A Multiple Kernels Interval Type-2 Possibilistic C-Means

Minh Ngoc Vu and Long Thanh Ngo

Abstract In this paper, we propose multiple kernels-based interval type-2 possibilistic c-Means (MKIT2PCM) by using the kernel approach to possibilistic clustering. Kernel-based fuzzy clustering has exhibited quality of clustering results in comparison with "routine" fuzzy clustering algorithms like fuzzy c-Means (FCM) or possibilistic c-Means (PCM) not only noisy data sets but also overlapping between prototypes. Gaussian kernels are suitable for these cases. Interval type-2 fuzzy sets have shown the advantages in handling uncertainty. In this study, multiple kernel method are combined into interval type-2 possibilistic c-Means (IT2PCM) to produce a variant of IT2PCM, called multiple kernels interval type-2 possibilistic c-Means (MKIT2PCM). Experiments on various data-sets with validity indexes show the performance of the proposed algorithms.

Keywords Type-2 fuzzy sets · Fuzzy clustering · Multiple kernels · PCM

1 Introduction

Clustering is one of the most important technique in machine intelligence, approach to unsupervised learning for pattern recognition. The fuzzy c-Means (FCM) algorithm [1] has been applied successfully to solve a large number of problems such as patterns recognition (fingerprint, photo) [2], image processing (color separated clustering, image segmentation) [3]. However, the drawbacks of FCM algorithm are sensitive to noise, loss of information in process and overlap clusters of different volume. To handle this uncertainty, Krishnapuram et al. used the compulsion of memberships in FCM, called Possibilistic c-Means (PCM) algorithm [4]. The PCM algorithm solves

M.N. Vu (✉) · L.T. Ngo
Department of Information Systems, Le Quy Don Technical University,
236 Hoang Quoc Viet, Hanoi, Vietnam
e-mail: ngoc114@gmail.com

L.T. Ngo
e-mail: ngotlong@mta.edu.vn

D. Król et al. (eds.), *Recent Developments in Intelligent Information and Database Systems*, Studies in Computational Intelligence 642,
DOI 10.1007/978-3-319-31277-4_6

so well some problems of actual data input which may be uncertainty, contradictory and vaguenesses that FCM algorithm cannot handle correctly. Though, PCM still exists some drawbacks such as: depending deeply on initializing centroid or sensitivity to noise data [5, 6]. To settle these weakness associated with PCM and improve accuracy in data processing, multiple kernel method based clustering is used. The main idea of kernel method by mapping the original data inputs into a higher dimensional space by transform function. However, different input features may impact on the results, the multiple kernel method can be combined from various kernels.

Type-2 fuzzy sets, which are the extension of fuzzy sets of type-1 [7], have shown advantages when solving uncertainty, have been researched and exploited to many different applications [8, 9], including fuzzy clustering. The interval type-2 Fuzzy Sets (IT2FS) and the kernel method [10, 11] considered separately have shown potential for the clustering problems. This combination of IT2FS and multiple kernels gives the extended interval type-2 possibilistic c-Means [12] in the kernel space. In this paper, we propose a multiple kernels interval type-2 possibilistic c-Means (MKIT2PCM) which is the kernel method based on the possibilistic approach to clustering. Because data-sets could come from many sources with different characteristics, the multiple kernels approach to IT2PCM clustering could be to handle the ambiguousness of data. In which, the final kernel function is combined from component kernel functions by sum of weight average. The algorithm is experimented with various data to measure the validity indexes.

The paper is organized as follows: Sect. 2 brings a brief background of type-2 fuzzy sets and kernel method. Section 3 describles the MKIT2PCM algorithm. Section 4 shows some experimental results and Sect. 5 concludes the paper.

