

## Interval Type-2 Possibilistic Fuzzy C-means 클러스터링을 위한 퍼지화 상수 결정 방법

Determining the Fuzzifier Values for Interval Type-2 Possibilistic Fuzzy C-means Clustering

---

저자 (Authors)	주원희, 이정훈 Won-Hee Joo, Frank Chung-Hoon Rhee
출처 (Source)	<a href="#">한국지능시스템학회 논문지 27(2)</a> , 2017.4, 99-105 (7 pages) <a href="#">Journal of Korean Institute of Intelligent Systems 27(2)</a> , 2017.4, 99-105 (7 pages)
발행처 (Publisher)	<a href="#">한국지능시스템학회</a> Korean Institute of Intelligent Systems
URL	<a href="http://www.dbpia.co.kr/Article/NODE07157714">http://www.dbpia.co.kr/Article/NODE07157714</a>
APA Style	주원희, 이정훈 (2017). Interval Type-2 Possibilistic Fuzzy C-means 클러스터링을 위한 퍼지화 상수 결정 방법. 한국지능시스템학회 논문지, 27(2), 99-105.
이용정보 (Accessed)	신라대학교 61.100.225.*** 2019/03/20 01:16 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.



## Interval Type-2 Possibilistic Fuzzy C-means 클러스터링을 위한 퍼지화 상수 결정 방법

### Determining the Fuzzifier Values for Interval Type-2 Possibilistic Fuzzy C-means Clustering

주원희\* · 이정훈\*\*†

Won-Hee Joo and Frank Chung-Hoon Rhee†

\*한양대학교 전자통신공학부, \*\*한양대학교 전자공학부

\*Department of Electronic and Communication Engineering, Hanyang University

\*\*School of Electronic Engineering, Hanyang University

#### 요약

일반적으로 type-1 fuzzy set에 존재하는 불확실성을 보다 효율적으로 다루고 제어하기 위하여 Type-2 fuzzy set (T2 FS)이 널리 사용되고 있다. T2 FS에서 퍼지화 상수 (fuzzifier value)  $m$ 은 이러한 불확실성을 처리하기 위한 가장 중요한 요소이다. 따라서 적절한 퍼지화 상수 값을 결정하는 연구는 여전히 지속되고 있고, 많은 방법들이 연구 되어 왔다. 본 논문에서는 주어진 패턴을 분류하기 위하여 Interval type-2 possibilistic fuzzy C-means (IT2 PFCM) 클러스터링 방법을 사용한다. 클러스터링을 위해 사용된 IT2 PFCM 방법에서 각 데이터에 대하여 적응적으로 적절한 퍼지화 상수의 값을 계산하는 방법을 제안한다. 히스토그램 접근법을 통하여 각각의 데이터 포인트로부터 정보를 추출해 내고 추출된 정보를 이용하여 두 개의 퍼지화 상수인  $m_1, m_2$  값을 결정한다. 이렇게 얻어진 값은 interval type-2 fuzzy의 최저 및 최고 멤버십 값을 결정하게 된다.

키워드 : Fuzzy sets, Interval type-2 fuzzy sets, Interval type-2 PFCM, 퍼지화 상수, 히스토그램 방법

#### Abstract

Type-2 fuzzy sets are preferred over type-1 sets as they are capable of addressing uncertainty more efficiently. The fuzzifier values play pivotal role in managing these uncertainties; still selecting appropriate value of fuzzifiers has been a tedious task. Generally, based on observation particular value of fuzzifier is chosen from a given range of values. In this paper we have tried to adaptively compute suitable fuzzifier values of interval type-2 possibilistic fuzzy c-means (IT2 PFCM) for a given data. Information is extracted from individual data points using histogram approach and this information is further processed to give us the two fuzzifier values  $m_1$  and  $m_2$ . These obtained values are bounded within some upper and lower bounds based on interval type-2 fuzzy sets.

Key Words: Fuzzy sets, Interval type-2 fuzzy sets, Interval type-2 PFCM, Fuzzifier, Histogram methods

Received: Mar. 4, 2017

Revised: Apr. 12, 2017

Accepted: Apr. 13, 2017

†Corresponding authors

frhee@fuzzy.hanyang.ac.kr

## 1. 서론

일반적으로 패턴을 분류하는 퍼지 클러스터링에 대한 많은 연구가 진행 되어 왔고 fuzzy C-means (FCM) 알고리즘이 가장 많이 사용 되어 왔다[1][2]. FCM 방법의 경우 상대적 유클리디안 거리를 이용하기 때문에 노이즈가 있는 데이터에서 좋지 않은 결과를 나타낸다. 또한, 모든 클래스의 멤버십 (Membership) 합의 값이 항상 1이 되어야 한다는 조건 때문에 잘못된 클러스터링을 할 수도 있다. 이러한 FCM 방법의 문제점을 해결하기 위해 하나의 패턴과 하나의 중심 값 간의 절대 거리로 typicality를 할당하는 possibilistic C-means (PCM) 클러스터링 알고리즘이 제안 되었다[3].

