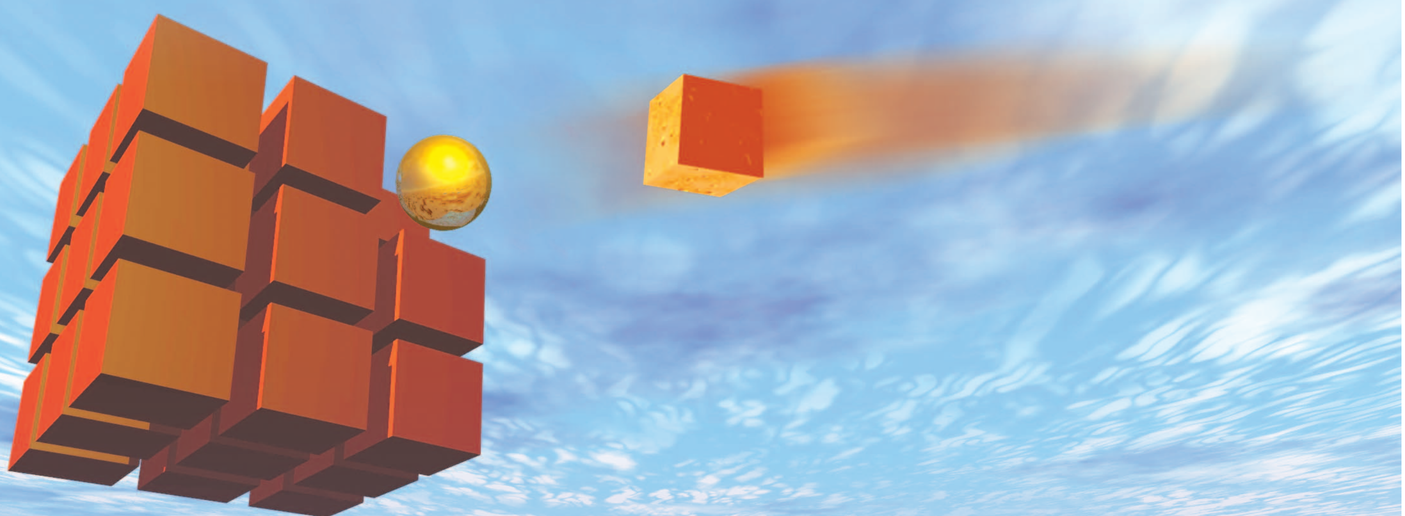
Frank Chung-Hoon Rhee
Hanyang University, KOREA

Uncertain Fuzzy Clustering:
Insights and Recommendations



I. Introduction

In many pattern recognition applications, the task of partitioning a pattern set can be considered to be the result of clustering procedures (algorithms) in which the cluster prototypes are estimated from the information of the pattern set. In many cases, it may be impossible to obtain exact knowledge or information from a given pattern set. Imprecise information can create imperfect representations of the pattern sets in various clustering algorithms. Therefore, various types of uncertainty may be taken into account when performing these algorithms. When one encounters fuzzy clustering, membership design includes various uncertainties such as ambiguous cluster membership assignment due to choice of distance measure, fuzzifier, prototype, and initialization of prototype parameters, to name a few.

For example, Figure 1 shows clustering results using the fuzzy C-means (FCM) clustering algorithm considering point prototypes for different fuzzifier parameter m that controls the amount of fuzziness of the final C-partition [1]. From the figure, results show that FCM can yield undesirable clustering results depending on the choice of fuzzifier m. From experience, the value for fuzzifier m is usually assigned to 2.0 and has been shown to be suitable for pattern sets that consist of clusters that are of similar volume and density. For pattern sets that contain

1556-603X/07/$25.00©2007IEEE

clusters of different volume or density, it may be difficult to determine an appropriate fuzzifier parameter value. Therefore, one entity of uncertainty present in FCM can be considered to be associated with the fuzzifier m.

As another example, the cluster membership that is assigned to a pattern can also be affected by the distance measure used between pattern and cluster prototype. Often, prototype-based fuzzy clustering, which assigns cluster memberships due to relative distance, can yield undesirable clustering results for pattern sets that include noise. To overcome this drawback, a possible approach to C-means (PCM) was proposed by Krishnapuram et al. [2]. It has been shown that the PCM approach, which assigns membership to pattern data according to cluster typicality, can yield robust clustering results for noisy pattern sets compared to FCM-based methods. However, PCM can be sensitive to the values of the initial cluster prototypes. Furthermore, for pattern sets that contain clusters that are relatively close to each other, the result of PCM can give overlapping clusters (i.e., see point prototype locations in Figure 2). As shown in the figure, we cannot always rely on PCM to yield more desirable clustering results compared with FCM. Unfortunately, it can be difficult to determine which cluster membership assignment method is appropriate for any given pattern set.

As mentioned above, various clustering algorithms contain various uncertainties that can control the performance. Proper management of uncertainty in the various parameters that are used in clustering algorithms is essential to the successful development of algorithms to further yield improved clustering results.

The ability of type-2 fuzzy sets to represent and manage uncertainty that exists in fuzzy sets allows us to model various uncertainties, which cannot be appropriately managed by type-1 fuzzy sets [3], [4]. Of course, employment of type-2 fuzzy sets usually increases the computational complexity in comparison with type-1 fuzzy sets due to the additional dimension of having to compute secondary grades for each primary membership. However, if we cannot obtain satisfactory results using type-1 fuzzy sets, then employment of type-2 fuzzy sets for managing uncertainty may allow us to obtain desirable results. Therefore, if the employment of type-2 fuzzy sets can provide significant improvement on performance (depending on the application), then the increase of computational complexity due to type-2 fuzzy sets may be a small price to pay. Also, if we simplify type-2 fuzzy sets to interval type-2 fuzzy sets, the computational complexity can be significantly reduced in appropriate applications that are applicable. The reduction on the computational complexity is due to the property that all the secondary grades for an interval type-2 fuzzy set are uniformly weighted (i.e., all equal to one). Hence, this simplifies the operations on the secondary grades.

In general, the management of uncertainty using interval type-2 fuzzy sets may be applied to various fields where we cannot obtain satisfactory performance with ordinary type-1 fuzzy sets. Most of the published studies focus mainly on type-2 fuzzy logic systems (FLS), and several interesting results have been reported [3]. Rhee and Hwang have studied how to define and manage uncertainty for distance measures when fuzzy membership functions are designed in pattern classification and clustering [5]–[9]. However, there remains plenty of research (compared to other areas) in applying type-2 fuzzy sets to the area of pattern recognition. This is believed to be a promising area of research and, therefore, this article focuses on a particular area, namely clustering.

If one can properly identify the various types of uncertainties embedded in various clustering methods, then expectation for developing algorithms to improve the performance may be highly considered. For this, we focus on the representation and management of uncertainty, which is present in the fuzzy memberships of the patterns that are associated with the various parameters involved in objective function-based clustering. To design and manage the uncertainties for the various parameters, we extend a pattern set to interval type-2 fuzzy sets, which creates a footprint of uncertainty (FOU) for uncertainties in the cluster parameters. Then, we can incorporate this interval type-2 fuzzy set into various objective-based clustering algorithms to observe the effect of managing the uncertainties in the cluster parameters. Finally, with the addition of solutions to type-reduction and defuzzification (i.e., cluster prototype updating and hard-partitioning), objective function-based interval type-2 fuzzy clustering methods can be achieved.

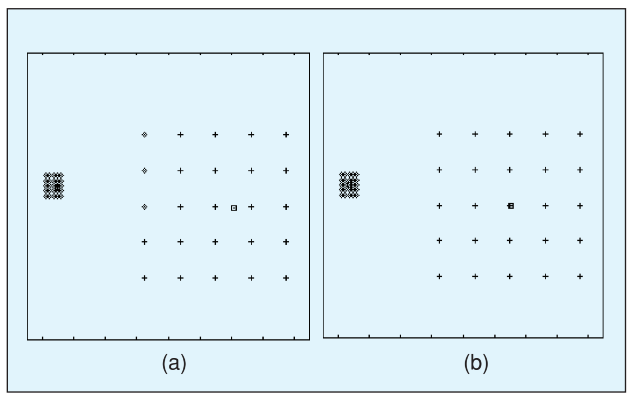


FIGURE 1 FCM clustering results on a pattern set consisting of two square distributed patterns of different volume and density: (a) using fuzzifier $m = 2.0$, (b) using $m = 1.1$.

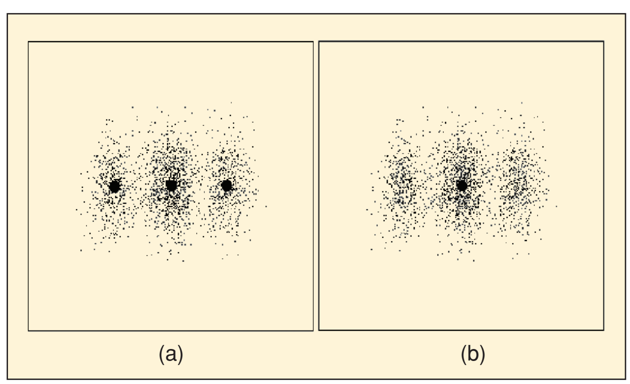


FIGURE 2 Clustering results on synthetically generated data for (a) FCM and (b) PCM.

The following are some uncertainties that may be associated with the parameters used in various objective function-based fuzzy clustering algorithms.

1) Uncertainty in the fuzzifier parameter m in FCM-based algorithms.
2) Uncertainty in the "bandwidth" η_j in PCM-based methods.
3) Uncertainty in the membership assignment involving cluster relativity (i.e., FCM) and typicality (i.e., PCM) based methods.

In what follows, suggestions on how to manage the uncertainties mentioned is discussed.