2 Preliminaries

2.1 Interval Type-2 Fuzzy Sets

A type-2 fuzzy set [9] defined in X denoted $\tilde{A}$, comes with a membership function of the form $\mu_{\tilde{A}}(x,u), u \in J_x \subseteq [0,1]$, which is a type-1 fuzzy set in [0, 1]. The elements of field of $\mu_{\tilde{A}}(x,u)$ are called primary membership grades of x in $\tilde{A}$ and membership grades of primary membership grades in $\mu_{\tilde{A}}(x,u)$ are called secondary ones.

Definition 1 A type-2 fuzzy set, denoted $\tilde{A}$, is described by a type-2 membership function $\mu_{\tilde{A}}(x,u)$ where $x \in X$ and $u \in J_x \subseteq [0,1]$, i.e.

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

or

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x,u))/(x,u), J_x \subseteq [0,1] \tag{2}$$

in which $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$

At each value of x, say $x = x'$, the 2-D plane whose axes are u and $\mu_{\tilde{A}}(x',u)$ is called a vertical slice of $\mu_{\tilde{A}}(x,u)$. A secondary membership function is a vertical slice of $\mu_{\tilde{A}}(x,u)$. It is $\mu_{\tilde{A}}(x = x',u)$ for $x \in X$ and $\forall u \in J_{x'} \subseteq [0,1]$, i.e.

$$\mu_{\tilde{A}}(x = x',u) \equiv \mu_{\tilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u)/u, J_{x'} \subseteq [0,1] \tag{3}$$

in which $0 \leq f_{x'}(u) \leq 1$ (Fig. 1).

A type-2 fuzzy set is called an interval type-2 fuzzy set if its secondary membership function $f_{x'}(u) = 1 \ \forall u \in J_x$ i.e. such constructs are defined as follows:

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

$$\tilde{A} = \{((x,u),1) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \tag{4}$$

Footprint of uncertainty (FOU) of a type-2 fuzzy set is union of primary functions i.e. $FOU(\tilde{A}) = \bigcup_{x \in X} J_x$. Upper/lower bounds of membership function (UMF/LMF) are denoted by $\bar{\mu}_{\tilde{A}}(x)$ and $\underline{u}_{\tilde{A}}(x)$.

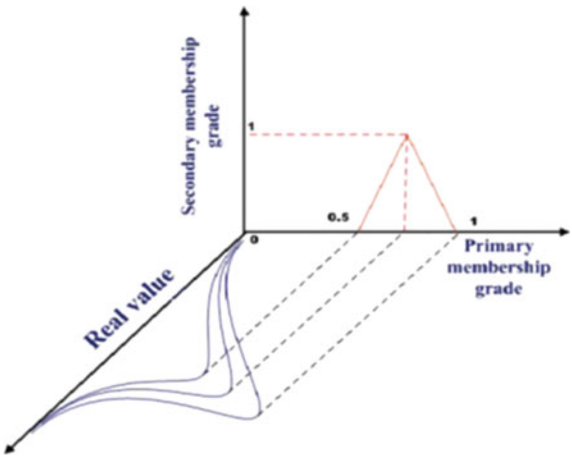

Fig. 1 Visualization of a type-2 fuzzy set

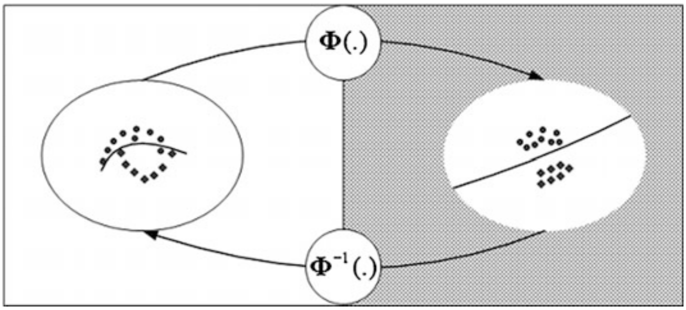

Fig. 2 Expression of the mapping from feature space to kernel space

2.2 Kernel Method

Kernel method gives an exclusive approach which maps all data ϕ from the basic dimensional feature space R^p to a different higher possibly dimensionality vector space (kernel space). The kernel space could be of infinite dimensionality.