그러나 PCM 알고리즘 또한 초기 파라미터 값에 따라 클러스터링 결과가 민감하게 반응 한다는 문제점이 있다. 이러한 문제점을 해결하기 위해 FCM 과 PCM을 가중치 합으로 결합한 PFCM 알고리즘이 연구 되었다[4]. 그러나 PFCM 방법 또한 퍼지화 상수 값을 결정해야 하는 불확실한 문제를 가지고 있다. 퍼지화 상수 값은 멤버십 함수를 구하는데 있어 결정적 작용을 하기 때문에 퍼지화 상수 값의 불확실성을 제어 하는 것은 중요한 문제이다. 일반적으로 사용자에게 의해 퍼지화 상수가  $m=2$  값으로 선택 되어 사용되고 있으며, 여러 번의 반복된 실험을 통하여 퍼지화 상수의

이 논문은 2015년 한양대학교 교내 연구비 지원으로 연구되었음 (H-Y-2015-p)

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

최저 및 최고 경계 (upper and lower bound)에 따른 최적 범위를 측정하는 연구가 진행 되고 있다[5][6]. 이러한 연구들이 진행되고 있지만, 각각 다른 모든 데이터에 동일한 퍼지화 상수 범위를 적용 할 수 없다[7]. 따라서 다르게 주어지는 여러 종류의 데이터들에 대하여 적응적으로 퍼지화 상수를 결정하는 방법이 연구 되어야 한다. 따라서 본 논문에서는 Interval type-2 possibilistic Fuzzy C-means (IT2 PFCM) 클러스터링 방법을 기반으로 하여 히스토그램을 이용한 퍼지화 상수 결정 방법을 제안하였다.

## 2. 배경이론

### 2.1 퍼지화 상수

주어진 각 클러스터의 밀도 혹은 부피가 다를 경우 퍼지화 상수는 클러스터링 멤버십 함수를 구하는 데 있어 결정적인 역할을 하게 된다. 두 개의 클러스터  $c_1, c_2$ 의 중심  $v_1$ 과  $v_2$  그리고 수직선 (vertical line)에 위치한 점들을 가정하였을 때 그림 1(a)과 같이 클러스터 센터에 대한 상대적 거리 (relative distance)는 0.5로 모두 같다. 이는 퍼지화 상수가  $m \rightarrow 1$ 인 경우로 결정 경계 (decision boundary)로 간주 된다. 위 조건에서의 퍼지화 영역은 존재 하지 않는다. 그림 1(b) 1(c)와 같이 퍼지화 상수  $m$  (degree of fuzziness) 값의 변화에 따라서 최대, 최소 퍼지 영역 또한 변화하게 된다. 따라서 퍼지화 상수  $m$ 의 선택에 따라 멤버십 함수가 다르게 계산되어 지고, 멤버십 함수에 의해 패턴이 속하는 소속도가 달라 질 수 있다. 또한 그림 1(d)와 같이 interval type-2 fuzzy set (IT2 FS)로 확장하여 두 개의 퍼지화 상수  $m_1, m_2$ 를 사용 하였을 경우 1개의 퍼지화 상수를 사용 하였을 때 보다 더 정확하게 패턴을 분류 할 수 있다[8].

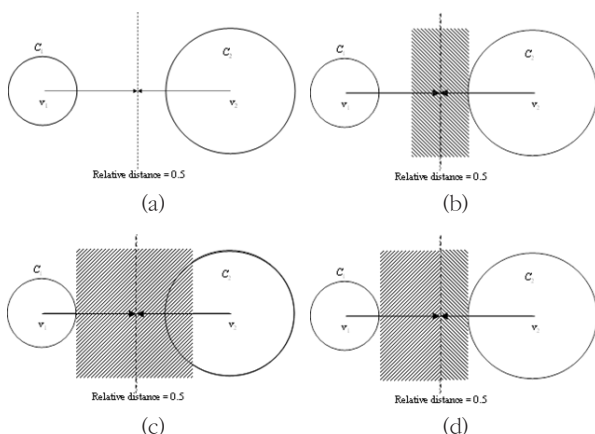


그림 1. 퍼지화 상수에 따른 퍼지화 영역: (a)  $m=1.1$  일 때 퍼지화 영역,

(b)  $m_1$ 이 작을 때, (c)  $m_2$ 가 커질 때, (d)  $m_1, m_2$ 를 적용할 때.

