Journal of Electrical and Electronic Engineering
2024, Vol. 12, No. 2, pp. 36-47
https://doi.org/10.11648/j.jeee.20241202.12

SciencePG
Science Publishing Group

Research Article

MST Initialization Based Intuitionistic Fuzzy c Means Clustering Using LINEX Hellinger Distance and Its Applications

Nithya M.[1], K. Bhuvaneswari Balasubramaium[1], Senthil S.[2,*]

[1]Department of Mathematics, Mother Teresa Women's University, Kodaikanal, India

[2]Department of Economics and Statistics, Integrated Child Development Services, Dindigul, India

Abstract

Due to the uncertainty and fuzziness of information, the traditional clustering analysis method sometimes cannot meet the requirement in practice. The clustering method based on intuitionistic fuzzy set has attracted more and more scholars attention nowadays. This paper discusses the intuitionistic fuzzy C-means clustering algorithm. There are a number of clustering techniques developed in the past using different distance/similarity measure. In this paper, we proposed a improved edge density minimal spanning tree initilization method using LINEX hellinger distance based weighted LINEX intuitionistic fuzzy c means clustering. IFCM considered an uncertainty parameter called hesitation degree and incorporated a new objective function which is based upon intutionistic fuzzy entropy in the conventional Fuzzy C-means. The clustering algorithm has membership and non membership degrees as intervals. Information regarding membership and typicality degrees of samples to all clusters is given by algorithm. Furthermore, the algorithm is extended for calculating membership and updating prototypes by minimizing the new objective function of weighted LINEX intuitionistic fuzzy c-means. Finally, the developed algorithms are illustrated through conducting experiments on random dataset, partition coefficient and validation function are used to evaluate the validity of clustering also this paper compares the results of proposed method with the results of existing basic intuitionistic fuzzy c-means.

Keywords

Intuitionistic Fuzzy C-means, Edge Density, Minimal Spanning Tree, Hellinger Distance, LINEX Function

1. Introduction

Clustering helps in finding natural boundaries in the data whereas fuzzy clustering can be used to handle the problem of vague boundaries of clusters. In fuzzy clustering, the requirement of crisp partition of the data is replaced by a weaker requirement of fuzzy partition, where the associations among data are represented by fuzzy relations. Clustering is the process of assigning data objects into a set of disjoint groups called clusters so that objects in each cluster are more similar to each other than objects from different clusters. Clustering algorithms work by assigning objects to a group if they show high level of similarity and by assigning objects to different groups if they are distinguished from

*Corresponding author: senthilsnc83@gmail.com (Senthil S.)

Received: 17 May 2024; **Accepted:** 5 June 2024; **Published:** 27 August 2024

each other.

Atanassov [1] presented "Intuitionistic fuzzy sets" that considers vulnerability in the definition of the membership value. Different approaches have been designed with the Intuitionistic theory for the process of image segmentation. Chaira [3] has given a novel Intuitionistic fuzzy c-means (IFCM) clustering algorithm for the process of partitioning an image into segments. It includes hesitation degree and a new parameter termed as intuitionistic fuzzy entropy (IFE) in the objective function. Although this strategy utilizes the irregular initialization of cluster prototype which produces inaccurate results when chosen randomly. This method does not include the local information and the spatial information, hence is sensitive to disturbances and other image artifacts [2, 10]. In order to make clustering more robust, Intuitionistic Fuzzy C-means clustering [6] assume that a pixel belongs to multiple clusters with different membership degrees [12]. These algorithms utilized only intensity of pixel as the only feature for the segmentation of images and failed to classify noisy pixels accurately. The pixels in an image are exceedingly associated, i. e. the every pixel in the prompt vicinity have about the equivalent feature information unless there is some curve or contour. Subsequently, integrating spatial information along with the membership value results in more homogenous regions as compared to other methods. However, these techniques use the arbitrary initialization of cluster centers which give inaccurate outcomes and more time for optimization.

A recent development is to use kernel method to construct the evolutionary kernel intuitionistic fuzzy c-means clustering algorithm, the new probability similarity measurement method and clustering technique are also proposed to design adaptive weights for intuitionistic fuzzy distances [7]. In Zhao F's research [13] the intuitionistic fuzzy set and rough set are combined with the statistical feature extraction technology. The intuitionistic fuzzy set is used to extract the regions of interest, and the gray scale co-occurrence matrix is used for feature extraction. Pop PC [9] proposed clustering algorithms based on the minimum spanning tree (MST) are able to detect clusters with arbitrary shapes. In Cheng D's research [4] the algorithm uses a new distance between local density peaks based on shared neighbors to construct a minimum spanning tree on the local density peaks, which excludes the interference of noise points and reduces the running time of MST-based clustering algorithms. A modified non-membership function to generate intuitionistic fuzzy set, which highlights the effect of uncertainty and makes good use of image information. A method of determining initial clustering centers based on pixel characteristics is also proposed by Jun Kong, Jian Hou, Min Jiang & Jinhua Sun [5]. Intuitionistic fuzzy C-means (IFCM) algorithm, as a successful extension and variant of FCM, has attracted extensive attention and has been widely used in many fields such as image processing and pattern recognition by [8, 11, 14]. Although the above methods are claimed to be robust to noise, they are confronted with the problem of selecting the parameters that control the role of the spatial constraints.