II. Uncertainties Associated with Objective Function-based Fuzzy Clustering

In general, objective function-based clustering methods are used to minimize the distance between a pattern and cluster prototype such as a point, line, hyperellipsoid or quadric shell, and the goal is to determine the prototype parameters (i.e., center and radius) that best represent the pattern set. In these approaches, iterative algorithms such as C-means are used to seek C partitions of the pattern set. If a pattern set forms "compact" clusters and each cluster is reasonably separable from one other, desirable clustering results can be obtained using C-means. However, in real-world applications, pattern sets of such may be seldom encountered. For example, in several practical cases, a few non-prototypical patterns in a pattern set (e.g., remote patterns distant from the cluster prototype) can cause the cluster prototypes to end up as "poor" representatives of the cluster partitions, since they contribute "fully" in the prototype update. Hence, these patterns may cause undesirable clustering results. To overcome this drawback of C-means, FCM-based clustering algorithms, which consider weight values (memberships) to control the contribution degree of a pattern in determining the parameters of cluster prototypes, were developed. A high-weighted pattern compared to others can play an important role in determining the resulting cluster prototype parameters. On the other hand, the role of a pattern is reduced for patterns that are far from the cluster centers in comparison with highly weighted patterns [1].

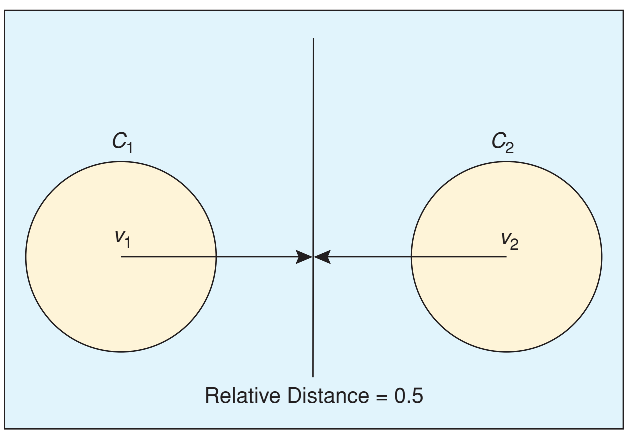


FIGURE 3 Crisp membership assignment in FCM.

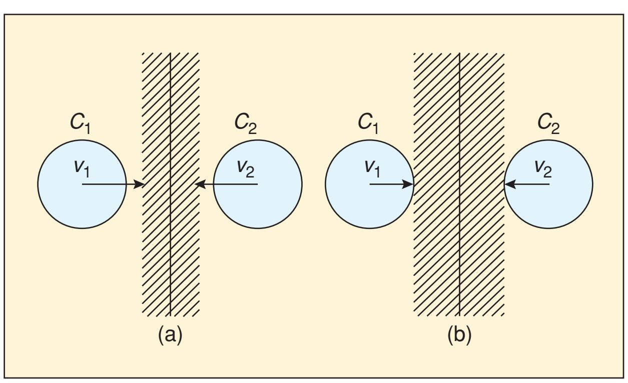


FIGURE 4 Maximum fuzzy region for two clusters of similar volume by fuzzifier m of (a) low degree and (b) high degree.

These approaches can give desirable clustering results for pattern sets that are not well separable; that is, pattern distributions that seem to contain overlapping clusters. The clustering results are due to the fuzzy membership assignment that considers relative distance of the patterns across all clusters. For example, a pattern that is distant from a cluster contributes less in the updating procedure of the cluster prototypes, whereas a pattern located nearby contributes more. If we consider point prototypes in describing the clusters for a pattern set, then the Euclidean distance measure is used to measure the distance between the cluster prototype (center) and patterns. As an example, Figure 3 shows two cluster boundaries C_1 and C_2, where the clusters are considered to be identical in structure and density.

The maximally fuzzy membership locations (vertical line) are where patterns are equally distant from the two cluster centers; that is, the relative distance between a pattern and each cluster center equaling 0.5. In most FCM-based methods, those patterns will contribute (i.e., help in locating the cluster centers) relatively less than a pattern that is located away from the maximally fuzzy membership location during the cluster center update. As a result, the vertical line in the figure can be considered as a "decision" boundary where patterns located to the left (right) of the boundary belong to cluster $C_1(C_2)$. This boundary can be expanded by a fuzzifier parameter m as shown in Figure 4. If the value of fuzzifier m is increased, then the maximum fuzzy boundary becomes wider. The width of the boundary indicates the range one desires to assign maximally fuzzy memberships (based on relative distance) to the patterns. In general, several studies show that reasonable clustering results can be obtained using $m = 2.0$; however, there is hardly a change in the clustering results regardless of fuzzzifier m for types of problems shown in Figure 4.

FCM-based methods are known to perform well on pattern sets that contain clusters that are of similar volume of hyper-spherical shape and density. However, if the clusters in a pattern set are of different density (i.e., different cluster size/different number of patterns ratios or same clusters size/different number of patterns ratios or different cluster size/same number of patterns ratios), then performance of FCM may significantly differ depending on the choice of fuzzifier m. Hence, a good choice of m may be considered to be related to the distribution of the patterns for the pattern set.

For the case shown in Figure 5, the maximum fuzzy boundary can cause undesirable clustering results if fuzzifier m is not assigned properly due to the difference in volume between two clusters.

Patterns that are on the left side of the maximum fuzzy boundary that is part of cluster C_2 intuitively contributes more for cluster C_1 rather than cluster C_2. This suggests that appropriate control of fuzzifier m is needed to improve the clustering performance in FCM for pattern distributions that contain clusters that differ significantly in size. If we set fuzzifier m to a large value, then a wide maximum fuzzy boundary such as in Figure 5 can be obtained. This may seem to be desirable for the viewpoint of cluster C_1; however, the estimated cluster center $\mathbf{v}_1$ will tend to move toward C_2. Hence, $\mathbf{v}_1$ will deviate from the ideal center of cluster C_1. This is also not desirable for the viewpoint of cluster C_2 (an undesirable cluster center may result for C_2). Therefore, the ideal situation is to have the maximum fuzzy region with a wide left region and narrow right region with respect to the maximum fuzzy membership locations as the one shown in Figure 6.

For each region with respect to the vertical line, the relative distance is 0.5. Unfortunately, a maximum fuzzy region of such cannot be represented by FCM since the change of fuzzifier m affects all clusters equally.

We now define and suggest how to manage uncertainty, which exists on the establishment of the maximum fuzzy boundary in FCM. For example, we could use two fuzzifiers to extend memberships of a pattern set to an interval type-2 fuzzy set and apply it to FCM. This will allow us to define and manage the uncertainty of fuzzifier m in FCM. In order to manage the uncertainty, applying interval type-2 fuzzy sets instead of type-2 fuzzy sets will help reduce the computational complexity and simultaneously improve clustering performance. We begin by defining the *uncertainty* for fuzzifier m in the FCM algorithm.

A. Uncertainty in the Fuzzifier m in the FCM Algorithm

Conventional FCM is an objective function-based iterative algorithm that assigns fuzzy memberships to pattern data and updates the prototype parameters of the cluster according to the assigned memberships. The assigned memberships play a role as weight values; that is, they represent the contribution degree of a pattern on estimating the cluster prototypes. The amount of contribution depends on the selection of fuzzifier m. Details are briefly described as follows.

FCM algorithms are based on minimizing the objective function

$$J(\mathbf{U},\mathbf{V}) = \sum_{j=1}^{C}\sum_{i=1}^{N}(u_j(\mathbf{x}_i))^m d_{ji}^2 \quad \text{subject to}$$

$$\sum_{j=1}^{C} u_{ji} = 1 \quad \text{for all } i, \tag{1}$$

where $\mathbf{U} = [u_j(\mathbf{x}_i)]$ is a fuzzy C-partition of the patterns $\{\mathbf{x}_i\}$, $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_c)$ are the cluster prototypes, and m denotes the fuzzifier. Minimizing (1), the update for membership $u_j(\mathbf{x}_i)$ can be described as

$$u_j(\mathbf{x}_i) = \frac{1}{\sum_{k=1}^{C}\left(\frac{d_{ji}}{d_{ki}}\right)^{2/(m-1)}}. \tag{2}$$

Distance $d_{ji}(d_{ki})$ denotes the distance between cluster prototype $\mathbf{v}_j(\mathbf{v}_k)$ and pattern $\mathbf{x}_i$. Equation (2) presents the fuzzy membership of pattern $\mathbf{x}_i$ that is assigned by the relative distance between pattern $\mathbf{x}_i$ and cluster prototype $\mathbf{v}_j$. The updating of the prototypes depends on the distance measure used. If the distance is defined as $d_{ji}^2 = (\mathbf{x}_i - \mathbf{v}_j)^T \mathbf{A}_j (\mathbf{x}_i - \mathbf{v}_j)$, where $\mathbf{v}_j$ is the prototype defined as cluster centers, then different hyper-quadric cluster shapes can be achieved depending on Matrix $\mathbf{A}_j$,. Table 1 shows other distance measures that can be used to create a variety of FCM algorithms.

If $\mathbf{V}$ are point prototypes (i.e., cluster centers), then the memberships in (2) depend on the relative distance between pattern $\mathbf{x}_i$ and all of the cluster centers, and fuzzifier m. Figure 7 illustrates the variation in the memberships according to the relative distance for various values of m.

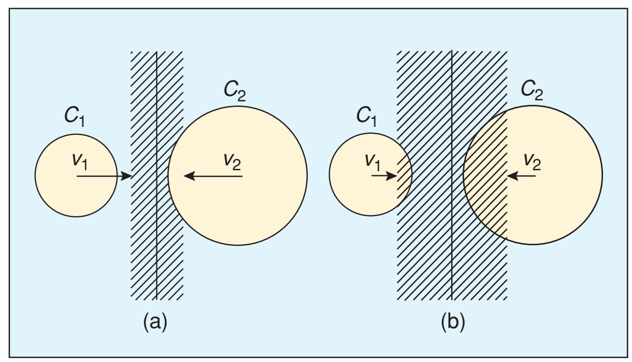


FIGURE 5 Maximum fuzzy region for two clusters of different volume by fuzzifier m of (a) low degree and (b) high degree.