Fig. 1. Degree of fuzziness for  $m$ : (a) No fuzzy region  $m=1.1$ , (b) low degree, (c) high degree, and (d) using both  $m_1$  and  $m_2$ .

### 2.2 퍼지화 상수의 범위 결정

앞에서 나는 바와 같이 주어진 데이터로부터 퍼지화 상수의 최저 및 최고 경계 범위 값을 결정하는 몇 가지 방법들이 제안 되었다[9]. 퍼지화 상수의 범위를 결정하기 위한 PFCM 멤버십 함수는 다음과 같이 주어진다. 클러스터  $i$ 에 대한  $k$ 번째 데이터 포인트에 대한 멤버십 함수는 식(1)과 같다.  $d_{ik}/d_{ij}$ 는 클러스터와 데이터 포인트 간의 유클리디안 거리 값을 의미한다.

$$u_{ik} = \frac{1}{\sum_{j=1}^c (d_{ik}/d_{ij})^{2/(m-1)}}. \quad (1)$$

퍼지화 상수의 범위를 결정하기 위하여 (1)에서 얻어진 멤버십 값을 이용하여 이웃의 멤버십 값들을 계산한다. 이를  $m$ 에 관한 표현식으로 정리하면 식(2)와 같고 퍼지화 상수의 최저 및 최고 경계 값을 얻을 수 있다.  $C$ 는 클러스터의 개수이고,  $m$ 은 퍼지화 상수이다.

$$1 + \frac{C-1}{C} \cdot \frac{2}{\delta} \cdot |\Delta| \leq m \leq \frac{2 \log d}{\log \left( \frac{\delta}{1-\delta} \cdot \frac{1}{c-1} \right)} + 1, \quad (2)$$

$$\text{where } \Delta = \frac{d_i - d_i^*}{d_i^*} \text{ and } \delta \text{ is threshold.}$$

### 2.3 Interval type-2 fuzzy membership function

일반적으로 type-1 fuzzy set (T1 FS)은 패턴 인식 분야에서 패턴의 불확실성을 나타내는데 널리 사용 되어져 왔다. 하지만 앞에서 본 바와 같이 T1 FS로 좋은 결과를 나타낼 수 없고, 불확실한 퍼지화 상수를 더욱 효율적으로 제어하기 위하여 type-2 fuzzy set (T2 FS)으로 확장한다[10][11][12]. T2 FS  $\tilde{A}$ 를 아래와 같이 나타낸다.

$$\tilde{A} = \int_{x \in X} u_{\tilde{A}}(x)/x = \int_{x \in X} \left[ \int_{u \in J_x} f_x(u)/u \right] / x. \quad (3)$$

불확실성을 더 강력하게 제어하는 T2 FS으로의 확장은 T1 FS에 비하여 우수한 결과를 얻어낼 수 있다. 그러나 계산이 복잡하고 많은 연산량을 필요로 한다[9]. 이를 보완하기 위하여 secondary membership degree가 1인 IT2 FS를 사용한다. IT2 FS  $\tilde{A}$ 를 식(4)와 같이 나타낸다. 식에서 알 수 있듯이 secondary membership degree가 모두 같기 때문에 T1 FS처럼 사용할 수 있다.

$$\tilde{A} = \left[ \int_{u \in J_x} 1/u \right] / x \quad (4)$$

### 3. Interval type-2 PFCM 알고리즘

IT2 PFCM 은 FCM 방법과 PCM 방법의 가중치의 합으로 표현되며, 위의 두 가지 특성을 모두 지니고 있다. 따라서 다음과 같이 PFCM 목적함수(objective function)를 최소화 하는 방향으로 클러스터링 한다.

$$J_{m,\eta}(U, T; V; X) = \sum_{k=1}^n \sum_{i=1}^c (au_{ik}^m + bt_{ik}^\eta) \times \|x_k - v_i\|^2 + \sum_{i=1}^c \gamma_i \sum_{k=1}^n ((1-t_{ik})^\eta) \quad (5)$$

$$\sum_{i=1}^c u_{ik} = 1, 0 \leq u_{ik}, t_{ik} \leq 1, m > 1, \eta > 1, \gamma_i > 0 \quad (6)$$