In this paper, to overcome the defects of the algorithms as mentioned, we proposed weighted–LINEX HELLINGER distance using IFCM, which requires the determination of the edge density and degree of minimal spanning, named the MST using weighted LINEX_IFCM. The intuitionistic fuzziness is embedded into the calculation process of similarity between the pixel and cluster centers, which achieves more accurate segmentation in the organization boundary. The algorithm is minimal spanning tree initialization method using LINEX HELLINGER_IFCM, which helps to speed up the convergence of the algorithm.

2. Minimal Spanning Tree Algorithm

In this section, we proposed LINEX HELLINGER distance measures for construct the minimal spanning tree.

2.1. Hellinger Distance

Let M and N discrete probabilistic distributions with $M=(m_1, m_2, \ldots\ldots m_k)$ and $N=(n_1, n_2, \ldots\ldots n_k)$. Then Hellinger distance will be

$$d_H(M,N)=\frac{1}{\sqrt{2}} * \sqrt{\sum_{i=1}^{k}\left(\sqrt{m_i}-\sqrt{n_i}\right)^2}$$

So Hellinger distance is directly related to the Euclidean norm [17], $d_H(M,N)=\frac{1}{\sqrt{2}} * \left\| \sqrt{M}-\sqrt{N} \right\|^2$

The Hellinger distance posseses the following characteristics:

i. Its values range between 0 and 1, where 0 signifies complete similarity between two distributions and 1 indicates complete dissimilarity.
ii. When two distributions are very similar, the Hellinger distance approcahes.
iii. The Hellinger distance is symmetric, meaning $d_H(M,N)=d_H(N,M)$. Compared to other distance metrices, such as Kullback-Leibler (KL) divergence ot total variation distance, Hellinger distance is more robust to outliers and in certain cases more accessible to compute. It finds widespread application in probability distribution comparisons and model fitting.

2.2. LINEX HELLINGER Distance Based MST

Given the grayscale point set D, the hierarchical methods starts by constructing a minimal spanning tree (MST) from the points in D. In $x=(x_1, x_2, \ldots\ldots\ldots x_n)^{\perp}$ and $y=(y_1, y_2, \ldots\ldots\ldots y_n)^{\perp}$ are two points of a MST and $e(x,y)$ is an edge between x and y then LINEX Hellinger distance function between x and y is denoted by $d(x,y)$ and calculated using equation (1)

$$d_{LINEX}(x,y)=\exp\left(-\frac{1}{\sqrt{2}}\left\|\sqrt{x}-\sqrt{y}\right\|\right)+\left(\frac{1}{\sqrt{2}}\left\|\sqrt{x}-\sqrt{y}\right\|\right)-1 \quad (1)$$

2.3. Edge Density Weight of MST

$T=(V,E)$ is the minimum spanning tree generated from dataset X of size N, where $V=\{v_1,v_2,\ldots\ldots\ldots v_N\}$ are vertices corresponding to objects, $E=\{e_1,e_2,\ldots\ldots\ldots e_{N-1}\}$ are edges corresponding to connection relationship between objects.

Edge density weight

$$EDW=d_{LINEX}(x,y)*\sqrt{\deg(x)*\deg(y)} \quad (2)$$

where

$$d_{LINEX}(x,y)=\exp\left(-\frac{1}{\sqrt{2}}\left\|\sqrt{x}-\sqrt{y}\right\|\right)+\left(\frac{1}{\sqrt{2}}\left\|\sqrt{x}-\sqrt{y}\right\|\right)-1,$$

deg (x) = degree of x from MST and

deg (y) = degree of y from MST.

2.4. Algorithm for Optimal Number of Clusters

The Algorithm: LINEX HELLINGER BASED MST

Input: Data points

Output: optimal number of cluster centers

Let e1 be an edge in the LINEX measure MST constructed from Data points

Let S_T be the set of disjoint subtrees of LINEX measure MST.

1. Create a node v, for each data points.
2. Compute the edge weight using equation (1)
3. Construct an LINEX measure MST from 2.
4. Compute the Edge density weight using equation (2)
5. $S_T=\phi$, $n_c=1$, $C=\phi$
6. For each $e1\in MST1$.
7. To remove longest edge or $w>\bar{w}+\left(\frac{\max(EDW)-\bar{w}}{2}\right)$ from MST
8. $S_T=S_T\cup\{T'\}$// T' is new disjoint subtrees (regions).
9. $n_c=n_c+1$.
10. Compute the center c_i of T_i using average of points.
11. $C=\cup T_i\in S_T\{c_i\}$.
12. $\delta(T_i)=$ Minimum standard deviation of edge density T_i.
13. $\Delta(T_i)=$ Maximum standard deviation of edge density T_i.
14. $CS=\dfrac{\delta(T_i)}{\Delta(T_i)}$
15. Until $CS<\bar{w}+\left(\frac{\max(EDW)-\bar{w}}{2}\right)$
16. Update the clusters points, repeat step 7 to step 14.
17. Finally we obtain the cluster centers.

