.7849	**0.8721**
DI	0.0123	0.0365	0.0428	**0.0452**
CSI	1.2512	0.7784	**0.7750**	0.7751
SSE	98.6389	78.8599	52.8752	**46.3986**

2.5 Summary

This chapter presents two unsupervised fuzzy clustering algorithms, DFCM and IFCM. The main idea of DFCM algorithm is to use density information as a preprocessing step to select initial centroids. IFCM algorithm based on local information and spectral clustering to make data separation better. In the next chapter, the dissertation presents the semi-supervised multiple kernel fuzzy c-means clustering algorithm and hybrid algorithms between semi-supervised fuzzy clustering and PSO technique.

Chapter 3

Improved fuzzy c-means clustering algorithm with semi-supervision

3.1 Introduction

In this chapter, the dissertation presents three semi-supervised fuzzy clustering techniques, including SMKFCM, SFCM-PSO and GIT2SPFCM-PSO.

3.2 Semi-supervised multiple kernel FCM clustering

The idea of the approach is to use the rudimentary centroids V^* to adjust centroids in the clustering process.

A semi-supervised multiple kernel fuzzy c-means clustering (SMKFCM) algorithm is extended from FCM by combining different kernels and semi-supervised method to obtain better results. SMKFCM maps the data from the feature space into kernel space H by using transform functions: $\psi = \{\psi_1, \psi_2, ..., \psi_M\}$ where $\psi_k(x_i)^T\psi_k(x_j) = K_k(x_i, x_j)$ and $\psi_k(x_i)^T\psi_{k'}(x_j) = 0\,|k \neq k'$

The prototypes v_i is constructed in the kernel space, the general framework of SMKFCM aims to minimize the objective function:

$$J_m(U, v) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{\mathrm{ij}}^m \left(\|\psi(x_j) - \psi(v_i)\|^2 + \|\psi(v_i^*) - \psi(v_i)\|^2\right) \tag{3.1}$$

In which, $\sum\limits_{i=1}^{c} u_{\mathrm{ij}} = 1$, n is the number of patterns, c is the number of clusters, $\psi(x) = \omega_1\psi_1(x) + \omega_2\psi_2(x), ..., \omega_M\psi_M(x)$.

Subject to $\omega_1 + \omega_2 + \omega_M = 1$ and $\omega_k \geq 0, \forall k$, where v_i is the centroid of the i^{th} cluster in the kernel space, $(\omega_1, \omega_2, ..., \omega_M)$ is a vector of weights for features, respectively. The distance d_{ij} concerns the j^{th} data (pattern) and thei^{th} prototype:

$$\|\psi(x_j) - \psi(v_i)\|^2 = (\psi(x_j) - \psi(v_i))^T(\psi(x_j) - \psi(v_i)) \tag{3.2}$$

Optimizing the objective function 3.1 we have:

$$v_i = \sum_{j=1}^{n} u_{\mathrm{ij}}^m(\psi(x_j) + v_i^*) \Big/ 2\sum_{j=1}^{n} u_{\mathrm{ij}}^m \tag{3.3}$$

$$u_{ij} = \frac{\left(\frac{1}{m((\psi(x_j)-v_i)^2+(v_i^*-v_i)^2)}\right)^{1/(m-1)}}{\sum\limits_{i=1}^{c}\left(\frac{1}{m((\psi(x_j)-v_i)^2+(v_i^*-v_i)^2)}\right)^{1/(m-1)}} \tag{3.4}$$

$$\omega_k = \frac{\beta_j + 2\sum\limits_{i=1}^{c} u_{\mathrm{ij}}^m v_i \psi_k(x_j)}{2\sum\limits_{i=1}^{c} u_{\mathrm{ij}}^m \psi_k^T(x_j)\psi_k(x_j)} \tag{3.5}$$

Now it can calculate the distance d_{ik} concerns the j^{th} data and the i^{th} prototype as:

$$d_{\mathrm{ij}}^2 = \|\psi(x_j) - \psi(v_i)\|^2 = \psi^T(x_j)\psi(x_j) - 2\psi(x_j)\psi(v_i) + \psi^T(v_i)\psi(v_i) \tag{3.6}$$

By replacing the v_i in Equation 3.3 and $\psi^T(x)\psi(y) = K(x, y) = \sum\limits_{k=1}^{M} \omega_k k_k(x, y)$ to the above equations and $\omega_1 + \omega_2 + \omega_M = 1$ and after some mathematical transformations, we have:

$$\beta_j = 2\sum_{k=1}^{M}\sum_{i=1}^{c} u_{\mathrm{ij}}^m K_k(x_j, x_j)\left(1 - \sum_{k=1}^{M}\frac{\sum\limits_{i=1}^{c}\sum\limits_{j=1}^{n} u_{\mathrm{ij}}^m(K_k(x_j, x_j) + K_k(x_j, v_i^*))}{2\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{c} u_{\mathrm{ij}}^m K_k(x_j, x_j)}\right) \tag{3.7}$$

$$d_{\mathrm{ij}}^2 = \frac{\beta_j + \sum\limits_{i=1}^{c}\sum\limits_{j=1}^{n} u_{\mathrm{ij}}^m(K_k(x_j, x_j) + K_k(x_j, v_i^*))}{2\sum\limits_{j=1}^{n} u_{\mathrm{ij}}^m K_k(x_j, x_j)} \tag{3.8}$$

To construct multi-kernel, we consider Gaussian kernel as K_1 and Polynomial kernel as K_2:

$$K_1(x, y) = \exp(-\|x - y\|^2/r^2), K_2(x, y) = (x^T y + d)^p \tag{3.9}$$

Where $r, d \in R^+, p \in N^+$.

The detailed steps of SMKFCM algorithm are described in 3.1. The computational complexity of SMKFCM is $O(n^2dcM)$ per iteration with M is the multiplier used.

3.3 Hybrid method of semi-supervised FCM and PSO

Usually, the supervised clustering technique requires large amounts of labeled data for training. In cases where labeled data is limited; the method often used is a semi-supervised clustering method.