식(5)에서,  $u_{ik}$ 는 입력 패턴  $k$ 가 클러스터  $i$ 에 소속되는 멤버십 값을 나타낸다.  $x_k$ 는  $k$ 번째 입력 패턴이고,  $v_i$ 는  $i$ 번째 클러스터의 중심 값을 나타낸다.  $m$ 은 퍼지화의 정도를 나타내는 상수로서  $m \in (1, \infty)$ 의 조건을 만족시킨다.  $t_{ik}$ 는 입력 패턴  $k$ 가 클러스터  $i$ 에 소속되는 typicality를 나타내는데 이는 절대적 거리를 이용하는 PFCM 방법의 특징이다.  $\gamma_i$ 는  $i$ 번째 클러스터의 typicality가 0.5가 되는 지점을 정의하는 스케일로서 값은 임의의 수이다. PFCM 방법으로 클러스터링을 하기 위해 위의 식(5)의 목적함수를 멤버십 함수  $u_{ik}$ 에 관하여 최소화해야 한다. 이를 위한 멤버십은 식(1)을 이용하여 구하고  $a$ 와  $b$ 는 fcm 과 pcm 의 가중치를 결정하는 변수이다.

앞서 2.3장에서 IT2 FS로 확장하기 위해서는 퍼지화 상수  $m$ 에 대한 불확실성을 표현해야 한다. 이를 위해서 primary membership 함수를 사용하여 최저 및 최고 멤버십 함수를 만들어야 한다. 퍼지화 상수  $m$ 에 따른 PFCM의 최저 및 최고 멤버십 함수는

$$\bar{u}_{ik} = \left( \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{ij}} \right)^{\frac{2}{m_1-1}}} \right) \quad (7)$$

$$u_{ik} = \left( \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{ij}} \right)^{\frac{2}{m_2-1}}} \right) \quad (8)$$

식(7),(8)의  $m_1, m_2$ 는 최고 및 최저 퍼지화 상수이다.

목적함수를 나타내는 식(5)에서 보듯이 최저 및 최고 멤버십 함수에 따라서  $\gamma_i$ 의 값 또한 변하게 된다. 식(7)(8)을 최소화하기 위한  $\gamma_i$ 는

$$\bar{\gamma}_i = K \frac{\sum_{k=1}^n u_{ik}^{-m_1} d_{ik}^2}{\sum_{k=1}^n u_{ik}^{-m_1}} \quad (9)$$

$$\underline{\gamma}_i = K \frac{\sum_{k=1}^n u_{ik}^{m_2} d_{ik}^2}{\sum_{k=1}^n u_{ik}^{m_2}}, K > 0. \quad (10)$$

$\gamma_i$ 를 사용하여 최저 및 최고 typicality는

$$\bar{t}_{ik} = \begin{cases} \frac{1}{1 + \left( \frac{d_{ik}^2}{\bar{\gamma}_i} \right)^{\frac{2}{m_1-1}}}, & \text{if } \frac{1}{1 + \left( \frac{d_{ik}^2}{\bar{\gamma}_i} \right)^{\frac{2}{m_1-1}}} > \frac{1}{1 + \left( \frac{d_{ik}^2}{\underline{\gamma}_i} \right)^{\frac{2}{m_2-1}}} \\ \frac{1}{1 + \left( \frac{d_{ik}^2}{\underline{\gamma}_i} \right)^{\frac{2}{m_2-1}}}, & \text{otherwise} \end{cases} \quad (11)$$

$$\underline{t}_{ik} = \begin{cases} \frac{1}{1 + \left( \frac{d_{ik}^2}{\underline{\gamma}_i} \right)^{\frac{2}{m_1-1}}}, & \text{if } \frac{1}{1 + \left( \frac{d_{ik}^2}{\underline{\gamma}_i} \right)^{\frac{2}{m_1-1}}} \leq \frac{1}{1 + \left( \frac{d_{ik}^2}{\bar{\gamma}_i} \right)^{\frac{2}{m_2-1}}} \\ \frac{1}{1 + \left( \frac{d_{ik}^2}{\bar{\gamma}_i} \right)^{\frac{2}{m_2-1}}}, & \text{otherwise} \end{cases} \quad (12)$$

위와 같이 멤버십을 구하고 난 후 각 클러스터들의 중심 값을 갱신해야 한다. 중심 값 갱신을 위해 KM 알고리즘을 이용하여 type-2 fuzzy set을 type-1으로 바꾸는 type reduction 과정을 수행 하여 각 클러스터들의 갱신한 중심 값은 식(13)과 같다.

$$v_i = \frac{\sum_{k=1}^n (au_{ik}^m + bt_{ik}^\eta) X_k}{\sum_{k=1}^n (au_{ik}^m + bt_{ik}^\eta)} \quad (13)$$

### 4. 히스토그램을 응용한 퍼지화 상수 결정

본 논문에서 제안하는 방법은 히스토그램 방법을 통하여 주어지는 데이터들에서 정보를 추출 한 후, 얻어진 정보를 바탕으로 적응적으로 퍼지화 상수를 계산한다.