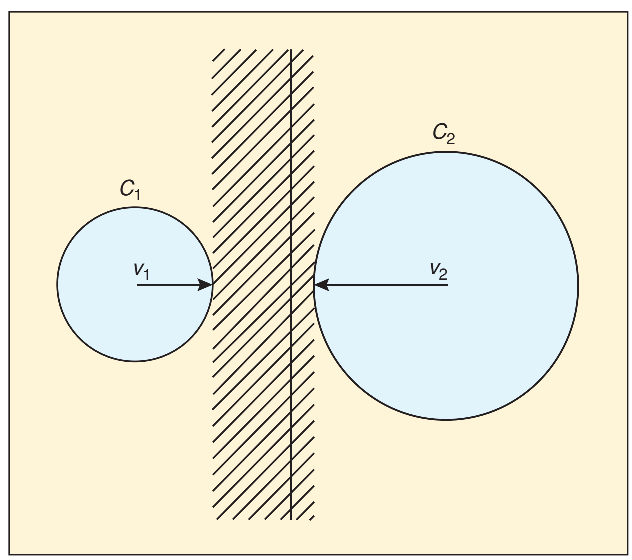


FIGURE 6 Desirable Maximum fuzzy region.

TABLE 1 Distance measures used to create a variety of objective function-based clustering algorithms.

DISTANCE d_{ji}^2	PROTOTYPE(S)	CLUSTER SHAPE	UPDATE EQUATIONS
$\lVert \mathbf{x}_i - \mathbf{v}_j \rVert^2$	CENTER $\mathbf{v}_j$	HYPER-SPHERICAL	$\mathbf{v}_j = \frac{\sum_{i=1}^{N} (u_j(\mathbf{x}_i))^m \mathbf{x}_i}{\sum_{i=1}^{N} u_j(\mathbf{x}_i)^m}$
$(\lVert \mathbf{x}_i - \mathbf{v}_j \rVert - r_j)^2$	CENTER $\mathbf{v}_j$ RADIUS r_j	HYPER-SPHERICAL SHELLS	$\mathbf{v}_j = \frac{\sum_{i=1}^{N} (u_j(\mathbf{x}_i))^m \mathbf{x}_i}{\sum_{i=1}^{N} u_j(\mathbf{x}_i)^m}$
			$r_j = \frac{\sum_{i=1}^{N} (u_j(\mathbf{x}_i))^m \mathbf{x}_i \lVert \mathbf{x}_i - \mathbf{v}_j \rVert}{\sum_{i=1}^{N} u_j(\mathbf{x}_i)^m}$
$\lvert \mathbf{F}_j \rvert^{1/n} (\mathbf{x}_i - \mathbf{v}_j)^T \mathbf{F}_j^{-1} (\mathbf{x}_i - \mathbf{v}_j)$	CENTER $\mathbf{v}_j$ FUZZY COVARIANCE $\mathbf{F}_j$	HYPER-QUADRIC	$\mathbf{v}_j = \frac{\sum_{i=1}^{N} (u_j(\mathbf{x}_i))^m \mathbf{x}_i}{\sum_{i=1}^{N} u_j(\mathbf{x}_i)^m}$
			$\mathbf{F}_j = \frac{\sum_{i=1}^{N} (u_j(\mathbf{x}_i))^m (\mathbf{x}_i - \mathbf{v}_j)(\mathbf{x}_i - \mathbf{v}_j)^T}{\sum_{i=1}^{N} (u_j(\mathbf{x}_i))^m}$

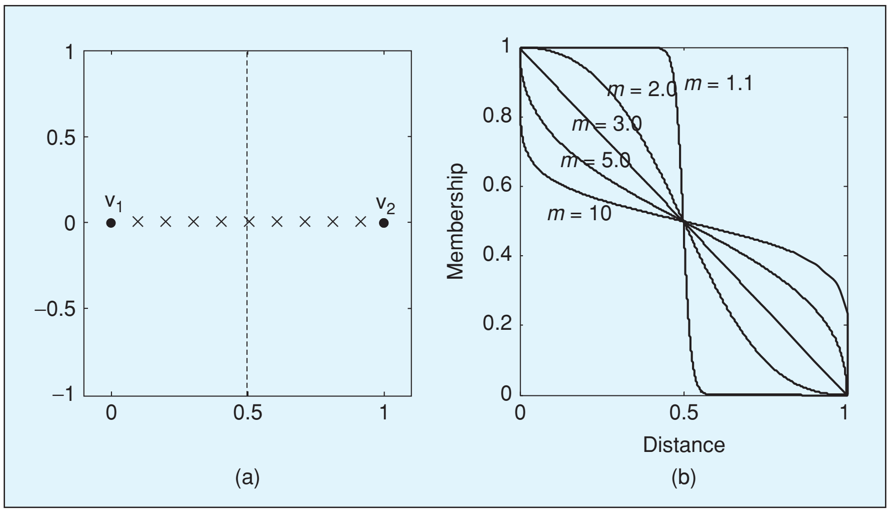


FIGURE 7 Example illustrating the variation of fuzzy memberships: (a) pattern set consisting of two centers and (b) membership plots corresponding to center v_1 for patterns located between v_1 and v_2 in (a) for various values of m.

Figure 7(a) shows two cluster centers $\mathbf{v}_1$ and $\mathbf{v}_2$ for a given pattern set (note that the patterns in which $\mathbf{v}_1$ and $\mathbf{v}_2$ are obtained are not shown). The membership plots corresponding to center $\mathbf{v}_1$ for patterns that lie between the two cluster centers for various values of fuzzifier m can be shown as in Figure 7(b). When $m \to 1$, the memberships are maximally crisp (hard); that is, patterns that are located just left (right) of the maximum fuzzy boundary are assigned full (0) membership for center $\mathbf{v}_1(\mathbf{v}_2)$. Conversely, when $m \to \infty$, the memberships are maximally fuzzy, which means patterns that are only located at the centers are assigned full (0) membership, otherwise they are assigned memberships of 0.5. As mentioned earlier, if the geometry of the clusters in a pattern set is of similar volume and density, then change in fuzzifier m will not significantly affect the clustering result in FCM. However, if there is a significant difference in density among clusters in a pattern set, then the choice of m will give inconsistent clustering results. The reason is that an impertinent establishment of maximum fuzzy boundary in FCM can provide undesirable updating of the cluster centers. This was previously emphasized in the explanation of Figure 5. A more desirable establishment of the maximum fuzzy region as in Figure 6 can allow for desirable clustering results in the FCM. Again, due to the constraint on the memberships, we cannot design this region with any particular single value of fuzzifier m to be used in the FCM. However, if we can somehow simultaneously incorporate various values of fuzzifier m, then we may be able to design a desirable maximum fuzzy region such as the one in Figure 6. This is shown in Figure 8.

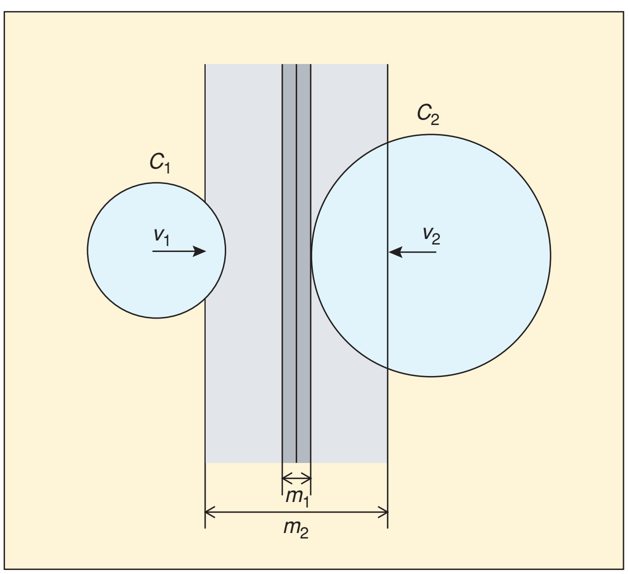


FIGURE 8 Uncertain maximum fuzzy region by two fuzzifiers ($m_1 < m_2$).

Unlike Figures 5 and 6, the established maximum fuzzy region in Figure 8 should be considered uncertain although we establish it with two fuzzifiers m_1 and m_2, which represent different fuzzy degrees. Therefore, we should consider memberships for a pattern as uncertain (fuzzy) instead of as certain (crisp) in the FCM. In other words, memberships for a pattern should not have to rely on a single specific fuzzifier m as in FCM. In this article, we define uncertainty of fuzzy memberships for a pattern associated with a cluster prototype as an *uncertain* maximum fuzzy boundary such as the case in Figure 8.

B. Uncertainty in the "Bandwidth" η in the PCM Algorithm

PCM is also an iterative algorithm that assigns fuzzy memberships to pattern data as in the FCM. The assigned memberships play a role of "compatibility with a prototype;" that is, they represent the contribution degree of possibility of a pattern belonging to a cluster. The amount of contribution depends on the selection of fuzzifier m and bandwidth η. Details are briefly described as follows.

PCM algorithms are based on minimizing the objective function

$$J(\mathbf{U},\mathbf{V}) = \sum_{j=1}^{C}\sum_{i=1}^{N}(u_j(\mathbf{x}_i))^m d_{ji}^2 + \sum_{j=1}^{C}\eta_j\sum_{i=1}^{N}(1-u_j(\mathbf{x}_i))^m, \tag{3}$$

where bandwidth η_j denotes the distance at which the membership value of a pattern in a cluster equals 0.5. All other parameters are defined the same as with the FCM objective function. Minimizing (3), the update for membership $u_j(\mathbf{x}_i)$ can be described as

$$u_j(\mathbf{x}_i) = \frac{1}{1+\left(\frac{d_{ji}^2}{\eta_j}\right)^{1/(m-1)}}. \tag{4}$$

Equation (4) presents the fuzzy membership of pattern $\mathbf{x}_i$ that is assigned by the absolute distance between pattern $\mathbf{x}_i$ and cluster prototype $\mathbf{v}_j$. As in FCM, the updating of the prototypes depends on the distance measure used. Distance measures in Table 1 can be used to create a variety of PCM algorithms. If $\mathbf{V}$ are point prototypes, the memberships in (4) depend on the absolute distance between pattern $\mathbf{x}_i$ and cluster centers, fuzzifier m, and bandwidth η. Figure 9 illustrates the variation in the memberships according to the absolute distance for various values of m and η. The plots can be described in a manner similar to the previous discussion regarding Figure 7.