Definition 3 K is a kernel function on X space if

$$K(x, y) = \phi(x).\phi(y) \tag{5}$$

where ϕ is a nonlinear map from X to a scalar space F: $\phi : R^p \to F$; $x, y \in X$. Figure 2 [13] represents the mapping from the original space to kernel space

The function $k : X \times X \to F$ is called Mercer kernel if $t \in I$ and $x = (x_1, x_2..., x_t) \in X^t$ the $t \times t$ matrix K is positive semi definite.

3 Multiple Kernels Interval Type-2 Possibilistic C-Means Clustering

3.1 Interval Type-2 Possibilistic C-Means Clustering

Interval type-2 possibilistic c-Means clustering (IT2PCM) is an completion of the PCM clustering in which we use two fuzzification coefficients m_1, m_2 corresponding to the upper and lower values of membership. Objective functions used for IT2PCM is shown in:

$$J_{m_1}(U, v) = \min\left[\sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij})^{m_1} d_{ij}^2 + \sum_{i=1}^{c}\eta_i \sum_{j=1}^{n}(1 - u_{ij})^{m_1}\right] \tag{6}$$

$$J_{m_2}(U,v) = \min\left[\sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij})^{m_2}d_{ij}^2 + \sum_{i=1}^{c}\eta_i\sum_{j=1}^{n}(1-u_{ij})^{m_2}\right] \quad (7)$$

where $d_{ij} = \| x_j - v_i \|$ is the Euclidean distance between the jth data and the ith cluster centroid, c is the number of clusters and n is number of patterns. m_1, m_2 is the degree of fuzziness and η_i is a suitable positive number. Upper/lower degrees of membership $\bar{u}_{ij}$ and $\underline{u}_{ij}$ are formed by involving two fuzzification coefficients m_1, m_2 $(m_1 < m_2)$ as follows:

$$\bar{u}_{ij} = \begin{cases} \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_1-1)}} & \text{if } \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_1-1)}} \ge \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_2-1)}} \\ \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_2-1)}} & \text{otherwise} \end{cases} \quad (8)$$

$$\underline{u}_{ij} = \begin{cases} \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_1-1)}} & \text{if } \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_1-1)}} \le \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_2-1)}} \\ \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m_2-1)}} & \text{otherwise} \end{cases} \quad (9)$$

$$i = 1,2,\ldots,c; j = 1,2,\ldots,n$$

where $\overline{\eta_{ij}}$ and $\underline{\eta_{ij}}$ are upper/lower degrees of typicality function calculated by

$$\overline{\eta_{ij}} = \frac{\sum_{j=1}^{n}\bar{u}_{ij}d_{ij}^2}{\sum_{j=1}^{n}\bar{u}_{ij}} \quad (10)$$

and

$$\underline{\eta_{ij}} = \frac{\sum_{j=1}^{n}\underline{u}_{ij}d_{ij}^2}{\sum_{j=1}^{n}\underline{u}_{ij}} \quad (11)$$

Because each pattern comes with the membership interval as the bounds $\bar{u}$ and $\underline{u}$. Cluster prototypes are computed as

$$v_i = \frac{\sum_{j=1}^{n}(u_{ij})^m x_j}{\sum_{j=1}^{n}(u_{ij})^m} \quad (12)$$

In which $i = 1,2,\ldots,c$

Algorithm 1: **Determining Centroids**.

Step 1: Find $\bar{u}_{ij}, \underline{u}_{ij}$ using (8), (9)
Step 2: Set $m \ge 1$

Compute $v'_j = (v'_{j1}, \ldots, v'_{jP})$ using (12) with $u_{ij} = \frac{(\bar{u}_{ij} + \underline{u}_{ij})}{2}$.
Sort n patterns on each of P features in ascending order.
Step 3: Find index k where: $x_{kl} \le v'_{jl} \le x_{(k+1)l}$ with $k = 1, .., N$ and $l = 1, .., P$
Update u_{ij}: if $i \le k$ then do $u_{ij} = \underline{u}_{ij}$ else $u_{ij} = \bar{u}_{ij}$
Step 4: Compute v''_j by (12) and compare v'_{jl} with v''_{jl}
If $v'_{jl} = v''_{jl}$ then do
$v_R = v'_j$.
Else
Set $v'_{jl} = v''_{jl}$.
Back to Step 3.