3. Formulation of LINEX HELLINGER Distance Based Intuitionistic Fuzzy C Means Clustering

3.1. Intuitionistic Fuzzy C-means (IFCM) Algorithm

Intuitionistic fuzzy set given by Atanassov [1] considers both membership $\mu(x),\ x\in X$ and non-membership $\nu(x),\ x\in X$. An intuitionistic fuzzy set A in X, is written as

$$A=\{x,\mu_A(x),\ \nu_A(x)\backslash x\in X\}$$

where $\mu_A(x)\to[0,1]$, $\nu_A(x)\to[0,1]$ are the membership and non-membership degrees of an element in the set A with the condition $0\le\mu_A(x)+\nu_A(x)\le1$ when $\nu_A(x)=1-\mu_A(x)$ for every x in the set A, then the set A becomes a fuzzy set. Also indicated a hesitation degree, $\pi_A(x)$ which arises due to lack of knowledge in defining the membership degree of each element x in the set A and is given by

$$\pi_A(x)=1-\mu_A(x)-\nu_A(x),\ 0\le\pi_A(x)\le1$$

In [1] intuitionistic fuzzy c-means, minimizes the objective function as:

$$J_{IFCM}=\sum_{i=1}^{C}\sum_{k=1}^{N}u_{ik}^{*m}\left\|x_i-v_k\right\|^2+\sum_{i-1}^{C}\pi_i^{*}e^{1-\pi_i^{*}} \quad 1<m<\infty \quad (3)$$

$u_{ik}^{*}=u_{ik}+\pi_{ik}$, where u_{ik}^{*} denotes the intuitionistic fuzzy membership and

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left[\frac{\| x_i - v_k \|}{\| x_i - v_j \|} \right]^{\frac{2}{m-1}}} \tag{4}$$

$$\pi_{ik} = 1 - u_{ik} - \left(\frac{1-u_{ik}}{1+\lambda u_{ik}} \right), \lambda > 0 \tag{5}$$

$$v_k = \frac{\sum_{i=1}^{N} u_{ik}{}^m x_i}{\sum_{i=1}^{N} u_{ik}{}^m} \tag{6}$$

$$\pi_i{}^* = \frac{1}{N} \sum_{k=1}^{N} \pi_{ik} \tag{7}$$

This iteration will stop when

$$\max_{ij} \left\{ \left| u_{ik}{}^{*k+1} - u_{ik}{}^{*k} \right| \right\} < \in,$$

where $\in$ is a termination criterion between 0 and 1, where as k is the iteration steps. This procedure converges to a local minimum or a saddle point of J_{IFCM}.

3.2. LINEX HELLINGER Distance Based Intuitionistic Fuzzy C-means (IFCM) Algorithm

In this section, we want to use the LINEX HELLINGER distance based Intuitionistic fuzzy c-means (IFCM) algorithm when the over estimating and the under estimating are not of the same importance. The procedures are the same as a Intuitionistic fuzzy c-means algorithm. All the entities are assigned to their nearest centroid from MST, using a LINEX loss function as the dissimilarity distance. The procedure continues until there is no change in clusters. Now consider the following optimization problem,

Objective function:

$$J_{IFCM}(U,C) = \sum_{i=1}^{c} \sum_{j=1}^{n} u^{*m}{}_{ij} L_{WLINEX}(x_j, v_i + \sum_{i=1}^{c} \pi_i{}^* e^{1-\pi_i{}^*} \quad 1 < m < \infty \tag{8}$$

where $d_{LINEX}(x,y) = w(x_j, v_i) \times \exp\left(-\frac{1}{\sqrt{2}} \left\| \sqrt{x} - \sqrt{y} \right\| \right) + \left(\frac{1}{\sqrt{2}} \left\| \sqrt{x} - \sqrt{y} \right\| \right) - 1$

$w(x_j, v_i) = \sqrt{\deg(x_j) * \deg(v_i)}$

$\deg(x_j) = \text{degree of } x_j \text{ from MST}$ and

$\deg(v_i) = \text{degree of } v_i \text{ from MST}$

3.3. Updating Membership

To obtain equation for calculating membership we minimizing the objective function $J_{WIFCM}(U,V)$ with constraint conditions

$$0 \le u_{ik} \le 1 \ \ \forall\ i = 1,2,\ldots\ldots\ldots c,\ \ k = 1,2,\ldots\ldots\ldots\ldots n$$

$$\sum_{i=1}^{c} u_{ik} = 1 \ \ \forall\ k = 1,2,\ldots\ldots\ldots\ldots n.$$

$$0 \le \sum_{k=1}^{n} u_{ik} \le n \ \ \forall\ i = 1,2,\ldots\ldots\ldots c$$

Can be solved by using the Lagrangian multiplication as follows:

To find min $J_{WIFCM}(U,V)$ it is sufficient to minimize the following inner sum for fixed k:

put $L_{ik} = L_{WLINEX}(u_{ik}, v_{ij})$

$$\text{Let } B = \left\{ \begin{array}{l} u_k = (u_{1k}, u_{2k}, \ldots\ldots u_{ck}) \in R^c \\ \text{such that } \sum_{i=1}^{c} u_{ik} = 1,\ 0 \le u_{ik} \le 1 \end{array} \right\}$$

and $g(u_k) = \sum_{i=1}^{c} u^m{}_{ik} L_{ik} - \lambda \left(\sum_{i=1}^{c} u_{ik} - 1 \right)$