C. Uncertainty in Selection of Objective Function

Another entity of uncertainty present in objective function-based clustering is in the choice of objective function to use. This can be illustrated by considering the two objective functions that are used for FCM and PCM. One is based on cluster relativity (FCM) and the other is based on cluster typicality (PCM). This is the same as considering the uncertainty between the membership assignment for cluster relativity (i.e., FCM) and typicality (i.e., PCM). For both objective functions, the fuzzifier m is assigned the same value. It may be difficult to determine whether to employ FCM or PCM to cluster a pattern set. For the case when a pattern set includes noise patterns, PCM may obtain better clustering results compared to FCM. However, overlapping clusters may result in PCM for clusters that are located close to each other. Hence, this can produce worse results than FCM. Therefore, when we execute C-means, we may consider a method to define and manage the uncertainty in cluster membership values that share the properties of cluster relativity and typicality in FCM and PCM, respectively. To obtain such a method, we may want to look for the following properties.

1) Property of cluster relativity: lowering the effect of noise patterns by the aid of membership assignment that is assigned to noise patterns that involves cluster typicality.

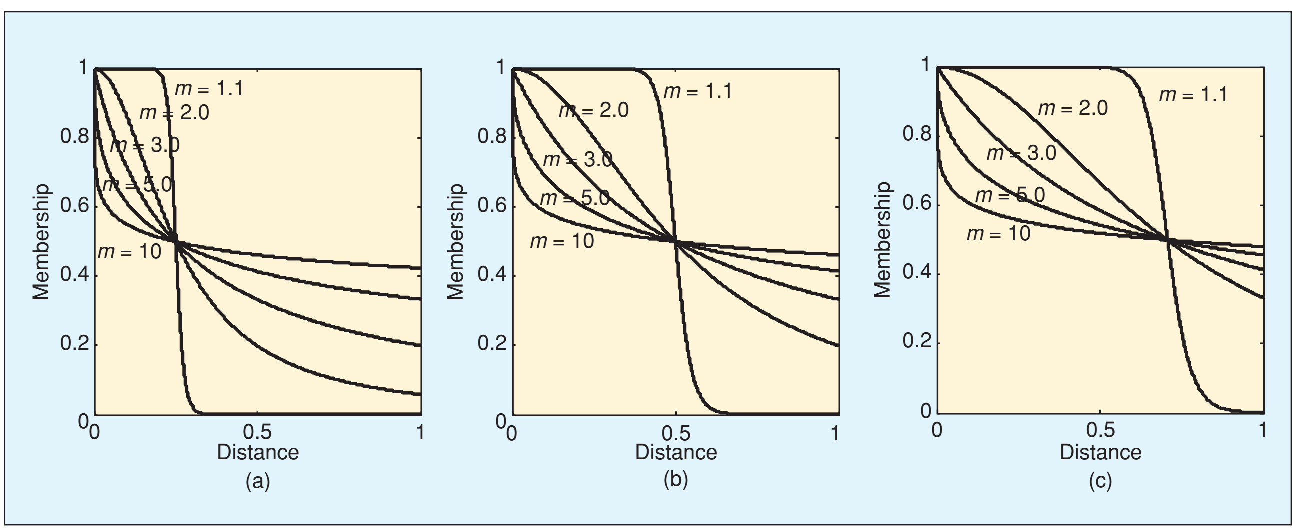


FIGURE 9 Example illustrating the variation of fuzzy memberships in PCM for (a) $\eta_j = 0.25$, (b) $\eta_j = 0.50$, and (c) $\eta_j = 0.75$ for various values of m.

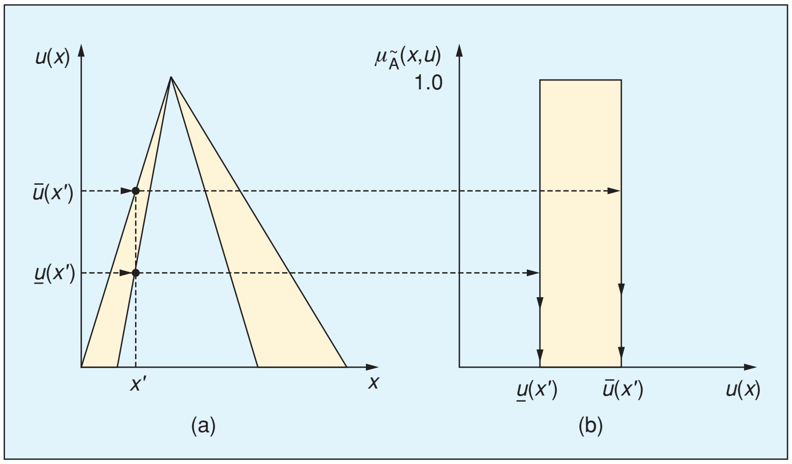


FIGURE 10 Example of (a) an interval type-2 fuzzy set and (b) the vertical slice of sample x'.

$$\tilde{\mathbf{A}} = \{((\mathbf{x}, u), \mu_{\tilde{\mathbf{A}}}(\mathbf{x}, u)) | \forall \mathbf{x} \in \mathbf{A}, \forall u \in J_{\mathbf{x}} \subseteq [0, 1], \mu_{\tilde{\mathbf{A}}}(\mathbf{x}, u) = 1\} \tag{5}$$

2) Property of cluster typicality: reduction in the sensitivity in the initialization by aid of membership assignment that involves cluster relativity to prevent "overlapping" cluster that occurs in mode-seeking algorithms such as PCM.

The appropriate management of defined uncertainty mentioned above may hopefully improve the performance in objective function-based clustering. This is explained in the following section by incorporating interval type-2 fuzzy sets.

III. Extension to Interval Type-2 Fuzzy Sets

When we represent uncertainty for a fuzzy set with an interval type-2 fuzzy set, the primary membership $J_{\mathbf{x}_i}$ of a pattern $\mathbf{x}_i$ can be represented as a membership interval with all secondary grades of the primary memberships equaling one [1]. An interval type-2 fuzzy set $\tilde{\mathbf{A}}$ can be represented as

Figure 10 illustrates an example of an interval type-2 fuzzy set where the gray-shaded region indicates the footprint of uncertainty (FOU).

In the figure, the membership value for a 1-D sample x' is represented by the interval between upper ($\bar{u}(x')$) and lower ($\underline{u}(x')$) membership. Therefore, each x' has a primary membership interval as $J_{x'} = [\underline{u}(x'), \bar{u}(x')]$ and the vertical slice of x' is shown in Figure 10(b), where the secondary grade for the primary membership of each x' equals one, according to the property of interval type-2 fuzzy sets.

In general, computation of fuzzy memberships in FCM is achieved by computing the relative distance among the patterns and cluster prototypes according to (2), mentioned above. Now, to define the interval of primary membership for a pattern, we define the lower and upper interval memberships using two different values of m. The primary memberships that extend pattern $\mathbf{x}_i$ by interval type-2 fuzzy sets become

$$\bar{u}_j(\mathbf{x}_i) = \begin{cases} \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_1-1)}} & \text{if } \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_1-1)}} > \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_2-1)}} \\ \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_2-1)}} & \text{otherwise} \end{cases}$$

and

$$\underline{u}_j(\mathbf{x}_i) = \begin{cases} \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_1-1)}} & \text{if } \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_1-1)}} \leq \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_2-1)}} \\ \dfrac{1}{\sum_{k=1}^{C} (d_{ji}/d_{ki})^{2/(m_2-1)}} & \text{otherwise} \end{cases} \tag{6}$$

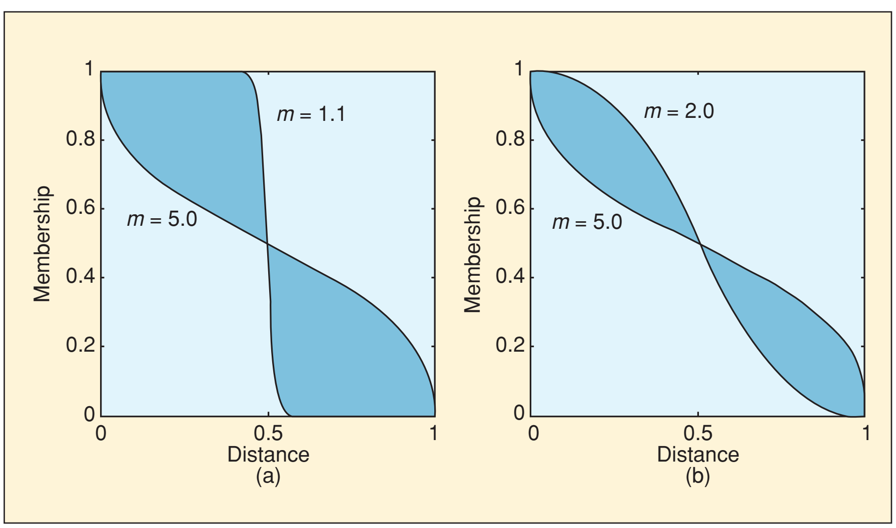


FIGURE 11 Footprint of uncertainty (FOU) of an interval type-2 fuzzy set for FCM: (a) $m = 1.1$ and 5.0 and (b) $m = 2.0$ and 5.0.

