Journal of The Korea Society of Computer and Information
Vol. 19, No. 7, July 2014
www.ksci.re.kr
http://dx.doi.org/10.9708/jksci.2014.19.7.001

분류 오류 최소화를 위한 클러스터링 기법

허 경 용*, 김 성 훈**

A New Clustering Method for Minimum Classification Error

Gyeong-Yong Heo*, Seong-Hoon Kim**

요 약

클러스터링은 대표적인 비교사 학습 방법의 하나로 균일한 특성을 가지는 데이터를 군집으로 묶기 위해 사용된다. 균일한 특성을 가지는 데이터 부분집합을 문맥으로 정의하고 문맥 내에서 국부적으로 분류를 행하는 융합 방법이 사용되고 있지만 클러스터링은 비교사 학습 방법이라는 한계로 인해 클러스터링 결과로 만들어지는 문맥이 분류에 있어 최선임을 보장하기 어렵다.

이 논문에서는 생성된 클러스터를 문맥으로 가정하고 각 문맥에서 분류를 시행하는 경우 최소의 오류를 보일 수 있는, 분류를 고려한 클러스터링 기법을 제안한다. 제안하는 방법은 선형 판별 분석에서와 유사하게 클러스터 내 동일한 클래스에 속하는 데이터 쌍은 작은 거리 값을, 서로 다른 클래스에 속하는 데이터 쌍은 큰 거리 값을 가지도록 하기 위한 제약 조건을 적용하여 분류 오류를 줄이도록 하였다. 제안한 방법의 실효성은 실험 결과를 통해 확인할 수 있다.

▸ Keywords : 클러스터링, 교사 클러스터링, 문맥, 분류

Abstract

Clustering is one of the most popular unsupervised learning methods, which is widely used to form clusters with homogeneous data. Clustering was used to extract contexts corresponding to clusters and a classification method was applied to each context or cluster individually. However, it is difficult to say that the unsupervised clustering is the best context forming method from the view of classification.

In this paper, a new clustering method considering classification was proposed. The proposed method tries to minimize classification error in each cluster when a classification method is applied

•제1저자 : 허경용 •교신저자 : 김성훈
•투고일 : 2014. 5. 15, 심사일 : 2014. 6. 9, 게재확정일 : 2014. 6. 17.
* 동의대학교 전자공학과(Dept. of Electronic Engineering, Dong-Eui University)
** 경북대학교 컴퓨터정보학부(School of Computer Information, Kyungpook National University)
※이 논문은 2013학년도 경북대학교 학술연구비에 의하여 연구되었음.

to each context locally. For this purpose, the proposed method adds constraints forcing two data points belong to the same class to have small distances, and two data points belong to different classes to have large distances in each cluster like in linear discriminant analysis. The usefulness of the proposed method is confirmed by experimental results.

▸ Keywords : Clustering, Supervised clustering, Context, Classification

I. 서 론

퍼지 클러스터링은 60년대 Zadeh[1]의 퍼지 집합 이론까지 거슬러 올라간다. 이를 바탕으로 Ruspini[2]가 퍼지 분할(fuzzy partition)을 클러스터링에 소개하여 기존의 Hard C-Means (HCM) 방법[3]을 퍼지 기반의 방법으로 확장할 수 있도록 하였으며, Dunn[4]이 처음으로 퍼지 클러스터링을 소개하였다. 이후 Bezdek[5]에 의해 일반적인 경우로 확장되었고 현재 Fuzzy C-Means(FCM)로 알려진 방법은 Bezdek의 방법을 일컫는다.