(λ, u_k) is stationary for F only if $\Delta_{\lambda, u_k} F(\lambda, u_k) = 0,\ \ 0 \in R^c$ that yields to

$$\frac{\partial F}{\partial \lambda}(\lambda, u_k) = \sum_{i=1}^{c} u_{ik} - 1 = 0 \tag{9}$$

$$\frac{\partial F}{\partial u_{ik}}(\lambda, u_k) = m u^{m-1}{}_{ik} L_{ik} - \lambda = 0 \tag{10}$$

$$m u^{m-1}{}_{ik} L_{ik} - \lambda = 0$$

$$m u^{m-1}{}_{ik} L_{ik} = \lambda$$

$$u^{m-1}{}_{ik} = \frac{\lambda}{m L_{ik}}$$

$$u_{ik}=\left[\frac{\lambda}{m L_{ik}}\right]^{\frac{1}{m-1}}$$

Then

$$u_{ik}=\left[\frac{\lambda}{m}\right]^{\frac{1}{m-1}}\left[\frac{1}{L_{ik}}\right]^{\frac{1}{m-1}} \tag{11}$$

From (9) $\sum_{i=1}^{c} u_{lk}=\sum_{l=1}^{c}\left[\frac{\lambda}{m}\right]^{\frac{1}{m-1}}\left[\frac{1}{L_{ik}}\right]^{\frac{1}{m-1}}=1$

$$\Rightarrow\left[\frac{\lambda}{m}\right]^{\frac{1}{m-1}}=\frac{1}{\sum_{l=1}^{c}\left[\frac{1}{L_{ik}}\right]^{\frac{1}{m-1}}} \tag{12}$$

Substitute (12) in (11). we obtain

$$u_{ik}=\frac{1}{\sum_{l=1}^{c}\left[\frac{1}{L_{ik}}\right]^{\frac{1}{m-1}}}\left[\frac{1}{L_{ik}}\right]^{\frac{1}{m-1}}=\left[\sum_{l=1}^{c}\left[\frac{L_{ik}}{L_{lk}}\right]^{\frac{1}{m-1}}\right]^{-1} \quad u_{ik}=\left[\sum_{l=1}^{c}\left[\frac{L_{WLINEX}(x_k,v_i)}{L_{WLINEX}(x_k,v_l)}\right]^{\frac{1}{m-1}}\right]^{-1} \tag{13}$$

The general equation is used to obtain membership ranks for objects in data for finding meaningful groups.

3.4. Obtaining Cluster Prototype Updating

To find min $J_{WLINEX}(U,V)$ it is sufficient to minimize the following inner sum for fixed i:

$$\sum_{k=1}^{n} u^{m}_{ik}\, w_{ji}\left[\exp\left(-\frac{1}{\sqrt{2}}\left(\sqrt{x_{kj}}-\sqrt{v_{ij}}\right)\right)+\left(\frac{1}{\sqrt{2}}\left(\sqrt{x_{kj}}-\sqrt{v_{ij}}\right)\right)-1\right]$$

Taking the partial derivative of objective function with respect to v_{ij} and setting the result to zero, we have the general form of updating center as

$$\frac{\partial\, WLINEX\,(U,V)}{\partial v_{ij}}=0$$

$$\frac{1}{2\sqrt{2}\sqrt{v_{ij}}}\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}\times\exp\left(-\frac{1}{\sqrt{2}}\left(\sqrt{x_{kj}}-\sqrt{v_{ij}}\right)\right)-\frac{1}{2\sqrt{2}\sqrt{v_{ij}}}\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}=0$$

$$\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}\times\exp\left(-\frac{1}{\sqrt{2}}\left(\sqrt{x_{kj}}-\sqrt{v_{ij}}\right)\right)=\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}$$

$$\Rightarrow e^{\frac{\sqrt{v_{ij}}}{\sqrt{2}}}=\frac{\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}}{\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}\times e^{-\frac{\sqrt{x_{kj}}}{\sqrt{2}}}}$$

$$\therefore\quad \sqrt{v_{ij}}=\sqrt{2}\,\log\left[\frac{\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}}{\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}\times e^{-\frac{\sqrt{x_{kj}}}{\sqrt{2}}}}\right]$$

$$\therefore\quad v_{ij}=2\left(\mathrm{Log}\left[\frac{\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}}{\sum_{k=1}^{n} u^{m}_{ik}\times w_{ji}\times e^{-\frac{\sqrt{x_{kj}}}{\sqrt{2}}}}\right]\right)^{2} \tag{14}$$

where $u_{ik}^{*}=u_{ik}+\pi_{ik}$, where u_{ik}^{*} denotes the intuitionistic fuzzy membership and

$$\pi_{ik}=1-u_{ik}-\left(\frac{1-u_{ik}}{1+\lambda u_{ik}}\right),\quad \lambda>0 \tag{15}$$

$$\pi_{i}^{*}=\frac{1}{N}\sum_{k=1}^{N}\pi_{ik} \tag{16}$$

The LINEX HELLINGER distance is suitable for clustering in which it can actually induce the necessary conditions.