In (6), m_1 and m_2 are fuzzifiers that represent different fuzzy degrees. When we define the interval of primary membership for a pattern, we use the highest and lowest primary membership of the interval for a pattern. These values are denoted by upper and lower membership for a pattern, respectively. The uses of fuzzifiers, which represent different fuzzy degrees, give different objective functions to be minimized in FCM (i.e., $m = m_1$ and $m = m_2$ in (1)).

Figure 11 illustrates an example of an interval type-2 fuzzy set according to (6) for a two-cluster case as in Figure 5(a) for fuzzifier combinations of $m_1 = 1.1$ and $m_2 = 5.0$, and $m_1 = 2.0$ and $m_2 = 5.0$.

The difference in fuzzy memberships denotes the FOU as indicted by the shaded gray areas. The membership for patterns located at each current cluster center (the center location at the current iteration) is set to 1.0 and set to 0 for the other corresponding center regardless of the fuzzy degree in a pattern set. Hence, there is no FOU for these patterns. For patterns, which are relatively equal in distance from the two cluster centers, memberships are set to 0.5. There is also no FOU for these patterns. Hence, it is reasonable to consider that there exists no uncertainty for the maximum and minimum fuzzy points.

In PCM, the fuzzy memberships are computed by the absolute distance among the patterns and cluster prototypes according to (3). For the interval of primary membership for a pattern, we define the lower and upper interval memberships using two different values of η_j. The primary memberships that extend pattern $\mathbf{x}_i$ by interval type-2 fuzzy sets become

$$\bar{u}_j(\mathbf{x}_i) = \begin{cases} \frac{1}{1+\left(d_{ji}^2/\eta_{j1}\right)^{1/(m-1)}} & \text{if } \frac{1}{1+\left(d_{ji}^2/\eta_{j1}\right)^{1/(m-1)}} > \frac{1}{1+\left(d_{ji}^2/\eta_{j2}\right)^{1/(m-1)}} \\ \frac{1}{1+\left(d_{ji}^2/\eta_{j2}\right)^{1/(m-1)}} & \text{otherwise} \end{cases}$$

and

$$\bar{u}_j(\mathbf{x}_i) = \begin{cases} \frac{1}{1+\left(d_{ji}^2/\eta_{j1}\right)^{1/(m-1)}} & \text{if } \frac{1}{1+\left(d_{ji}^2/\eta_{j1}\right)^{1/(m-1)}} \le \frac{1}{1+\left(d_{ji}^2/\eta_{j2}\right)^{1/(m-1)}} \\ \frac{1}{1+\left(d_{ji}^2/\eta_{j2}\right)^{1/(m-1)}} & \text{otherwise} \end{cases} \quad (7)$$

In (7), η_{j1} and η_{j2} are bandwidths that represent different locations where the memberships equal 0.5. The use of different bandwidths gives different objective functions to be minimized in PCM (i.e., $\eta_j = \eta_{j1}$ and $\eta_j = \eta_{j2}$ in (3)). Figure 12 illustrates an example of an interval type-2 fuzzy set according to (7) for the two-cluster case as in Figure 7(a) for m set to 2.0, and bandwidth combinations of $\eta_j = 0.25$ and $\eta_j = 0.75$. As in the FCM described above, a different interval may be defined by using two different values of m. The primary memberships that extend pattern $\mathbf{x}_i$ by interval type-2 fuzzy sets can be obtained in a manner similar to that of (6). Figure 13 shows the FOU for $\eta_j = 0.5$ and fuzzifier combinations of $m_1 = 2.0$ and $m_2 = 5.0$.

For the uncertainty in the selection of objective function, the fuzzy memberships are computed by both the relative and absolute distances among the patterns and cluster prototypes according to (1) and (3), respectively. To define the interval of primary membership for a pattern, the lower and upper interval memberships can be obtained using the two memberships. The primary memberships that extend pattern $\mathbf{x}_i$ by interval type-2 fuzzy sets become

$$\bar{u}_j(\mathbf{x}_i) = \begin{cases} \frac{1}{\sum_{k=1}^{C}(d_{ji}/d_{ki})^{2/(m-1)}} & \text{if } \frac{1}{\sum_{k=1}^{C}(d_{ji}/d_{ki})^{2/(m-1)}} > \frac{1}{1+\left(d_{ji}^2/\eta_j\right)^{1/(m-1)}} \\ \frac{1}{1+\left(d_{ji}^2/\eta_j\right)^{1/(m-1)}} & \text{otherwise} \end{cases}$$

and

$$\underline{u}_j(\mathbf{x}_i) = \begin{cases} \frac{1}{\sum_{k=1}^{C}(d_{ji}/d_{ki})^{2/(m-1)}} & \text{if } \frac{1}{\sum_{k=1}^{C}(d_{ji}/d_{ki})^{2/(m-1)}} \le \frac{1}{1+\left(d_{ji}^2/\eta_j\right)^{1/(m-1)}} \\ \frac{1}{1+\left(d_{ji}^2/\eta_j\right)^{1/(m-1)}} & \text{otherwise} \end{cases} \quad (8)$$

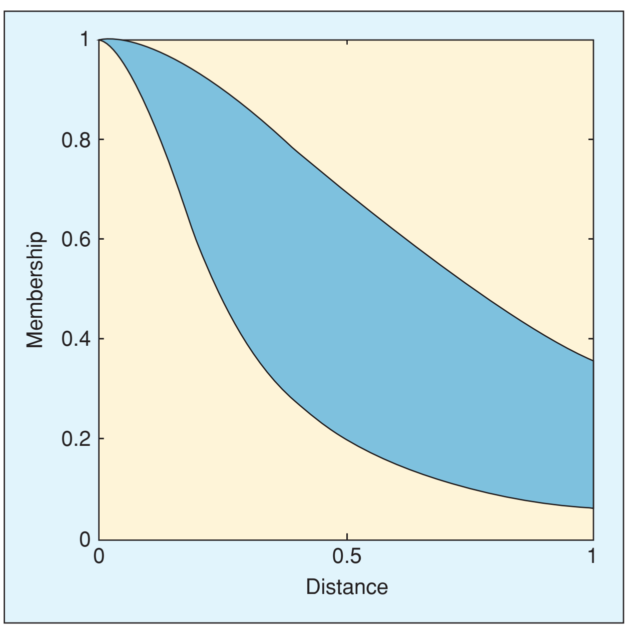


FIGURE 12 Footprint of uncertainty (FOU) of an interval type-2 fuzzy set for PCM ($m = 2.0$) for bandwidth combinations of $\eta_j = 0.25$ and $\eta_j = 0.75$.

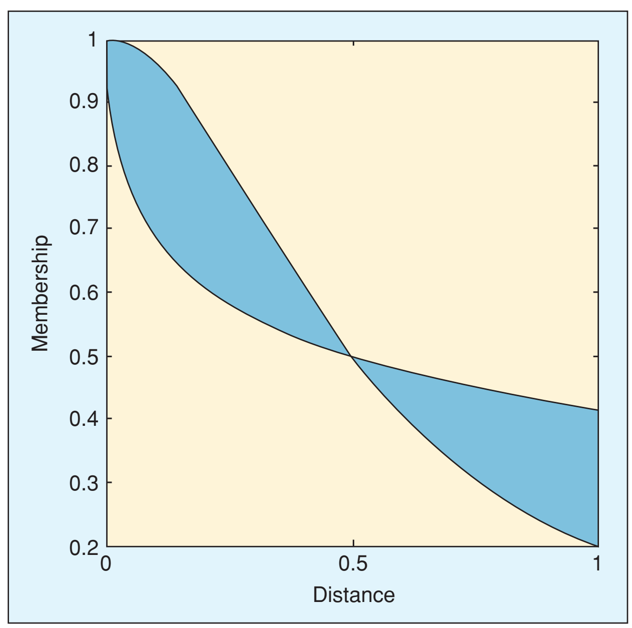


FIGURE 13 Footprint of uncertainty (FOU) of an interval type-2 fuzzy set for PCM ($\eta_j = 0.5$) for fuzzifier combinations of $m_1 = 2.0$ and $m_2 = 5.0$.

Figure 14 illustrates an example of an interval type-2 fuzzy set according to (8) for a two-cluster case as in Figure 7(a) using fuzzifier $m = 2.0$ and for $\eta_j = 0.25$, $\eta_j = 0.50$, and $\eta_j = 0.75$.

The application of the various interval type-2 fuzzy sets described above can be included in the modification of prototype-updating and hard-partition procedure in clustering algorithms. Hence, it can lead us to objective function-based interval type-2 fuzzy clustering methods.

IV. Overview of Objective Function-based Interval Type-2 Fuzzy Clustering

The uncertainty defined in an interval type-2 fuzzy set should be managed appropriately through all steps in objective function-based clustering. For this, we manage the uncertainty of the parameters discussed above by the following two procedures.

1) Procedure for updating cluster prototypes.
2) Procedure for hard-partitioning (defuzzification) in the final clustering decision.

First by modification of the procedure for updating cluster prototypes is by the extension of interval type-2 fuzzy sets while executing the clustering algorithm. For this, type-reduction and defuzzification methods using type-2 fuzzy operations are needed. Among the various choices of type-reduction methods, the generalized centroid (GC) type-reducer is most attractive for clustering since it is similar to the center-update method in prototype-based clustering algorithms. It is to be noted that there exists other type-reduction methods. For example, if we choose height type-reduction, computation complexity can be reduced. However, due to its simple estimation, this can give us inaccurate cluster prototypes leading to undesirable clustering results. Details on other methods can be found in [3].