먼저 3장에서 정의한 IT2 PFCM 알고리즘을 통하여 데이터가 어느 클러스터에 속하는지 대략적으로 추정한다. 대략적으로 추정하여 분류된 클러스터들의 데이터를 기반으로 하여 히스토그램을 얻어낸다. 이렇게 얻어진 히스토그램은 triangular

window를 통해 보다 완만하고 부드러운 히스토그램으로 만들고, 이 히스토그램 위에 곡선적합방법을 사용하여 멤버십 함수를 얻어낸다. IT2 FS를 얻기 위하여 T2 FS의 모든 주요 멤버십 집합인 Footprint of uncertainty (FOU)를 결정해야 한다. 그러므로 멤버십 값보다 큰 히스토그램의 값들은 최고 멤버십의 히스토그램으로 할당 되고, 멤버십 값보다 작은 값을 가지는 히스토그램의 값들은 최저 멤버십의 히스토그램으로 저장된다. 이와 같이 곡선적합된 히스토그램을 이용하여 최저 및 최고 멤버십 함수를 구할 수 있다. 이렇게 구해진 멤버십 함수를 이용하여 퍼지화 상수  $m_1, m_2$ 를 추정하는 알고리즘을 제안한다. 그림 2는 클래스와 차원에 따라 결정된 히스토그램과 FOU 예를 보여주고 있다.

함수  $f(X)=0$ 을 만족하는  $X$ 를 구하기 위해 고정 소수점 반복법을 이용하여  $X=g(X)$  형태로 나타낼 수 있고 다음  $X$ 은

$$X_{i+1} = g(X_i), \quad i = 0, 1, \dots, N \quad (14)$$

식(7)과 (8)을 (14)식의 형태로 표현하면

$$u_i = \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{ji}} \right)^{\frac{2}{m-1}}} \quad (15)$$

$$\Rightarrow \frac{1}{u_i} = \sum_{j=1}^c \left( \frac{d_{ik}}{d_{ji}} \right)^{\frac{2}{(m-1)}}. \quad (16)$$

식(16)에서 양변에 모두  $\log$ 를 취하면

$$\log \frac{1}{u_1} = \log \left( \sum_{j=1}^c \frac{d_{ik}}{d_{ji}} \right)^{\frac{2}{m-1}} \quad (17)$$

$$\therefore \log \sum_{k=0}^N a_k = \log a_0 + \log \left( 1 + \sum_{i=1}^N \frac{a_k}{a_0} \right). \quad (18)$$

식(17)을 아래와 같이 정리하면

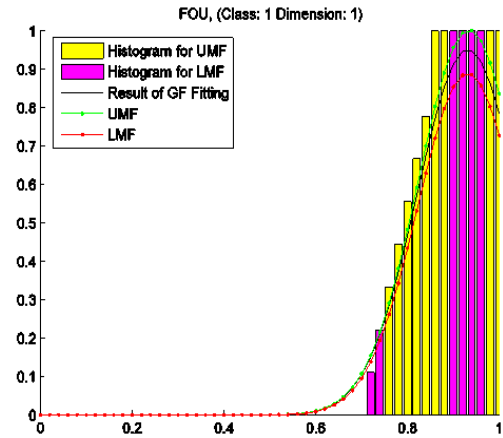
$$\begin{aligned} \log \left( \frac{1}{u_1} \right) &= \frac{2}{m-1} \log \left( \frac{d_{ki}}{d_{1i}} \right) \\ &+ \log \left( 1 + \sum_{j=2}^c \left( \frac{d_{ki}}{d_{ji}} \right)^{\frac{2}{m_{old}-1}} \right). \end{aligned} \quad (19)$$

식(19)로부터 본 논문에서 구해야 할 퍼지화 상수  $m$ 에 관하여 정리하면 다음과 같이 새로운  $m$ 의 값은

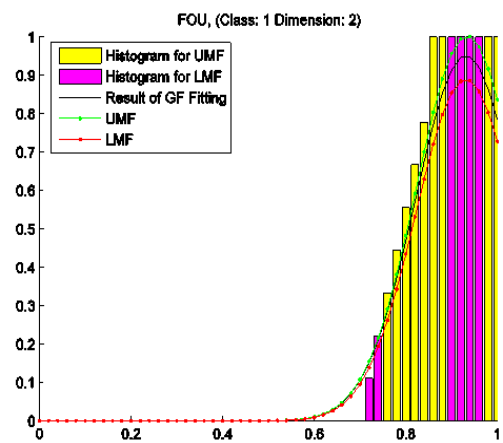
$$m_{i \text{ new}} = \frac{\log \left( \frac{1}{u_1} \right) - \log \left( 1 + \sum_{j=2}^c \left( \frac{d_{ki}}{d_{ji}} \right)^{\frac{2}{m_{old}-1}} \right)}{\log \left( \frac{d_{ki}}{d_{1i}} \right)}. \quad (20)$$