This iteration will stop when $\max_{ij}\left\{ \left| u_{ik}^{*k+1} - u_{ik}^{*k} \right| \right\} < \in$

The MST based LINEX HELLINGER _IFCM algorithm iteratively optimizes J_{WLINEX} by continuous updating u^*_{ik} and v_{ij} until the difference in successive u^*_{ik} values is very small $\leq \in$, where $\in$ is a small positive value between 0 and 1.

4. Efficient LINEX HELLINGER Distance Based IFCM

4.1. Efficient LINEX HELLINGER_IFCM Algorithm

Stage 1: Set the cluster centroids $\{v_j\}_{j=1}^{c}$ by using LINEX measure MST initialization method.

Stage 2: Compute the membership function using (13)

Stage 3: Update the cluster centroids using (14)

Stage 4: Go to stage (3-5), repeat until convergence.

Stage 7: Image segmentation after defuzzification and then a region labeling procedure is proposed.

Stage 8: The termination criterion is as follows

$|J_m - J_{m-1}| < \in$, where m is the iteration count,

$\in$ is a small number that can be set by the user.

The proposed efficient LINEX HELLINGER distance based MST obtained cluster centers; the LINEX HELLINGER distance based IFCM algorithm continues iteratively updates, membership and centroids with these values. When this improved, Efficient LINEX HELLINGER distance based IFCM algorithm has converged, another defuzzification process takes place in order to convert the fuzzy partition matrix to a crisp partition matrix that is segmented.

4.2. Validation Function Based on Feature Structures

Two representative functions for the fuzzy partition namely; Partition coefficient V_{pc} and Validation function V_p are used to evaluate the validity of clustering [15, 16].

$$V_{pc} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{c} {u^*_{ij}}^2 \tag{17}$$

The proposed efficient LINEX HELLINGER distance based MST obtained cluster centers; the LINEX HELLINGER distance based IFCM algorithm continues iteratively updates, membership and centroids with these values. When this improved, Efficient MHIFCM algorithm has converged, another defuzzification process takes place in order to convert the fuzzy partition matrix to a crisp partition matrix that is segmented.

$$V_p = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n} u^m_{ij} L_{WLINEX}(x_j, v_i) + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*}}{N \times \min\left\{ \left\| v_i - v_j \right\|^2 \right\}} \tag{18}$$

5. Results and Discussion

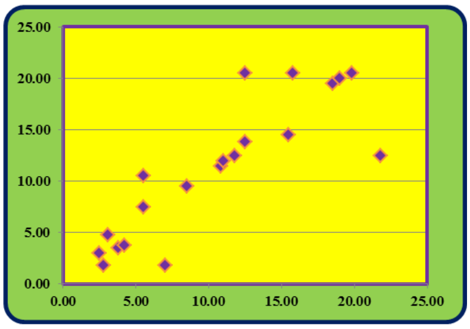

Figure 1. *Scatter diagram for random dataset.*

Table 1. *Random Data.*

Data			**Intensity**	**Data**			**Intensity**
S. No	X	Y	I (v)	S. No	X	Y	I (v)
1	2.50	3.00	0.75	11	15.80	20.50	0.25
2	3.80	3.50	0.50	12	11.80	12.50	0.35
3	7.00	1.80	0.15	13	15.50	14.50	0.80
4	3.10	4.80	0.18	14	5.50	10.50	0.70
5	5.50	7.50	0.45	15	18.50	19.50	0.40

Data			Intensity	Data			Intensity
6	8.50	9.50	0.75	16	12.50	13.80	0.25
7	10.80	11.50	0.60	17	21.80	12.50	0.95
8	4.20	3.80	0.25	18	19.80	20.50	0.25
9	2.80	1.80	0.45	19	19.00	20.00	0.60
10	12.50	20.50	0.65	20	11.00	12.00	0.30

Table 2. Dissimilarity matrix.

S. No	Co-ordinate		intensity											
	x	y	I (v)	vertex	1	2	3	4	5	6	7	8	9	10
1	2.50	3.00	0.75	1	0.0000	0.0408	0.2892	0.0942	0.3091	0.6035	0.8693	0.0867	0.0449	1.5020
2	3.80	3.50	0.50	2		0.0000	0.1761	0.0490	0.1841	0.4338	0.6711	0.0140	0.0780	1.3012
3	7.00	1.80	0.15	3			0.0000	0.2876	0.3871	0.5657	0.7566	0.1521	0.2048	1.4524
4	3.10	4.80	0.18	4				0.0000	0.1455	0.4181	0.6427	0.0337	0.1618	1.2158
5	5.50	7.50	0.45	5					0.0000	0.1029	0.2566	0.1525	0.4303	0.7407
6	8.50	9.50	0.75	6						0.0000	0.0539	0.3941	0.7374	0.4416
7	10.80	11.50	0.60	7							0.0000	0.6146	1.0036	0.2621
8	4.20	3.80	0.25	8								0.0000	0.1144	1.2360
9	2.80	1.80	0.45	9									0.0000	1.6847
10	12.50	20.50	0.65	10										0.0000

Table 3. LINEX measure based minimal spanning tree edges.