FCM은 여러 분야에서 성공적으로 적용되었으며 융합을 위한 문맥(context) 분할 방법으로도 사용되었다[6]. 융합에서 사용된 클러스터링은 분류기를 적용할 영역, 즉, 문맥을 나누기 위해 사용되었으며 이는 교사 학습 방법의 일부로 사용된 경우이다. 하지만 교사 학습 방법의 일부로 사용되는 경우 클러스터링은 비교사 학습 방법이라는 한계로 인해 분류기의 성능을 최적화한다고 보기 어렵다. 교사 학습 방법으로서의 퍼지 클러스터링 기법도 존재하지만[7], 교사 퍼지 클러스터링의 경우 클래스 정보를 특징값의 일부로 사용하여 클러스터링을 수행함으로써 동일한 클래스에 속하는 데이터 포인트들로 클러스터를 구성하도록 한다. 하지만 교사 클러스터링 역시 분류의 관점에서 최적임을 보장하지는 못한다.

이 논문에서는 분류의 관점에서 최적의 클러스터링을 수행하는 새로운 클러스터링 기법, 특히 융합을 위한 문맥 분할의 방법으로서의 클러스터링 기법을 제안한다. 융합은 전체 특징공간을 균일한 여러 개의 부분공간 또는 문맥으로 나누고 각 부분공간에서 하나 이상의 분류기를 선택적으로 적용하는 방법으로 특정 부분공간에서 우수한 성능을 보이는 분류기가 존재한다는 사실에 근거하고 있다[7]. 이전 방법에서는 FCM을 이용하여 특징공간을 분할하였지만 FCM은 비교사 클러스터링으로 특징값이 비슷한 데이터들로 클러스터를 형성하기 때문에 생성된 부분공간에 분류기를 적용할 경우 오류값을 최소화할 수 있다는 보장은 없다. 따라서 특징공간을 분할함에 있어 각 부분공간에서 국부적으로 오류가 최소화되는 방식으로 클러스터링을 수행함으로써 융합 결과에서 오류값을 줄일 수 있다. 이를 위해 제안한 방법에서는 선형 판별 분석[8]에서와 유사하게 같은 클래스에 속하는 두 데이터 포인트 사이의 거리는 작고 다른 클래스에 속하는 두 데이터 포인트 사이의 거리는 크다는 제약 조건을 FCM의 목적함수에 추가함으로써 분류 오류를 최소화하고자 하였으며 실험 결과를 통해 제안한 방법이 효과적임을 확인할 수 있었다.

이 논문의 구성은 다음과 같다. 먼저 2장에서는 기존 FCM을 다루고 이의 변형으로 클러스터 내에서 분류 오류를 최소화하는 클러스터링 기법인 Within-Cluster-Discriminant Fuzzy Clustering 기법을 제안한다. 제안한 방법의 유용성은 3장에서 실험 결과를 통해 보이며, 결론 및 향후 연구 방향에 관해서는 4장에서 언급한다.

II. 본 론

1. Fuzzy C-Means (FCM)

N개의 D차원 데이터 $X=\left\{x_i|1 \leq i \leq N, x_i \in R^D\right\}$가 주어졌을 때 K개의 클러스터를 형성하기 위해 FCM은 식 (1)의 목적 함수를 최소로 한다[5].

$$J=\sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^{m}\|x_i-v_k\|_A^2=\sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^{m}d_{ki}^{2} \tag{1}$$

이 때 u_{ki}는 x_i가 k번째 클러스터에 소속되는 정도를 나

타내는 소속도 값(membership value)을, v_k는 k번째 클러스터의 중심을, m은 퍼지화 상수(fuzzifier constant)를 나타낸다. 거리 d_{ki}^2은 식 (2)와 같이 표현되며 $A=I$인 경우 유클리드 거리(Euclidean distance)가 된다.

$$d_{ki}^2 = \|x_i - v_k\|_A^2 = (x_i - v_k)^T A (x_i - v_k) \tag{2}$$

소속도 u_{ki}는 데이터 x_i가 모든 클러스터에 소속되는 정도의 합이 1이 되어야 한다는 식 (3)의 제약조건(constraint)을 만족시켜야 한다.

$$\sum_{k=1}^{K} u_{ki} = 1 \tag{3}$$

식 (1)과 (3)을 이용하여 라그랑지 방정식(Lagrange equation)은 식 (4)와 같이 나타낼 수 있다.