The centroid of a type-1 fuzzy set can be computed as

$$\mathbf{v}_{\mathbf{x}} = \frac{\sum_{i=1}^{N} x_i u(x_i)}{\sum_{i=1}^{N} u(x_i)}, \tag{9}$$

and by applying the extension principle and considering interval type-2 fuzzy sets, estimated centers by the centroid type-reduction can be obtained by interval

$$\begin{aligned}\mathbf{v}_{\tilde{\mathbf{X}}} &= [\mathbf{v}_L, \mathbf{v}_R] \\ &= \sum_{u(\mathbf{x}_1)\in J_{\mathbf{x}_1}} \cdots \sum_{u(\mathbf{x}_N)\in J_{\mathbf{x}_N}} 1 \Bigg/ \frac{\sum_{i=1}^{N} \mathbf{x}_i u(\mathbf{x}_i)^m}{\sum_{i=1}^{N} u(\mathbf{x}_i)^m}.\end{aligned} \tag{10}$$

Details on how to obtain the above equations can be found in [1].

All centers among the interval are needed to be computed in order to represent interval $[\mathbf{v}_L, \mathbf{v}_R]$. For this, embedded type-2 fuzzy sets for an interval type-2 fuzzy set that are represented by all upper and lower values of primary membership $J_{\mathbf{x}_i}$ are considered. Primary membership $J_{\mathbf{x}_i}$ can be organized for a pattern $\mathbf{x}_i$ with $\bar{u}(\mathbf{x}_i)$ and $\underline{u}(\mathbf{x}_i)$.

In general, there exists 2^N embedded fuzzy sets in an interval type-2 fuzzy set that consist of N patterns. This is also the number of center intervals that can be obtained from the embedded fuzzy sets as in (10). The union of estimated center intervals from the embedded fuzzy sets constitutes an interval type-1 fuzzy set as

$$\begin{aligned}\mathbf{Y} &= \{1/\mathbf{v}_1, \ldots, 1/\mathbf{v}_{2^N}\} \\ &= 1.0/[\mathbf{v}_L, \mathbf{v}_R] \qquad \mathbf{v}_L \le \forall \mathbf{v}_i \le \mathbf{v}_R.\end{aligned} \tag{11}$$

From this approach, we can obtain a type-reduced fuzzy set for an interval type-2 fuzzy set. However, this approach should be avoided in practical cases since the amount of computation can drastically increase if the number of patterns increases.

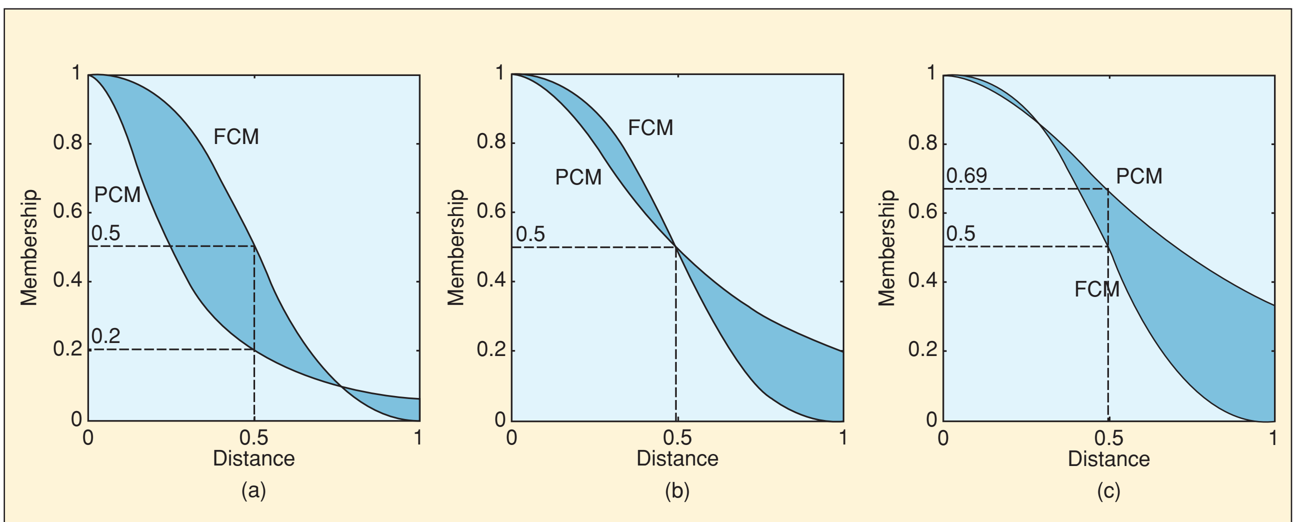


FIGURE 14 Footprint of uncertainty (FOU) of an interval type-2 fuzzy set in (8) for $m = 2.0$ and (a) $\eta_j = 0.25$, (b) $\eta_j = 0.50$, and (c) $\eta_j = 0.75$.

Therefore, we consider an alternative approach, which can reduce the computational load. Karnik et al. have suggested an iterative algorithm that estimates both ends of an interval; that is, estimating $\mathbf{v}_R$ and $\mathbf{v}_L$ without using all of the embedded fuzzy sets for an interval type-2 fuzzy set. These can be estimated by arranging the pattern set in ascending order and finding the location between the upper and lower membership for a pattern set. It is required for an ascending-ordered pattern set to represent a left value $\mathbf{v}_L$ and a right value $\mathbf{v}_R$ of an interval cluster center. For our proposed method, we apply the iterative algorithm to estimate the cluster centers. The iterative algorithm is used to estimate the maximum value $\mathbf{v}_R$ and minimum value $\mathbf{v}_L$. Details can be found in [3].

If we consider point prototypes, the estimated interval type-1 fuzzy set for the cluster centers can be represented as $\mathbf{v}_j = 1.0/[\mathbf{v}_L, \mathbf{v}_R]$.

A crisp center for estimated center $\mathbf{v}_j$ is simply obtained by a defuzzification method for a fuzzy set since estimated center $\mathbf{v}_j$ is an interval type-1 fuzzy set. This can be computed as

$$\mathbf{v}_j = \frac{\sum_{\mathbf{v}\in J_{\mathbf{v}_j}} (u(\mathbf{v}))\mathbf{v}}{\sum_{\mathbf{v}\in J_{\mathbf{v}_j}} (u(\mathbf{v}))} = \frac{\mathbf{v}_L + \mathbf{v}_R}{2}. \tag{12}$$

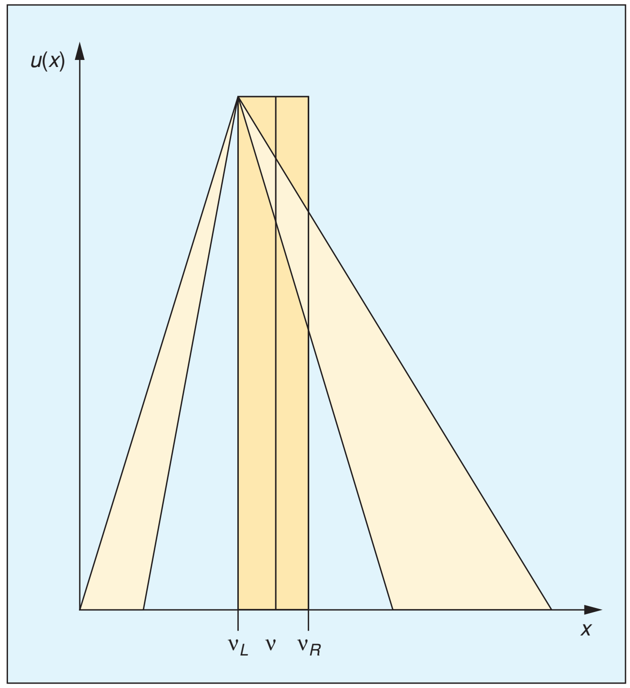


FIGURE 15 Example of a crisp center and interval type-1 fuzzy center.

When applying the above procedures (i.e., type-reduction and defuzzification) as for the example in Figure 10, the center of an interval type-2 fuzzy set can be obtained as shown in Figure 15.

The estimated center from type-reduction and defuzzification signifies the management of uncertainty, which is represented by a membership interval that applies two different values of m in FCM. In this article, we manage the membership interval according to two different values of m. This is achieved by the modified procedure of center-updating, which contains type-reduction and defuzzification. The management of uncertainty with an interval type-2 fuzzy set can yield more desirable cluster centers compared with FCM that uses a single value of m in the case when pattern distributions contain partitions of different size volumes. This is greatly due to the extra degree of freedom where two values of m are used. These values are chosen to appropriately describe the fuzzy degree for the clusters.

As a final clustering result, estimation of the cluster centers is obtained by hard-partitioning of the fuzzy memberships for a pattern. In other words, a pattern is assigned to a cluster according to a fuzzy membership. Hence, we can consider hard-partitioning as defuzzification for a fuzzy set. Hard-partitioning in FCM can be summarized as:

$$\begin{aligned}&\text{If } u_j(\mathbf{x}_i) > u_k(\mathbf{x}_i) \quad k = 1, \ldots, C \text{ and } j \neq k\\ &\text{then } \mathbf{x}_i \text{ is assigned to cluster } j.\end{aligned} \tag{13}$$

Unfortunately, we cannot apply (13) directly to our proposed method since the pattern set is extended to an interval type-2 fuzzy set. We need to reduce the type of an interval type-2 fuzzy set before hard-partitioning. Type-reduction before hard-partition should be handled carefully since upper and lower membership $\bar{u}(\mathbf{x}_i)$ and $\underline{u}(\mathbf{x}_i)$ cannot be used direct-

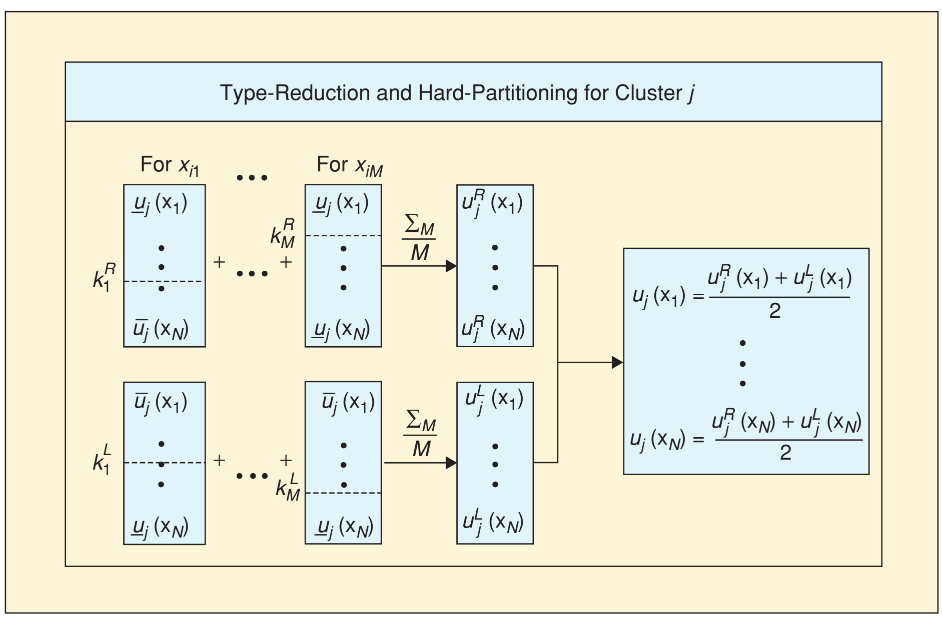


FIGURE 16 Type-reduction for pattern set X.

ly in achieving type-reduction before hard-partitioning.