S. No	Edges	LINEX HELLINGER distance	S. No	Edges	LINEX HELLINGER distance
1	(1, 2)	0.0408	11	(20, 12)	0.0051
2	(2, 8)	0.0140	12	(12, 16)	0.0120
3	(8, 4)	0.0337	13	(16, 13)	0.0714
4	(8,)	0.1144	14	(13, 15)	0.1204
5	(8, 3)	0.1521	15	(15, 19)	0.0064
6	(8, 5)	0.1525	16	(15, 18)	0.0124
7	(5, 14)	0.0619	17	(15, 11)	0.0313
8	(14, 6)	0.0767	18	(11, 10)	0.0635
9	(6, 7)	0.0539	19	(11, 17)	0.3183
10	(7, 20)	0.0136			

***Table 4**. Degree of minimal spanning tree edges.*

S. No	Degree of MST	S. No	Degree of MST
1	1	11	3
2	2	12	2
3	1	13	2
4	1	14	2
5	2	15	4
6	2	16	2
7	2	17	1
8	5	18	1
9	1	19	1
10	1	20	2

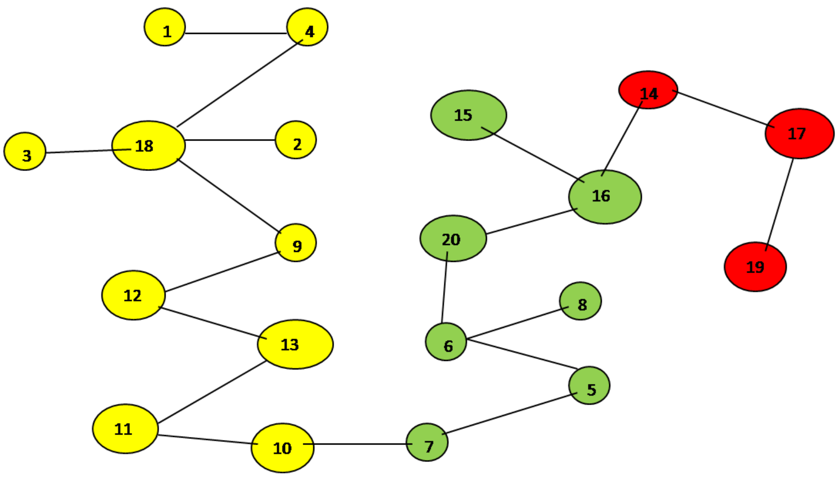

***Figure 2**. LINEX HELLINGER distance based Minimal spanning tree connected through points.*

***Table 5**. Edge density minimal spanning tree edges.*

S. No	Edges	Edge density	S. No	Edges	Edge density
1	(1, 2)	0.0577	11	(20, 12)	0.0102
2	(2, 8)	0.0443	12	(12, 16)	0.0240
3	(8, 4)	0.0754	13	(16, 13)	0.1428
4	(8, 9)	0.2558	14	(13, 15)	0.3405
5	(8, 3)	0.3401	15	(15, 19)	0.0128
6	(8, 5)	0.4822	16	(15, 18)	0.0248
7	(5, 14)	0.1238	17	(15, 11)	0.1084
8	(14, 6)	0.1534	18	(11, 10)	0.1100
9	(6, 7)	0.1078	19	(11, 17)	0.5513
10	(7, 20)	0.0272			

Our weighted LINEX Hellinger distance based minimal spanning tree algorithm constructs MST from the dissimilarity matrix is shown figure 2. Figure 2 shows a typical example of weighted LINEX Hellinger measure MST1 constructed from point set (from Dissimilarity matrix), in which inconsistent edges are removed to create subtree (clusters/regions). our algorithm finds the center of each clusters, which will be useful in many applications. Generally in most of the clustering algorithm data points can be represented as dissimilarity matrix representation. It contains the distance values between the data points represented as lower or upper triangular matrix. First to identify the longest edge or $w > \overline{w} + \left(\frac{\max(EDW) - \overline{w}}{2}\right)$ in the MST to generate subtree (clusters). Table 4, the longest edge weight 0.5513 connecting the data points 11 and 17 is find to be inconsistent one. By removing the inconsistent edge from the weighted LINEX HELLINGER_MST, data points partitioned into two subtrees or clusters T_1 and T_2 namely. $T_1 = \{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,18,19,20\}$ and $T_2 = \{17\}$. Now minimal spanning tree produced outlier is {17}. Next to find minimum and maximum standard deviation of T_i then Cluster Separation value less than 0.5625 is valid. Secondly to remove next longest edge, weighted LINEX HELLINGER distance based MST partitioned into Three subtrees or clusters T_1, T_2 and T_3 namely $T_1 = \{1,2,3,4,8,9\}$, $T_2 = \{5,6,7,12,13,14,16,20\}$ and $T_3 = \{10,11,15,18,19\}$. Next to find minimum and maximum edge and to calculate standard deviation of T_i then Cluster separation value less than 0.5625 is valid. Here standard deviation of edge density $T_1 = 0.1350$, Standard deviation of edge density $T_2 = 0.0620$ and Standard deviation of edge density $T_3 = 0.0520$ then $CS = \frac{0.0520}{0.1350} = 0.3890 < 0.5740$ then we stop the removing inconsistent edges.. Finally minimal spanning tree produced three cluster center and 1 outliers. Then the center of the cluster and its convergence of standard FCM and LINEX HELLINGER distance based IFCM are determined under successive interations of experiments using data points. With the new efficient objective function based LINEX Hellinger distance induced weighted measure the termination value is achieved, with very less iteration and with much better performance in getting membership (Table 6) than standard FCM. Table 7 gives the number of iteration to achieve the results of cluster on the data points by standard FCM and weighted LINEX_IFCM. It is clear from the final cluster, membership (Table 6), scatter diagram (Figure 1), minimal spanning tree (Figure 2) that our proposed LINEX Hellinger _IFCM induced weighted degree of minimal spanning tree is much faster than the standard FCM and the method is converged fast to terminate condition with less run time. To test the effectiveness of weighted LINEX_IFCM, the edge density minimal spanning tree based IFCM is used as center. This is done to find out the fuzzy membership and appropriate number of clusters. Thus, we have concluded the final optimal clusters formed as 3 (Figure 3). This algorithm has also reduced the number of iterations. Best result is achieved by this measure fuzzy partition coefficient V_{pc} maximum and validation function V_p minimum (Table 8). The weighted LINEX_IFCM clustering algorithm has the following membership value intimacy (Table 6).