$$\mathcal{L} = \sum_{k=1}^{K}\sum_{i=1}^{N} u_{ki}^m d_{ki}^2 - \sum_{i=1}^{N}\lambda_i\left(\sum_{k=1}^{K} u_{ki} - 1\right) \tag{4}$$

식 (4)를 u_{ki}에 대하여 편미분하면 식 (5)를 얻을 수 있다.

$$\frac{\partial\,\mathcal{L}}{\partial\,u_{ki}} = m u_{ki}^{m-1} d_{ki}^2 - \lambda_i \tag{5}$$

식 (5)를 영(zero)으로 놓고 u_{ki}에 대하여 정리하면 소속도에 대한 갱신식(update equation)인 식 (6)을 얻을 수 있다.

$$u_{ki} = \left(\frac{\lambda_i/m}{d_{ki}^2}\right)^{\frac{1}{m-1}} \tag{6}$$

식 (6)에서 라그랑지 상수 λ_i는 식 (3)의 제약조건을 통해 구할 수 있다. 식 (6)를 식 (3)에 대입하여 정리하면 식 (7)을 얻을 수 있다.

$$\left(\frac{\lambda_i}{m}\right)^{\frac{1}{m-1}} = \frac{1}{\sum_{a=1}^{K}\left(\frac{1}{d_{ai}^2}\right)^{\frac{1}{m-1}}} \tag{7}$$

식 (7)을 식 (6)에 대입하면 소속도에 대한 갱신식은 식 (8)과 같이 구해진다.

$$u_{ki} = \frac{\left(\frac{1}{d_{ki}^2}\right)^{\frac{1}{m-1}}}{\sum_{a=1}^{K}\left(\frac{1}{d_{ai}^2}\right)^{\frac{1}{m-1}}} \tag{8}$$

식 (4)를 v_k에 대하여 편미분하면 식 (9)를 얻을 수 있다.

$$\frac{\partial\,\mathcal{L}}{\partial\,v_k} = -2\sum_{i=1}^{N} u_{ki}^m (x_i - v_k) \tag{9}$$

식 (9)를 영(zero)으로 놓고 정리하면 클러스터의 중심 v_k에 대한 갱신식인 식 (10)을 얻을 수 있다.

$$v_k = \frac{\sum_{i=1}^{N} u_{ki}^m x_i}{\sum_{i=1}^{N} u_{ki}^m} \tag{10}$$

식 (8)과 (10)를 이용하여 FCM (Fuzzy C-Means) 알고리듬은 그림 1과 같이 나타낼 수 있다. 이 때 U와 V는 u_{ki}와 v_k로 구성되는 행렬을 나타낸다.

```
Input X : data, K : number of clusters
Initialize u_ki^(0) randomly, t = 0
do
    t ← t + 1
    Update v_k^(t) with Equation (10)
    Update u_ki^(t) with Equation (8)
while ||U^(t) − U^(t−1)|| > ε
return V, U
```

그림 1. FCM 알고리듬
Fig. 1. FCM algorithm

2. Within-Cluster-Discriminative Fuzzy Clustering (WCD-FC)

클러스터링은 주어진 데이터를 균일한(homogeneous) 특

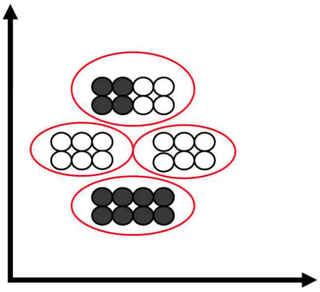
성을 가지는 K개의 부분집합으로 묶는 대표적인 비교사(unsupervised) 학습 방법이다. 클러스터링에서 균일한 정도는 거리 척도(distance measure)를 통해 결정되며 대표적인 거리 척도로는 유클리드 거리(Euclidean distance)가 있다. 하지만 유클리드 거리는 구형의 클러스터만을 찾아낼 수 있는 단점이 있어 타원형의 클러스터를 찾아내기 위한 마할라노비스 거리(Mahalanobis distance)[9, 10], 오목한 클러스터를 찾아내기 위한 측지 거리(geodesic distance)[11] 등 다양한 거리 척도가 이용되고 있다.