For example we compute v^L

In step 3
Update u_{ij}:
If $i \le k$ then do
$u_{ij} = \bar{u}_{ij}$.
Else
$u_{ij} = \underline{u}_{ij}$.
In Step 4 change V_R with v_L.
After having v_i^R, v_i^L, the prototypes is computed as follows

$$v_i = \left(\frac{v_i^R + v_i^L}{2} \right) \tag{13}$$

For the membership grades we obtain

$$u_i(x_k) = (u_i^R(x_k) + u_i^L(x_k))/2, j = 1, \ldots, c \tag{14}$$

where

$$u_i^L = \sum_{l=1}^{M} u_{il}/P, u_{il} = \begin{cases} \bar{u}_i(x_k) & \text{if } x_{il} \text{ uses } \bar{u}_i(x_k) \text{ for } v_i^L \\ \underline{u}_i(x_k) & \textit{otherwise} \end{cases} \tag{15}$$

$$u_i^R = \sum_{l=1}^{M} u_{il}/P, u_{il} = \begin{cases} \bar{u}_i(x_k) & \text{if } x_{il} \text{ uses } \bar{u}_i(x_k) \text{ for } v_i^R \\ \underline{u}_i(x_k) & \textit{otherwise} \end{cases} \tag{16}$$

Next, defuzzification follows the ruleif $u_i(x_k) > u_j(x_k)$ for $j = 1, \ldots, c$ and $i \neq j$ then x_k is assigned to cluster i

3.2 Multiple Kernel Interval Type-2 Possibilistic C-Means Algorithm

Here, we introduce MKIT2PCM algorithm which can use a combination of different kernels. The pattern $x_j, j = 1, \ldots n$ and prototype v_i, $i = 1, \ldots c$ are mapped to the kernel space through a kernel function ϕ. The objective functions used for MKIT2PCM is shown in

$$Q_{m_1}(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m_1} \parallel \phi_{com}(x_j) - \phi_{com}(v_i) \parallel^2 + \sum_{i=1}^{c}\eta_i\sum_{j=1}^{n}(1-u_{ij})^{m_1} \quad (17)$$

$$Q_{m_2}(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m_2} \parallel \phi_{com}(x_j) - \phi_{com}(v_i) \parallel^2 + \sum_{i=1}^{c}\eta_i\sum_{j=1}^{n}(1-u_{ij})^{m_2} \quad (18)$$

where $\parallel \phi_{com}(x_j) - \phi_{com}(v_i) \parallel = d_{\phi ij}$ is the Euclidean distance from the prototype v_i to the pattern x_j which is computed in the kernel space.

Where ϕ_{com} is the transformation defined by the combined kernels

$$k_{com}(x,y) = <\phi_{com}(x), \phi_{com}(y)> \quad (19)$$

The kernel k_{com} is the function that is combined from multiple kernels. A commonly used the combination is a linearly summary that reach as follows:

$$k_{com} = w_1^b k_1 + w_2^b k_2 + \cdots + w_l^b k_l \quad (20)$$

where $b > 1$ is a certain coefficient and satisfying the constraint $\sum_{i=1}^{l} w_i = 1$.