Table 6*. Final membership of three clusters of LINEX HELLINGER distance based Intuitionistic FCM method and object allocation.*

Co-ordinate (x, y)			**intensity**				
S. No	**x**	**y**	**I (v)**	**Mem-1**	**Mem-2**	**Mem-3**	**appropriate cluster**
1	2.50	3.00	0.75	1	0.9213	0.0523	0.0265
2	3.80	3.50	0.50	2	0.9895	0.0072	0.0032
3	7.00	1.80	0.15	1	0.7820	0.1444	0.0736
4	3.10	4.80	0.18	1	0.8808	0.0830	0.0362
5	5.50	7.50	0.45	2	0.4630	0.4292	0.1077
6	8.50	9.50	0.75	2	0.0794	0.8548	0.0658
7	10.80	11.50	0.60	2	0.0046	0.9858	0.0096
8	4.20	3.80	0.25	5	0.9710	0.0202	0.0087
9	2.80	1.80	0.45	1	0.9149	0.0551	0.0299

Co-ordinate (x, y)			intensity				appropriate cluster
S. No	x	y	I (v)	Mem-1	Mem-2	Mem-3	
10	12.50	20.50	0.65	1	0.0559	0.2632	0.6809
11	15.80	20.50	0.25	3	0.0199	0.0788	0.9013
12	11.80	12.50	0.35	2	0.0269	0.8865	0.0866
13	15.50	14.50	0.80	2	0.0480	0.3319	0.6201
14	5.50	10.50	0.70	2	0.2367	0.6315	0.1318
15	18.50	19.50	0.40	4	0.0026	0.0093	0.9882
16	12.50	13.80	0.25	2	0.0490	0.7068	0.2442
17	21.80	12.50	0.95	1	0.0852	0.2753	0.6395
18	19.80	20.50	0.25	1	0.0161	0.0513	0.9326
19	19.00	20.00	0.60	1	0.0050	0.0171	0.9779
20	11.00	12.00	0.30	2	0.0171	0.9420	0.0410

Table 7. *Comparison of iteration count.*

	No. of iterations	No. of clusters
FCM	11	3
KFCM	8	3
Edge density MST initialization method based LINEX Hellingerdistance based Intuitionistic FCM	3	3

Table 8. *Cluster validity function.*

	V_{pc}	V_p
FCM	0.8325	0.1825
KFCM	0.8238	0.1720
MST initialization method based LINEX Hellinger Intuitionistic FCM	0.8889	0.0661

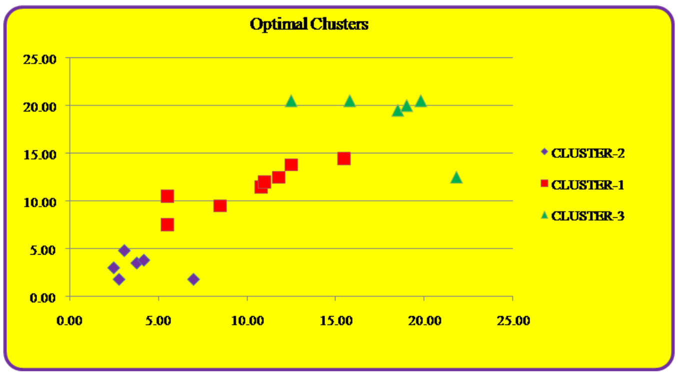

Figure 3. *MST initialization method based LINEX Hellinger distance using Intuitionistic FCM, final cluster three.*

6. Conclusion

In this paper, we have proposed a modified intuitionistic fuzzy c-means algorithm (IFCM) induced weighted LINEX Hellinger distance measure using degree of MST and solved analytically the objective function of the weighted LINEX Hellinger_IFCM method using Lagrange method of undetermined multiplier. Besides, the intuitionistic fuzzy set and non-membership degree are generated by an improved method, which highlight the role of uncertainty effectively. The algorithm overcomes problems involved with membership values of objects to each cluster by generalizing degrees of membership of objects to each cluster. For the initial partition matrix, if the membership degrees and non-membership degrees of the classified objects to the categories are obviously different, the iteration times will be reduced accordingly. Compared with FCM clustering algorithm, Experimental results demonstrate that the performance of random image data points, the pixel clustering effect and the robustness to noise of the proposed algorithm are all significantly better than other algorithms method.