클러스터링의 또 다른 변형으로는 교사 학습으로의 변형인 교사(supervised) 클러스터링[7]이 있다. 비교사 학습에서 사용하는 데이터는 특징 벡터 X인 반면, 교사 학습에서 사용하는 데이터는 특징 벡터에 더해서 클래스 라벨 $Y=\{y_i|1 \le i \le N, y_i \in R\}$가 주어진다. 구별(discriminative) 클러스터링[12], Support Vector Machine(SVM)에서 최대 분리 폭의 개념을 도입한 최대 분리 폭 클러스터링(large margin clustering)[13] 등도 기본 개념은 교사 클러스터링과 동일하다. 교사 클러스터링은 클래스 라벨 Y를 특징 벡터 X의 일부로 간주하여 새로운 특징벡터 $X'=\left\{(x_i, f(y_i))|1 \le i \le N, x_i \in R^D, f(y_i) \in R\right\}$을 구성하고 X'을 이용하여 클러스터링을 수행한다. 이 때 $f(\cdot)$은 클래스 라벨을 입력으로 하는 임의의 함수이다. 교사 클러스터링은 동일한 클래스 라벨을 가지는 데이터 포인트들이 동일한 클러스터에 소속되도록 클러스터링을 수행한다.

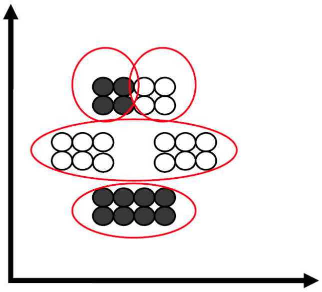
그림 2는 비교사 클러스터링과 교사 클러스터링의 전형적인 예를 비교한 것으로 비교사 클러스터링의 경우 거리 척도에서 가까운 데이터 포인트들이 클러스터를 형성하는 반면, 교사 클러스터링에서는 거리 척도뿐만이 아니라 클래스 라벨도 거리 척도에 반영되어 동일한 라벨의 데이터 포인트들로 클러스터가 구성됨을 알 수 있다.

교사 클러스터링을 분류(classification)의 일종으로 생각할 수 있는 것은 분류가 동일한 클래스 라벨을 가지는 데이터 포인트들을 묶는 방법이기 때문이다. 하지만 두 개의 클래스가 주어진 경우 분류는 두 개의 클러스터를 생성하는 것이 일반적이지만 교사 클러스터링은 지정된 개수로 클러스터링을 수행하는 점에서 다르다. 그림 2의 교사 클러스터링 예에서도 각 클래스별로 2개, 총 4개의 클러스터를 생성하고 있다.

이 논문에서 제안하는 클러스터링은 전통적인 비교사 및 교사 클러스터링과 달리 각 클러스터에 속하는 데이터 포인트들이 클러스터 내에서 클래스 별로 분리가 되도록 클러스터링을 수행하는 방법이다. 제안하는 방법은 FCM의 변형으로 클래스 라벨을 부가 정보로 사용한다는 점에서 교사 클러스터링



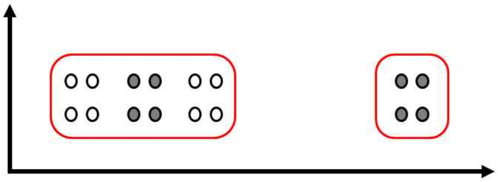
(a) 비교사 클러스터링



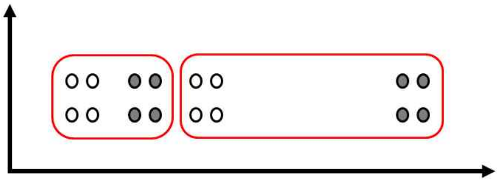
(b) 교사 클러스터링

그림 2. 비교사 및 교사 클러스터링
Fig. 2. Unsupervised and supervised clustering

과 유사하다. 하지만 교사 클러스터링이 클러스터 내에 하나의 클래스 라벨만이 존재하도록 클러스터링을 수행하는 것과 달리, 동일한 클러스터 내의 서로 다른 클래스에 속하는 데이터 포인트들의 분리도를 높이는 방향으로 클러스터링을 수행한다는 점에서 교사 클러스터링과 구별된다. 그림 3은 비교사 클러스터링과 WCD-FC의 차이점을 보인 것으로 제안한 방법의 경우 클러스터 내에서 동일 클래스 라벨을 갖는 데이터