위의 과정과 같이  $u_i \in \{u_i(X_k), \bar{u}_i(X_k)\}$ 와  $m_{i \text{ new}}$ 의 값은 멤버십 값을 얻기 위한 함수로 쓰이게 된다. 식(12)를 모든 클러스터 데이터에 적용하여 계산하면 새로운  $m_{1 \text{ new}}$ 과  $m_{2 \text{ new}}$  값을 구할 수 있다. 식(20)을 통해 구해진 퍼지화 상수값의 평균값을 취하여 최종적으로 새로운 퍼지화 상수  $m_1$ 과  $m_2$ 를 결정하게 되고, 클러스터링 위해 이렇게 구해진 새로운 퍼지화 상수는

$$m_j = \left( \sum_{i=1}^N m_{ji} \right) / N \quad j = 1, 2. \quad (21)$$



(a)



(b)

그림 2. 각 클래스와 차원에 따라 결정된 히스토그램과 FOU  
(a) 클래스1과 차원1 (b) 클래스1과 차원2

Fig. 2. FOU obtained for individual class and dimension  
(a) class 1 dimension 1, and (b) class 1 dimension 2.

## 5. 결과 및 고찰

실험의 결과 분석을 위해 기존의 방법을 따른 IT2 PFCM 과 본 논문에서 제안된 방법을 비교 진행 하였다. 다양한 패턴과 이미지를 클러스터링 하고 그 결과를 도출 해 낸다. 기존의 IT2 PFCM 방법은 여러 번 반복된 학습을 통하여 가장 최적의 퍼지화 상수  $m_1$ ,  $m_2$ 를 취사선택 하였다. 이를 토대로  $m$ 의 값을 1~5 안에서 적용 하였다. 다양한 패턴들에 대하여 퍼지화 상수 이외의 모든 변수들은 동일하게 적용 하였다.

### 5.1 Iris34 데이터

본 논문 실험의 데이터로는 널리 알려진 Iris 데이터를 사용한다. Iris 데이터는 총 3개의 클러스터로 나누어져 있으며 각 클러스터에는 50개씩 총 150개의 패턴 데이터와, 각각의 벡터에 4개의 피쳐로 구성되어 있다. 본 실험에서는 4개의 피쳐 중 피쳐3과, 피쳐4를 이용한 Iris34 데이터를 사용하여 클러스터링 하였다.

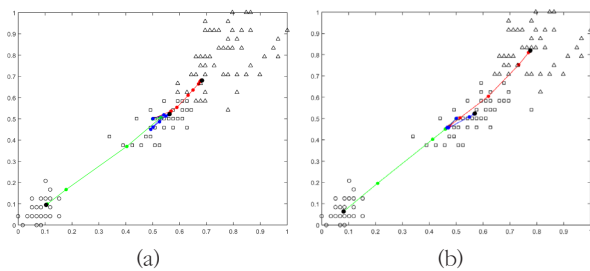


그림 3. Iris34 데이터에 대한 클러스터링 결과:  
(a) IT2 PFCM (b) 제안 방법.

Fig. 3. Clustering results of Iris34 data:  
(a) IT2 PFCM (b) proposed method.

그림 3은 Iris34 데이터의 실험의 결과를 나타낸다. 총 3개의 클러스터로 나누어져 있으며 각 클러스터에는 50개씩 총 150개의 데이터를 클러스터링 하였다. 3(a)에서는 기존의 반복 학습을 통해 퍼지화 상수  $m_1=1$ ,  $m_2=2$  찾았고 3(b)에서는 제안된 알고리즘을

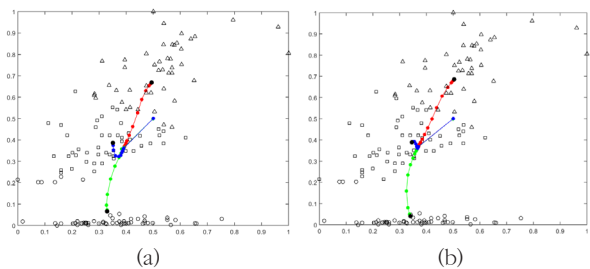


그림 4. Iris56 데이터에 대한 클러스터링 결과:  
(a) IT2 PFCM (b) 제안 방법.

Fig. 4. Clustering results of Iris56 data:  
(a) IT2 PFCM (b) proposed method.