The typical kernels defined on $\mathbb{R}^p \times \mathbb{R}^p$ are used: Gaussian kernel $k(x,y) = \exp\left(\frac{-x-y^2}{2\sigma^2}\right)$ and polynomial kernel $k(x_i,x_j) = (x_i * x_j + d)^2$

The next is how to compute $\bar{u}_{ij}$ and $\underline{u}_{ij}$ which can be reached by:

$$\bar{u}_{ij} = \begin{cases} \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_1-1)}} & if\ \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2)}{\eta_i}\right)^{1/(m_1-1)}} \geq \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_2-1)}} \\ \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_2-1)}} & otherwise \end{cases} \quad (21)$$

$$\underline{u}_{ij} = \begin{cases} \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_1-1)}} & if\ \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_1-1)}} \leq \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_2-1)}} \\ \frac{1}{1+\sum_{j=1}^{c}\left(\frac{d_{\phi com}(x_j,v_i)^2}{\eta_i}\right)^{1/(m_2-1)}} & otherwise \end{cases} \quad (22)$$

Respectively, where $d_{\phi_{com}}(x_j, v_i)^2 = \parallel \phi_{com}(x_j) - v_i \parallel^2$

$$= k_{com}\left(x_j, x_j\right) - \frac{2\sum_{h=1}^{n}(u_{ih})^m k_{com}\left(x_j, x_h\right)}{\sum_{h=1}^{n}(u_{ih})^m} + \frac{\sum_{h=1}^{n}\sum_{l=1}^{n}(u_{ih})^m (u_{il})^m k_{com}(x_h, x_l)^2}{\left(\sum_{h=1}^{n}(u_{ih})^m\right)} \tag{23}$$

where $\overline{\eta_{ij}}$ and $\underline{\eta_{ij}}$ are calculated by

$$\overline{\eta_{ij}} = \frac{\sum_{j=1}^{n}\bar{u}_{ij}\phi_{com}(x_j) - v_i^2}{\sum_{j=1}^{n}\bar{u}_{ij}} \tag{24}$$

$$\underline{\eta_{ij}} = \frac{\sum_{j=1}^{n}\underline{u}_{ij}\phi_{com}(x_j) - v_i^2}{\sum_{j=1}^{n}\underline{u}_{ij}} \tag{25}$$

Algorithm 2. **Multiple kernel interval type 2 possibilistic c-Means clustering**

Input: n patterns $X = \{x_i\}_{i=1}^{n}$, kernel functions $\{k_i\}_{i=1}^{l}$, fuzzifiers m_1, m_2 and c clusters.

Output: a membership matrix $U = \{u_{ij}\}_{i=1,..,n}^{j=1,..,c}$ and weights $\{w_i\}_{i=1}^{l}$ for the kernels.

Step 1: Initialization of membership grades

1. Initialize centroid matrix $V^0 = \{v_i\}_{i=1}^{C}$ is randomly chosen from input data set
2. The membership matrix U^0 can be computed by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{lj}}\right)^{\frac{2}{m-1}}} \tag{26}$$

where $m > 1$ and $d_{ij} = d\left(x_j - v_i\right) = \left\|x_j - v_i\right\|$.

Step 2: Repeat

1. Compute distance $d(x_j, v_i)^2$ using formula (23)
2. Compute $\overline{\eta_{ij}}$ and $\underline{\eta_{ij}}$ using formulas (24), (25)
3. Calculate the upper and lower memberships $\bar{u}_{ij}$ and $\underline{u}_{ij}$ using formulas (21), (22)
4. Use **Algorithm 1** for finding centroids and using formula (12) to update the centroids.
5. Update the membership matrix using formula (14)

Step 3: Conclusion criteria satisfied or maximum iterations reached:

1. Return U and V
2. Assign data x_j to cluster c_i if data $(u_j(x_i) > u_k(x_i))$, $k = 1, \ldots, c$ and $j \neq k$.

Else

Back to step 2.

4 Experimental Results

In this paper, the experiment is implemented on remote sensing images with validity index comparisons. The vector of pixels is acquired from various sensors with different spectral. So that, we can identify different kernels for different spectrals and use the combined kernel in a multiple-kernel method.