(a) 비교사 클러스터링



(b) WCD-FC

그림 3. 비교사 클러스터링 및 WCD-FC
Fig. 3. Unsupervised clustering and WCD-FC

포인트들이 뭉쳐서 나타나므로 클러스터 내에서 분류가 쉽도록 클러스터링 됨을 알 수 있다. WCD-FC에서 클러스터 내 클래스의 분리도는 '동일 클래스에 속하는 데이터 포인트들 사이의 거리는 가깝고 다른 클래스에 속하는 데이터 포인트들 사이의 거리는 멀다.'라는 규칙을 사용하였다.

N개의 D차원 데이터 $X=\{x_i|1 \leq i \leq N, x_i \in R^D\}$와 클래스 라벨 $Y=\{y_i|1 \leq i \leq N, y_i \in R\}$가 주어졌을 때 이를 K개의 클러스터로 클러스터링하기 위해서 WCD-FC는 식 (11)의 목적 함수를 최소화 한다.

$$\begin{aligned} J= & \sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^m\|x_i-v_k\|_A^2 \\ & +\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}u_{ki}^m u_{kj}^m\left(\exp\left(\alpha\|x_i-x_j\|_A^2\right)-1\right)I(y_i=y_j) \\ & +\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}u_{ki}^m u_{kj}^m\left(\exp\left(-\beta\|x_i-x_j\|_A^2\right)-1\right)I(y_i\neq y_j) \\ = & \sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^m d_{ki}^2+\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}u_{ki}^m u_{kj}^m\left(d_{\alpha,ij}^2 I_{ij}+d_{\beta,ij}^2\overline{I_{ij}}\right) \end{aligned} \tag{11}$$

상수 $\alpha(>0)$와 $\beta(>0)$는 클래스 분리 정도를 클러스터링에 반영하는 상수이며 $I(\cdot)$는 매개 변수가 참이면 1의 값을, 거짓이면 0의 값을 갖는 지시 함수(indicator function)로 $I(y_i=y_j)=I_{ij}$, $I(y_i\neq y_j)=\overline{I_{ij}}$로 나타낸다. $d_{ki}^2=\|x_i-v_k\|_A^2$는 클러스터 중심에서 데이터 포인트까지의 거리를 나타내며, $d_{\alpha,ij}^2=\exp\left(\alpha\|x_i-x_j\|_A^2\right)-1$와 $d_{\beta,ij}^2=\exp\left(-\beta\|x_i-x_j\|_A^2\right)-1$는 각각 데이터 포인트들 사이의 거리에 비례 및 반비례하는 거리 척도로 두 데이터 포인트가 동일한 경우 영(zero)의 값을 가지며 상수 α와 β에 의해 감쇄 정도가 제어되도록 하였다. 식 (11)에서 두 번째 항은 동일 클러스터에 속하면서 동일한 클래스 라벨을 갖는 데이터 포인트 사이의 거리는 가까워야 함을, 세 번째 항은 동일 클러스터에 속하면서 서로 다른 클래스 라벨을 갖는 데이터 포인트 사이의 거리는 멀어야 함을 반영한다. 식 (11)은 식 (12)와 같이 나타낼 수 있다.

$$J=\sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^m\left(d_{ki}^2+\sum_{j=1}^{N}u_{kj}^m\left(d_{\alpha,ij}^2 I_{ij}+d_{\beta,ij}^2\overline{I_{ij}}\right)\right) \tag{12}$$