Acknowledgments

We would like to thank the reviewers for their constructive comments. We thank to Dr. R. David Chandrakumar, Professor, Department of Mathematics, Vickram College of Engineering for his encouragement and support given. I also thank Dr. A. Sathya, Assistant Professor, Department of Mathematics, Government College of Daman, Daman & Diu for her timely help throughout this work.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Atanassov KT, “Intuitionistic fuzzy sets. Fuzzy Sets Systems”, vol. 20, pp. 87-96 (1986).

[2] Arora J and Tushir M,“Robust spatial intuitionistic fuzzy C-means with city-block distance clustering for image segmentation” Journal of Intelligent & Fuzzy Systems, vol. 35, pp. 5255-5264 (2018).

[3] Chaira T, “A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images” Applied Soft computing, vol. 11(2), pp. 1711-1717 (2011).

[4] Cheng D, Zhu Q, Huang J, Wu Q, and Yang L, ‘Clustering with local density peaks-based minimum spanning tree’, IEEE Transactions on Knowledge and Data Engineering, vol. 33(2), pp. 374–387 (2021).

[5] Jun Kong, Jian Hou, Min Jiang & Jinhua Sun, ‘‘A Novel Image Segmentation Method Based on Improved Intuitionistic Fuzzy C-Means Clustering Algorithm’, KSII transactions on internet and information systems, vol. 13(6), pp. 3121-3143 (2019).

[6] Li G, Long C, Chen CLP and Zhou J, “Integrating guided filter into fuzzy clustering for noisy image segmentation” Digit. Signal Process, vol. 83, pp. 235–248 (2018).

[7] Lohani QMD, Solanki R and Muhuri PK, ‘Novel adaptive clustering algorithms based on a probabilistic similarity measure over Atanassov intuitionistic fuzzy set’, IEEE Trans. Fuzzy Syst., vol. 26(6), pp. 3715–3729 (2018).

[8] Nithya M et al., “A Fiedler’s approach to LINEX Intuitionistic fuzzy c means clustering induced spectral initialization for data analysis”, Applied and Computational Mathematics, vol. 12 (4), pp. 82-91(2023).

[9] Pop PC, ‘The generalized minimum spanning tree problem: an overview of formulations, solution procedures and latest advances’, European Journal of Operational Research, vol. 283(1), pp. 1–15 (2020).

[10] Szmidt E and Kacprzyk J, “Distances between intuitionistic fuzzy sets” Fuzzy sets Systems, vol. 114, pp. 505-518(2000).

[11] Senthil S et al., “Efficient kernel induced fuzzy c-means based on Gaussian function for image data analyzing”, Journal of Intelligent & Fuzzy Systems, https://doi.org/10.3233, IFS-151820 IOS Press, 13 October 13. (2015).

[12] Xu Z and Wu J, “Intuitionistic fuzzy C-means clustering algorithms” J. Syst. Eng. Electron, vol. *21*, pp. 580–590 (2010).

[13] Zhao F, Chen Y, Liu H and Fan J, ‘Alternate PSO-based adaptive interval type-2 intuitionistic fuzzy C-means clustering algorithm for color image segmentation’, IEEE Access, vol. 7, pp. 64028–64039 (2019).

[14] Zhao F, Liu HQ, Fan JL, Chen CW, Lan R & N. Li N, ‘Intuitionistic fuzzy set approach to multi-objective evolutionary clustering with multiple spatial information for image segmentation’, Neurocomputing, vol. 312, pp. 296-309 (2018).

[15] Xie X. L and Beni G, A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 3(8), pp-841-847(1991).

[16] Bezdek J. C., Hall L. O., Clarke L. P, Review of MR image segmentation techniques using pattern recognition, medical physics 20(4), pp. 1033-1048 (1993).

[17] Rudolf Beran, Minimum Hellinger distance estimates for parametric models, University of California, Berkeley, The Annals of Statistics, Vol. 5(3), pp. 445-463 (1977).

Biography

Nithya M. was born in 1992. She is an Research Scholar of Mathematics, Mother Teresa Womens University, Kodaikanal, Tamilnadu, India. His current research interests include digital topology, data mining and image processing.

K. Bhuvaneswari Baasubramaium was born in 1968. He received the Ph. D degree from Bharathiar University, Coimbatore in 2005. She is working as a Professor in Department of Mathematics, Mother Teresa Womens University, Kodaikanal, Tamilnadu, India. His current research interests include digital topology, data mining and image processing.

Senthil S. was born in 1983. He received the Ph. D degree from Anna University, Chennai in 2016. He is working as a Statistical Inspector in Department of Economics and Statistics, Integrated Child Development Services (ICDS), Collectorate, Dindigul, Tamilnadu, India. His current research interests include data mining and image processing.