The data inputs involve Gaussian kernel k_1 of pixel intensities and Gaussian kernel k_2 of spatial information. The spatial function is defined as follows [14]

$$p_{ij} = \sum_{k \in NB(x_j)} u_{ik} \tag{27}$$

where $NB(x_j)$ represents a squared neighboured 5×5 window with the center on the pixel x_j. The spatial function p_{ij} express membership grade that pixel x_j belongs to the cluster i.

The kernel algorithm produces Gaussian kernel k of pixel intensities as data inputs. We use Lansat-7 images data in this experiment with study data of Cloud Landsat TM at IIS, U-Tokyo, (SceneCenterLatitude: 11.563043 S SceneCenterLongitude: 15.782159 W).

The results are shown in Fig. 3 where (a) is for VR channel, (b) is for NIR channel, (c) is for SWIR channel and (d), (e) and (f) are image results of the

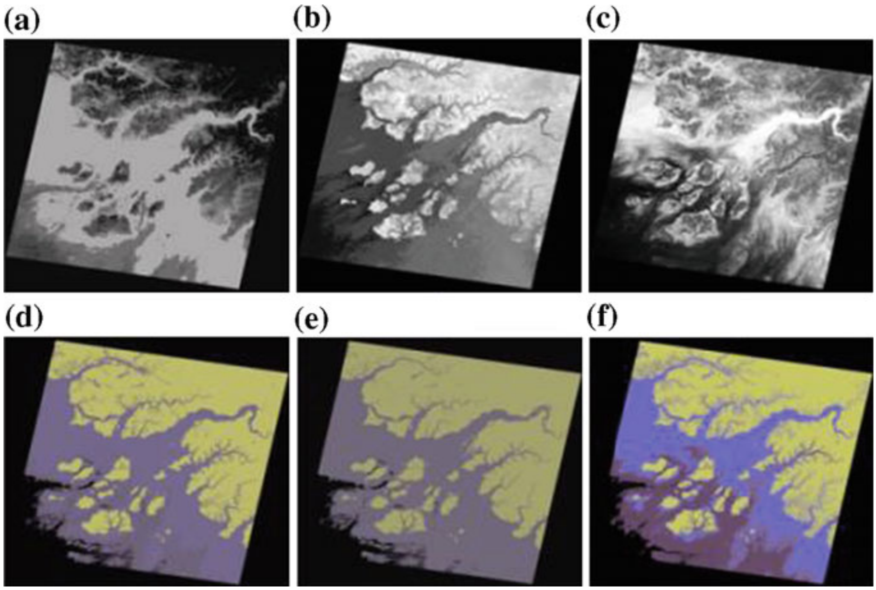

Fig. 3 Data for test: land cover classification. **a** VR channel image; **b** NIR channel image; **c** SWIR channel image; **d** MKPCM classification; **e** IT2PCM classification; **f** MKIT2PCM classification

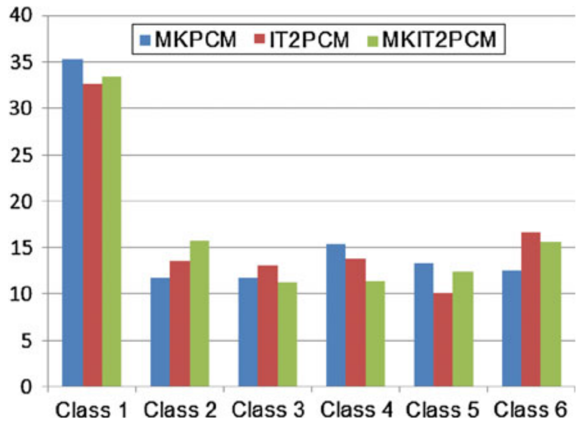

Fig. 4 Comparisons between the result of MKPCM, IT2PCM, MKIT2PCM

Table 1 The various validity indexes

Validity Index	MKPCM	IT2PCM	MKIT2PCM
PC-I	0.334	0.225	**0.187**
CE-I	1.421	1.645	**1.736**
S-I	0.019	0.171	**0.0608**
DB-I	2.42	2.051	**1.711**

classification of MKPCM, IT2PCM and MKIT2PCM algorithms. Figure 4 compares results between MKPCM, IT2PCM-F, MKIT2PCM algorithms (in percentage %).