통하여 적응적으로  $m_1=0.8$ ,  $m_2=2$ 의 값을 찾아내었다. 클러스터링 결과 역시 더 뛰어난 것을 확인 할 수 있다.

### 5.2 Iris56 데이터

그림 4는 Iris56 데이터의 실험의 결과를 나타낸다. Iris56 데이터는 Iris 데이터의 피쳐 1과, 피쳐 2를 이용하였다. 실험 방법은 5.1과 동일하다. 그림 4(a)에서 역시 기존의 반복 학습을 통해 가장 적합한 퍼지화 상수  $m_1=2$ ,  $m_2=3$  찾았고 그림 4(b)에서는 제안된 알고리즘을 통하여 적응적으로  $m_1=2.1$ ,  $m_2=2.8$ 의 값을 찾아내었다. 그림 4와 같이 클러스터링 결과도 더 뛰어난 것을 확인 할 수 있다.

### 5.3 이미지 데이터1 (길, 숲, 하늘)

길, 숲, 하늘 영상을 분할한다. 실험 영상은 길, 숲, 하늘 세 부분으로 이루어져 있으며 이미 만들어진 intensity 와 가우시안 속성의 영상을 이용한다. 두 영상에서 100개의 샘플 데이터를 추출하여 학습 시킨 후 학습된 결과를 이용해 영상을 분할한다(그림 5(6)와 표 1(2) 참조).

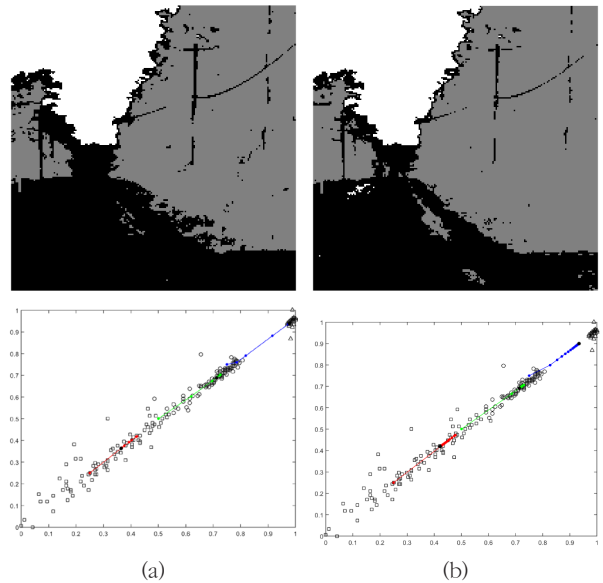


그림 5. 길 숲 영상에 대한 영상 분할 결과 (a) IT2 PFCM (b) 제안 방법.  
Fig. 5. Clustering results of image segmentation 1: (a) IT2 PFCM and (b) proposed method.

표 1. 길, 숲, 하늘 영상에 대한 인식률 및 iteration 비교.

Table 1. Comparison on the basis of classification rate and number of iterations for image 1.

Method	IT2 PFCM	Proposed
Recognition rate(%)	81.57	84.65
Iterations	16	35



## 5.4 이미지 데이터2(길, 숲, 하늘)

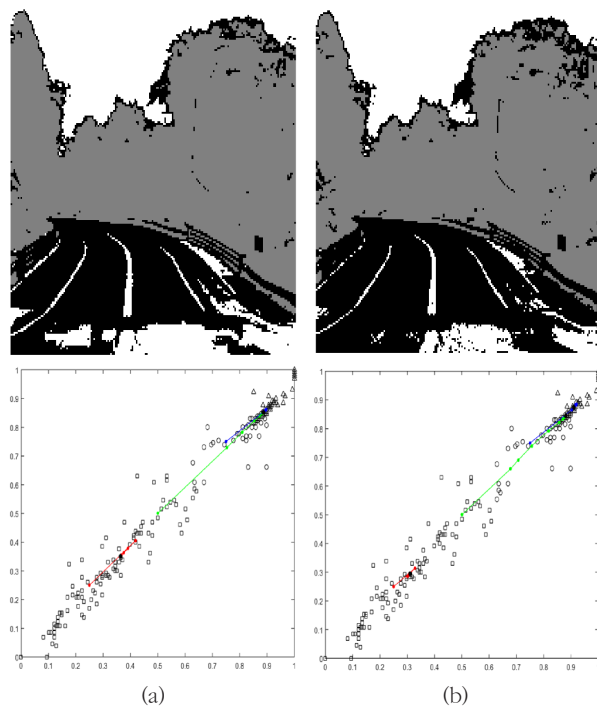


그림 6. 길 숲 영상2에 대한 영상 분할 결과: (a) IT2 PFCM (b) 제안 방법.