식 (12)는 식 (1)의 FCM 목적 함수와 동일한 형태를 가지며 거리 척도가 마할라노비스 거리에서 클러스터 내 다른 데이터 포인트들과의 거리를 고려한 척도로 바뀐 점에서 다르다. 식 (12)와 식 (3)의 제약조건을 이용하여 라그랑지 방정식은 식 (13)과 같이 나타낼 수 있다.

$$\begin{aligned} \mathcal{L}= & \sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^m\left(d_{ki}^2+\sum_{j=1}^{N}u_{kj}^m\left(d_{\alpha,ij}^2 I_{ij}+d_{\beta,ij}^2\overline{I_{ij}}\right)\right) \\ & -\sum_{i=1}^{N}\lambda_i\left(\sum_{k=1}^{K}u_{ki}-1\right) \end{aligned} \tag{13}$$

식 (13)에서 $i=j$인 경우 $d_{\alpha,ij}^2=d_{\beta,ij}^2=0$이므로 식 (13)은 식 (14)와 같이 쓸 수 있다.

$$\begin{aligned} \mathcal{L}= & \sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^m\left(d_{ki}^2+\sum_{j=1,\dots,N\ j\neq i}u_{kj}^m\left(d_{\alpha,ij}^2 I_{ij}+d_{\beta,ij}^2\overline{I_{ij}}\right)\right) \\ & -\sum_{i=1}^{N}\lambda_i\left(\sum_{k=1}^{K}u_{ki}-1\right) \\ = & \sum_{k=1}^{K}\sum_{i=1}^{N}u_{ki}^m\overline{d_{ki}}^2-\sum_{i=1}^{N}\lambda_i\left(\sum_{k=1}^{K}u_{ki}-1\right) \end{aligned} \tag{14}$$

식 (14)를 u_{ki}에 대하여 편미분하고 식 (3)의 제약조건을 이용하여 정리하면 소속도 u_{ki}에 대한 갱신식, 식 (15)를 얻을 수 있다.

$$u_{ki}=\frac{\left(\frac{1}{\overline{d_{ki}}^2}\right)^{\frac{1}{m-1}}}{\sum_{a=1}^{K}\left(\frac{1}{\overline{d_{ai}}^2}\right)^{\frac{1}{m-1}}} \tag{15}$$

식 (14)를 v_k에 대하여 편미분하여 정리하면 클러스터 중심 v_k에 대한 갱신식, 식 (16)을 얻을 수 있다.

$$v_k=\frac{\sum_{i=1}^{N}u_{ki}^m x_i}{\sum_{i=1}^{N}u_{ki}^m} \tag{16}$$

식 (16)은 FCM에서 클러스터 중심의 갱신식인 식 (10)과 동일하다. 식 (15)과 (16)을 이용하여 WCD-FC 알고리

들은 그림 4와 같이 나타낼 수 있다.

```
Input X : data, K : number of clusters
Initialize u_ki^(0) randomly, t = 0
do
    t ← t + 1
    Update v_k^(t) with Equation (16)
    Update u_ki^(t) with Equation (15)
while ||U^(t) − U^(t−1)|| > ϵ
return V, U
```

그림 4. WCD-FC 알고리듬
Fig. 4. WCD-FC algorithm

III. 실험 결과

그림 5의 데이터에 대하여 3번 그룹과 4번 그룹 사이의 거리 d를 변화시키면서 FCM과 제안한 방법인 WCD-FC을 적용하였다.