In our proposed method, we perform type-reduction in order to estimate cluster centers. For this, left memberships $u_j^L(\mathbf{x}_i)$ and right memberships $u_j^R(\mathbf{x}_i)$ for all patterns have already been estimated to organize $\mathbf{v}_L$ and $\mathbf{v}_R$, respectively. Therefore, type-reduction can be achieved using $u_j^L(\mathbf{x}_i)$ and $u_j^R(\mathbf{x}_i)$ to partition a pattern set into clusters. In this approach, hard-partitioning can be obtained as follows.

$$\text{Type-reduction: } u_j(\mathbf{x}_i) = \frac{u_j^R(\mathbf{x}_i) + u_j^L(\mathbf{x}_i)}{2}$$
$$j = 1, \ldots, C$$

and

$$\text{Hard partitioning: Same as in (13).} \tag{14}$$

In (14), memberships $u_j^L(\mathbf{x}_i)$ and $u_j^R(\mathbf{x}_i)$ are also different according to each feature for a pattern. Therefore, a representative value for left and right membership needs to be computed for each feature. This can be obtained as

$$u_j^R(\mathbf{x}_i) = \frac{\sum_{l=1}^{M} u_{jl}(\mathbf{x}_i)}{M},$$
$$\text{where } u_{jl}(\mathbf{x}_i) = \begin{cases} \bar{u}_j(\mathbf{x}_i) & \text{if } x_{il} \text{ uses } \bar{u}_j(\mathbf{x}_i) \text{ for } \mathbf{v}_j^R \\ \underline{u}_j(\mathbf{x}_i) & \text{otherwise} \end{cases}$$

and

$$u_j^L(\mathbf{x}_i) = \frac{\sum_{l=1}^{M} u_{jl}(\mathbf{x}_i)}{M},$$
$$\text{where } u_{jl}(\mathbf{x}_i) = \begin{cases} \bar{u}_j(\mathbf{x}_i) & \text{if } x_{il} \text{ uses } \bar{u}_j(\mathbf{x}_i) \text{ for } \mathbf{v}_j^L \\ \underline{u}_j(\mathbf{x}_i) & \text{otherwise} \end{cases} \tag{15}$$

In (15), M denotes the number of features for pattern $\mathbf{x}_i$. Figure 16 illustrates the procedure of type-reduction before hard-partitioning. Our overall proposed interval type-2 method can be summarized in Figure 17.

As a final comment, the method described above expresses the uncertainty for a pattern set with an interval of fuzzy membership, which can be obtained by using two fuzzifiers m_1 and m_2. However, if the same value for the two fuzzifiers is used, ($m_1 = m_2$), we can naturally obtain a type-1 fuzzy set that has no uncertainty for its memberships. In this case, the result of interval type-2 FCM becomes exactly the same as FCM since the results of center-updating and hard-partitioning in our proposed method are identical to the results in FCM. Therefore, we conclude that our proposed method can be considered as a generalized form of FCM.

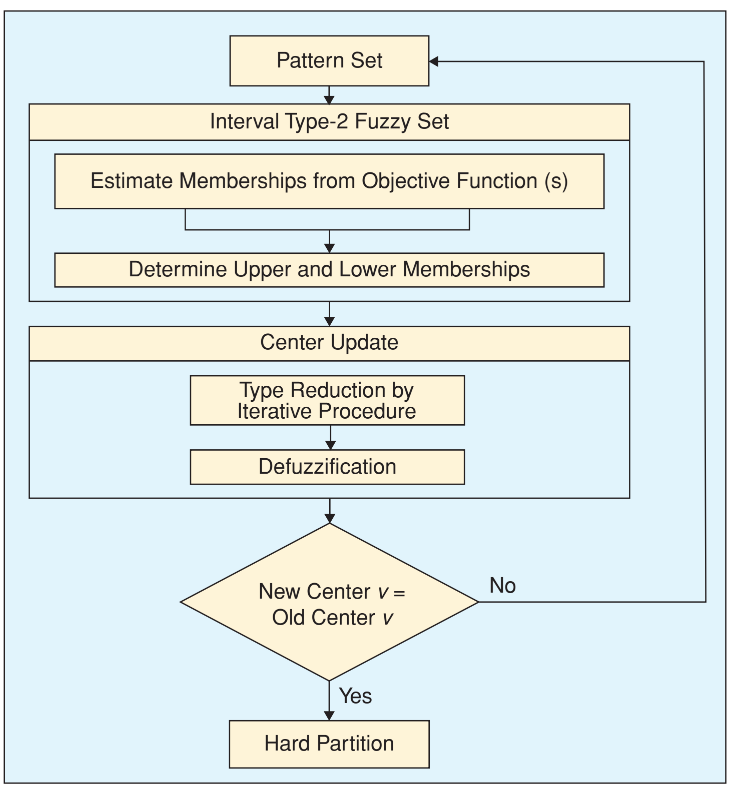


FIGURE 17 Interval type-2 fuzzy clustering.

V. Experimental Results

In this section, we compare the results between the type-1 FCM and the proposed interval type-2 fuzzy clustering for several pattern sets.

A. Different volume/Equal low number "Squares" Data

By changing the volume of clusters in a pattern set, we observe the effectiveness of interval type-2 FCM in comparison with FCM. Figure 18 shows a simple example that illustrates the effect of fuzzifier m in FCM for a pattern set that contains clusters of different volume and equal number of patterns.

The pattern set contains two clusters each containing nine patterns with different cluster density. Figure 18(a) shows the final cluster partition for the best possible clustering result using FCM ($m = 4.0$). Figure 18(b) shows the FCM result using $m = 2.0$, which is generally used when performing FCM. As shown in the figure, the three left vertical patterns of the big cluster are located closer toward the small cluster from the midpoint of the two ideal cluster centers (locations where pattern memberships are maximally fuzzy, i.e., relative distance is 0.5). Therefore, with respect to the small cluster, memberships assigned to those patterns are larger when $m = 2.0$, than for when $m = 4.0$. This is due to the fact that the maximally fuzzy boundary is less wide when $m = 2.0$, than for when $m = 4.0$. Hence, the center of the small cluster will eventually move more to the right than for the case when $m = 4.0$. As a result, final partitioning of FCM using $m = 4.0$ gives better (but not desirable) results than for when $m = 2.0$. However, in Figure 18(c), the result of interval type-2 FCM by establishing a FOU using $m_1 = 2.0$ and

$m_2 = 10.0$ gives desirable results with cluster centers located at their prototypical locations. The reason for this is that our proposed method utilizes the appropriate value of m during center-updating as explained earlier. As a result, undesirable partitioning of the data as in the FCM can be avoided.

Figure 19 shows the recognition plots that compare both methods.

The plot for the interval type-2 method shows the best result for each value of m (i.e., best result from fixing m_1 and varying m_2, or varying m_1 and fixing m_2). From the figure, it shows that interval type-2 FCM outperforms FCM in all cases.

B. Gaussian Random Generated Data Set with Noise

In the next example, Figure 20 shows results for Gaussian random generated data in the presence of noise. For this example, uncertainty in the selection of objective function is considered. The purpose of this experiment is to observe the result of overlapping cluster prototypes, which can occur in mode-seeking algorithms such as PCM. Figure 20(a) and (b) represent clustering results for FCM and PCM, respectively, showing the final locations of the prototypes. The FCM result shows the cluster centers of left cluster and right cluster to appear slightly away from the desirable cluster centers due to the noise. Using these final cluster centers of FCM to initialize PCM, results show the final location of all three cluster prototypes are the same. In Figure 2(c), the interval type-2 fuzzy method performs somewhat as a compromise between FCM and PCM. The left cluster center and right cluster center move a little toward middle cluster center due to the cluster membership that describes cluster typicality (PCM).

C. Image Segmentation

For our final examples, we consider problems that involve images. For our first example, we apply the interval type-2 FCM method to a problem that involves a 364 × 236 image of "wolf." This image is considered to have three distinctive

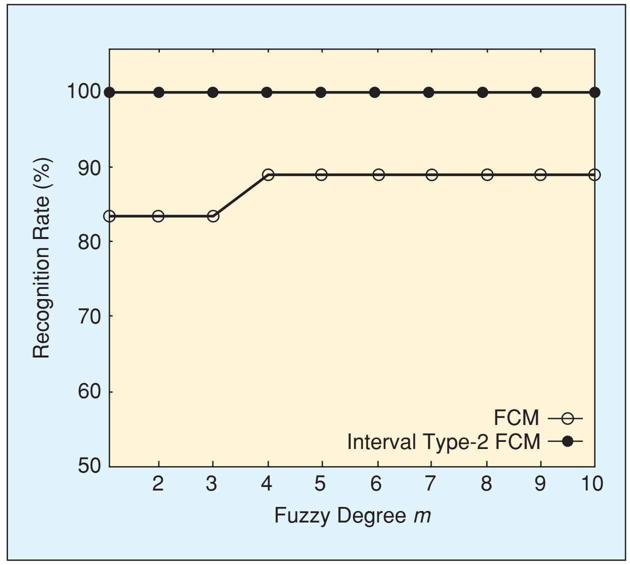


FIGURE 19 Recognition plot for FCM and interval type-2 FCM clustering results (m and $m_1 = 1.1 \sim 10.0$, $m_2 =$ value for best result).