Fig. 6. Clustering results of image segmentation 2:

(a) IT2 PFCM and (b) proposed method.

표 2. 길, 숲, 하늘 영상에 대한 인식률 및 iteration 비교.

Table 2. Comparison on the basis of classification rate and number of iterations for image 2.

Method	IT2 PFCM	Proposed
Recognition rate(%)	80.46	81.65
Iterations	32	81

## 6. 결론 및 향후 연구

Fuzzy set을 이용한 클러스터링을 수행 하려면 선행적 작업이 필요하고, 반복학습을 통해 적절한 값을 찾아야 했다. 더욱이 T2 FS 으로 확장되면서 불확실성을 내포한 퍼지화 상수의 결정은 무엇보다 중요한 요소였다. 본 논문에서 제안한 알고리즘으로 인하여 반복적 학습을 수행 할 필요 없이 어떠한 data 대하여서도 적응적으로 최적화 된 퍼지화 상수를 구할 수 있음을 실험을 통해 확인 하였다. 앞으로 general type-2 fuzzy set 에 근거하여 최적화된 퍼지화 상수를 찾아내는 연구로 확장 할 것이다.

## References

- [1] G. Raju, B. Thomas, S. Tobgay, and S. Kumar, "Fuzzy clustering methods in data mining: a comparative case analysis," *ICACTE*, pp. 489-493, 2008.
- [2] J. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, New York: Plenum, 1981.
- [3] R. Krishnapuram and J. Keller, "A possibilistic approach to clustering," *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 2, pp. 98-110, May 1993.
- [4] N. Pal and J. Bezdek, "A possibilistic fuzzy c-means clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 13, pp. 517 - 530, Aug. 2005.
- [5] Z. Wa, W. Xie, and J. Yu, "Fuzzy c-means clustering algorithm based on kernel method," *Int. Conf. on Computational Intelligence and Multimedia Applications*, pp. 49-54, Sep. 2003.
- [6] J. Bezdek, R. Hathaway, M. Sabin, and W. Tucker, "Convergence theory for fuzzy c-means: counter examples and repairs," *IEEE Trans. Systems, Man and Cybernetics*, vol. 17, pp. 873-877, Oct. 1987.
- [7] J. Yu, Q. Cheng, and H. Huang, "Analysis of the weighting exponent in the FCM," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 34, pp. 634-639, Feb. 2004.
- [8] Y. Cho, H. Lee, and H. Jeon, "Interval Type-2 Fuzzy Logic Control System of Flight Longitudinal Motion," *Journal of Korean Institute of Intelligent Systems*, vol. 25, no.2, pp. 168-173, 2015.
- [9] I. Ozkan and I. Turksen, "Upper and lower values for the level of fuzziness in FCM," *Information Sciences*, vol. 177, pp. 5143-5152, Dec. 2007.
- [10] J. Min and F. Rhee, "An Interval Type-2 Fuzzy PCM Algorithm for Pattern Recognition," *Journal of Korean Institute of Intelligent Systems*, vol. 19, no.1, pp. 102-107, 2009.
- [11] C. Hwang and F. Rhee, "Uncertain fuzzy clustering: interval type-2 fuzzy approach to C-means," *IEEE Trans. Fuzzy Syst.*, vol. 15, no.1, pp. 107-120, Feb. 2007.
- [12] F. Rhee, "Uncertain fuzzy clustering: insights and recommendations," *IEEE Computational Intelligence Magazine*, vol. 2, no. 1, Feb. 2007.

## 저자 소개



### 주원희(Won-Hee Joo)

2008년 : 숭실대학교 정보통신공학부 공학사  
2010년 : 숭실대학교 대학원 정보통신공학부  
공학석사  
2010년~현재 : 한양대학교 대학원  
전자통신공학과 박사과정

관심분야 : Fuzzy Systems, Pattern Recognition

E-mail : jinilove@hanyang.ac.kr



### 이정훈(Frank Chung-Hoon Rhee)

1985년 : University of Southern California(USC),  
Dept. of Electrical Eng., BSEE  
1989년 : University of Missouri,  
Dept. of Electrical & Computer Eng.,  
MSEE

1990년~1993년 : University of Missouri,  
Dept. of Electrical & Computer Eng., Ph.D  
1995년~현재 : 한양대학교 전자공학부 교수  
2015년~현재 : 한국지능시스템학회 부회장,  
세계 퍼지학회 (IFSA) 한국지부 대표

관심분야 : Computer Vision, Pattern  
Recognition, and all aspects of Computational  
Intelligence

E-mail : frhee@fuzzy.hanyang.ac.kr