We consulted the various validity indexes -which maybe the Bezdek's partition coefficient (PC-I), the Classification Entropy index (CE-I), the Separation index (S-I) and the Davies-Bouldin's index (DB-I) [15]. We can see the results of these indexes in the Table 1

The value of validity indexes coming from MKIT2PCM algorithm has smaller values of PC-I, S-I, DB-I and larger value only of CE-I. So that, the results in Table 1 show that the MKIT2PCM have better quality clustering than KIT2PCM and MKPCM in comparison by validity indexes.

5 Conclusion

In this paper, we applied the multiple kernel method to interval type-2 PCM clustering algorithm which enhances the efficiency of the clustering results. The experiments completed for remote sensing images dataset that the proposed algorithm quite better results than others produced by the familiar clustering methods of PCM. We demonstrated that the proposed algorithm can determine proper clusters, and they can realize the advantages of the possibilistic approach. Some future studies may be focused on other possible extensions can be incorporated interval type 2 fuzzy sets to hybrid clustering approach to data classification use multiple kernel fuzzy possibilistic c-Means clustering (FPCM).

References

1. Bezdek, J.C., Ehrlich, R., Full, W.: The fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2–3), 191–203 (1984)
2. Das, S.: Pattern recognition using the fuzzy c-means technique. Int. J. Energy Inf. Commun **4** (1), February (2013)
3. Patil, A., Lalitha, Y.S.: Classification of crops using FCM segmentation and texture, color feature. IJARCCE **1**(6), (2012)
4. Krishnapuram, R., Keller, J.: A possibilistic approach to clustering. IEEE Trans. Fuzzy Syst. **1**, 98–110 (1993)
5. Kanzawa, Y.: Sequential cluster extraction using power-regularized possibilistic c-means. JACIII **19**(1),67–73 (2015)
6. Krishnapuram, R., Keller, J.M.: The possibilistic c-means: insights and recommendations. IEEE Trans. Fuzzy Syst. **4**, 385–393 (1996)
7. Sanchez, M.A., Castillo, O., Castro, J.R., Melin, P.: Fuzzy granular gravitational clustering algorithm for multivariate data. Inf. Sci. **279**, 498–511 (2014)
8. Rubio, E., Castillo, O.: Designing type-2 fuzzy systems using the interval type-2 fuzzy c-means algorithm. Recent Adv. Hybrid Approaches Des. Intell. Syst., 37–50 (2014)
9. Nguyen, D.D., Ngo, L.T., Pham, L.T., Pedrycz, W.: Towards hybrid clustering approach to data classification: multiple kernels based-interval type-2 fuzzy C-means algorithms. J. Fuzzy Sets Syst. **279**, 17–39 (2015)
10. Zhao, B., Kwok, J., Zhang, C.: Multiple kernel clustering. In: Proceedings of 9th SIAM International Conference Data Mining, pp. 638–649, 2009
11. Filippone, M., Masulli, F., Rovetta, S.: Applying the possibilistic C-means algorithm in Kernel-induced spaces. IEEE Trans. Fuzzy Syst. **18**(3), 572–584 (2010)
12. Rubio, E., Castillo, O.: Interval type-2 fuzzy clustering for membership function generation. HIMA, 13–18, (2013)
13. Raza, M.A., Rhee, F.C.H.: Interval type-2 approach to Kernel possibilistic C-means clustering. In: IEEE International Conference on Fuzzy Systems, 2012
14. Chuang, K.S., Tzeng, H.L., Chen, S., Wua, J., Chen, T.-J.: Fuzzy c-means clustering with spatial information for image segmentation. Comput. Med. Imaging Graph. **30**(1), 9–15 (2006)
15. Wang, W., Zhang, Y.: On fuzzy cluster validity indices. Fuzzy Sets Syst. **158**, 2095–2117 (2007)