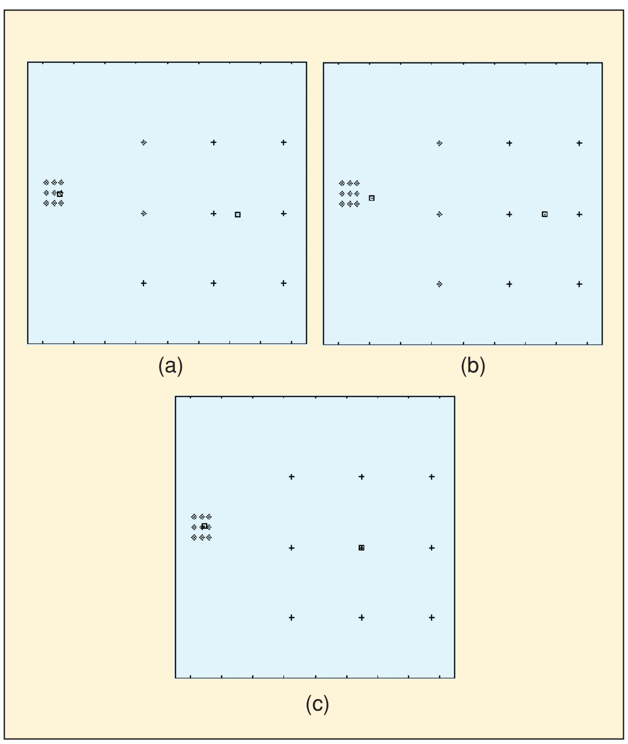


FIGURE 18 Clustering results on a pattern set consisting of two square-distributed patterns of different volume: (a) FCM with $m = 4.0$, (b) FCM $m = 2.0$, and (c) interval type-2 FCM with $m_1 = 2.0$ and $m_2 = 10.0$.

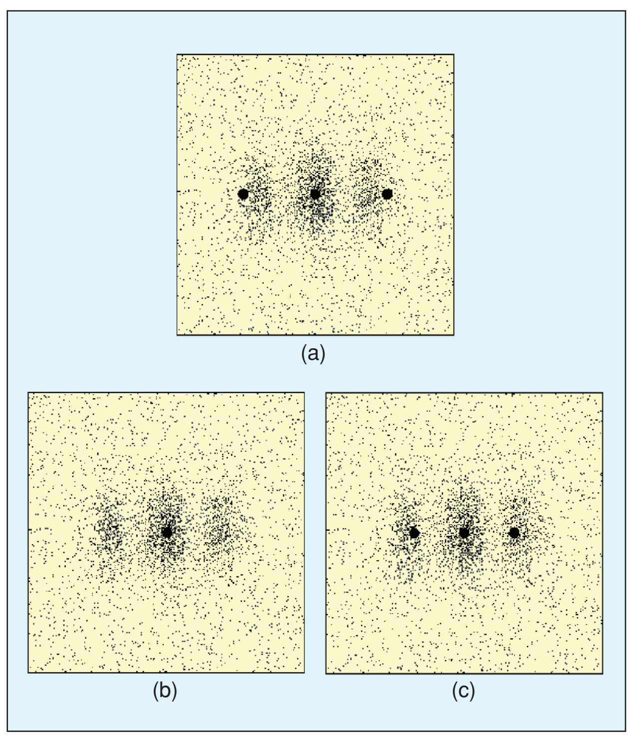


FIGURE 20 Clustering results on Gaussian random generated data: (a) FCM, (b) PCM, and (c) interval type-2 method.

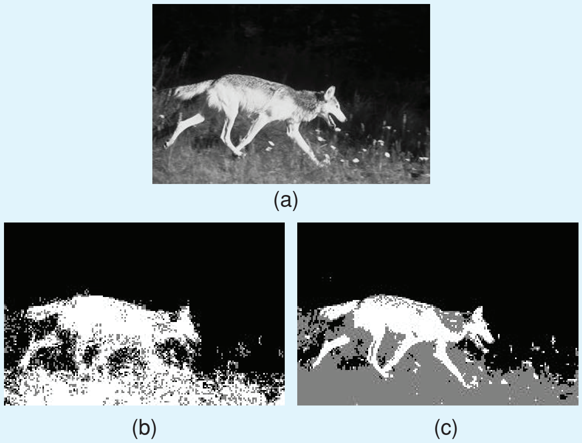


FIGURE 21 Segmented results for "Wolf" image: (a) original image, (b) FCM, and (c) interval type-2 FCM.

regions; namely, wolf, grass, and background. We extracted the color information (R,G,B) from each image and constructed 3-dimensional pattern vectors. Figure 21(b) shows the FCM result, where two of the clusters centers almost overlap; hence, the wolf and grass become almost indistinguishable. Whereas shown in Figure 21(c), the proposed method gives three distinct regions.

The next example shows a 200×200 image of "elan," which consists of a car and background. Two feature images (intensity image and histogram-equalized median filtered image) are used to extract pattern data. For each feature image, 100 samples are randomly extracted from each object (car and background). Here, the uncertainty in the selection of objective function is considered. Clustering result for $C = 2$ is shown in Figure 22, where the segmented images show the cluster partitioning of the image data for FCM, PCM, and interval type-2. Although the car was separated almost perfectly from the background, many parts of the background were grouped incorrectly for FCM. The PCM result show groups of the background included in areas of the car; however, hardly any part of the car was included in the background (see Figure 22(c)). The type-2 result in Figure 22(d) is shown to give slightly more desirable results compared with Figure 22 (b) and (c).

VI. Conclusions

In this article, interval type-2 fuzzy sets were used to model the uncertainty that is associated with the various parameters in objective function-based clustering. The purpose was to represent and manage the uncertainty in the cluster memberships by incorporating interval type-2 fuzzy sets. As a result, interval type-2 clustering methods were obtained by modifying the prototype-updating and hard-partitioning procedures in the type-1 fuzzy objective function-based clustering. As a consequence, the management of uncertainty by an interval type-2 fuzzy approach aids cluster prototypes to converge to a more desirable location than a type-1 fuzzy approach. Several examples illustrated the effectiveness of interval type-2 fuzzy approach methods. Furthermore, the uncertainty associated with the parameters (e.g., in Table 1) for other existing clustering algorithms can be considered in the development of several other interval type-2 clustering algorithms. They are currently under investigation.

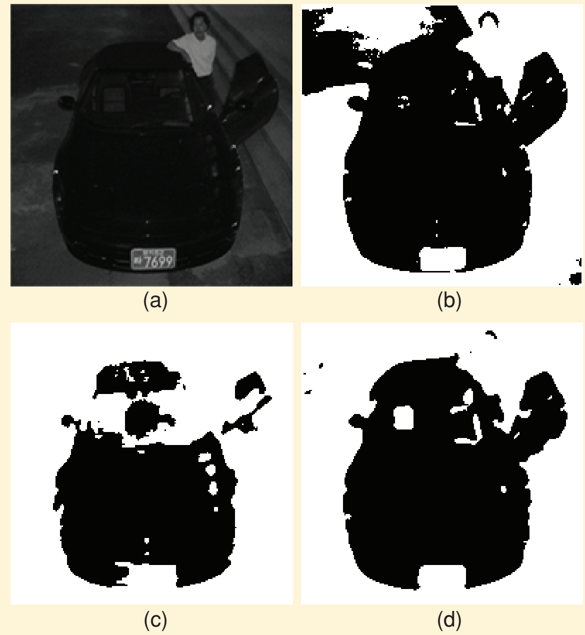


FIGURE 22 Segmentation results for "elan" image: (a) original image, (b) FCM, (c) PCM, and (d) interval type-2 FCM.

Acknowledgment

This research was supported by the Agency for Defense Development, Korea, through the Image Information Research Center at Korea Advanced Institute of Science & Technology. The author would also like to thank his former Ph.D. student Dr. Cheul Hwang for useful suggestions for the improvement of this article and computer simulations.

References

[1] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, 1981.

[2] R. Krishnapuram and J. Keller, "A possible approach to clustering," *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 2, pp. 98–110, May 1993.

[3] J. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice Hall, Upper Saddle River, NJ, 2001.

[4] J. Mendel and R. John, "Type-2 fuzzy set made simple," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, Apr. 2002.

[5] F. Rhee and C. Hwang, "A type-2 fuzzy *C*-means clustering algorithm," in *Proc. 2001 Joint Conf. IFSA/NAFIPS*, pp. 1926–1919, July 2001.

[6] F. Rhee and C. Hwang, "An interval type-2 fuzzy perceptron," in *Proc. 2002 Int. Conf. Fuzzy Syst.*, vol. 2, pp. 1331–1335, May 2002.

[7] F. Rhee and C. Hwang, "An interval type-2 fuzzy K-nearest neighbor," in *Proc. 2003 Int. Conf. Fuzzy Syst.*, vol. 2, pp. 802–807, May 2003.

[8] C. Hwang and F. Rhee, "An interval type-2 fuzzy *C* spherical shells algorithm," in *Proc. 2004 Int. Conf. Fuzzy Syst.*, vol. 2, pp. 1117–1122, July 2004.

[9] C. Hwang and F. Rhee, "Uncertain fuzzy clustering: interval type-2 fuzzy approach to *C*-means," to appear in *IEEE Transactions on Fuzzy Systems*.