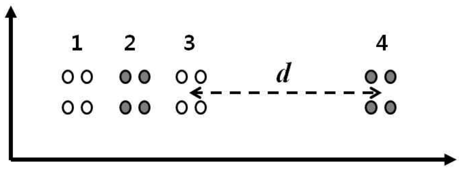
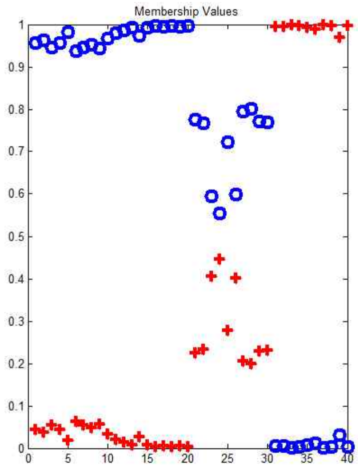
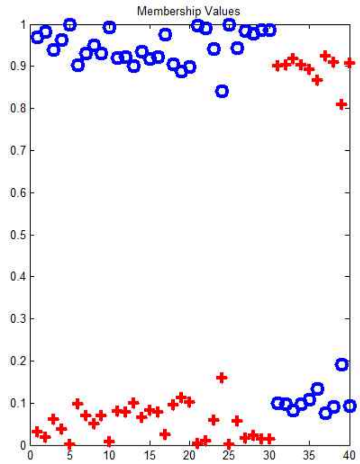


그림 5. 데이터 집합
Fig. 5. Data set

d가 큰 경우, WCD-FC에서 사용하는 거리 척도에 분류 오류보다는 마할라노비스 거리가 많은 영향을 미치므로 {1, 2, 3}과 {4}의 두 개 클러스터를 형성하는 점에서는 FCM에서와 동일하였다. 하지만 그림 6에 나타난 바와 같이 3번 그룹에 속하는 데이터 포인트들의 소속도 값은 FCM의 경우 두 클러스터의 경계에 걸쳐져 0.5 근처의 값이 나온 반면 WCD-FC의 경우에는 1번이나 2번 그룹에서와 크게 차이가 나지 않는 것을 확인할 수 있다. 즉, 데이터 포인트들이 보다 명확하게 특정 클래스에 소속됨으로써 클러스터 내에서의 분류 알고리듬 적용이 보다 용이해 진다.



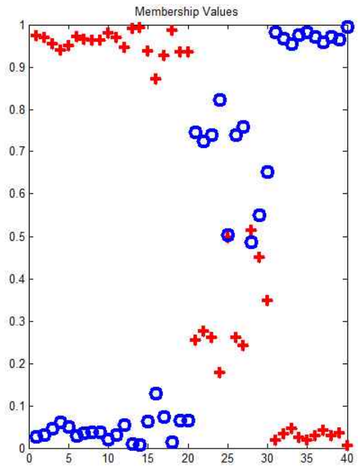
(a) FCM



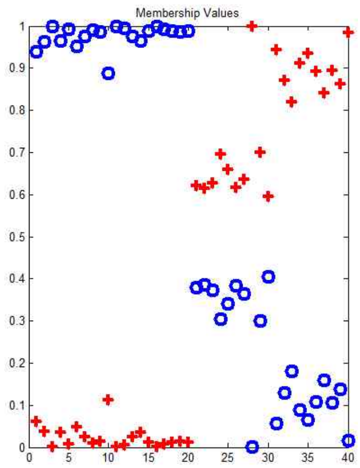
(b) WCD-FC

그림 6. 실험 결과 1
Fig. 6. Experimental result 1

d값이 작아지면 3번과 4번 그룹의 거리가 가까워져서 FCM에서도 {1, 2}와 {3, 4}의 두 개 클러스터를 형성하기 시작한다. 하지만 d가 충분히 작은 값을 가지지 않으면 3번 그룹의 데이터 포인트들은 두 클러스터에 나뉘어 소속되는 현상이 발생한다. 하지만 WCD-FC에서는 분류 오류가 거리 척도에 반영됨으로써 3번 그룹의 데이터 포인트들이 서로 다른 클러스터에 소속되는 현상은 발생하지 않으므로 분류 측면에서 보다 유리함을 알 수 있다.



(a) FCM



(b) WCD-FC

그림 7. 실험 결과 2
Fig. 7. Experimental result 2

IV. 결 론

클러스터링은 대표적인 비교사 학습 방법으로서 다양한 분야에 적용되고 있으며 융합을 위한 문맥 결정도 그 중 하나이다. 하지만 클러스터링은 비교사 학습 방법이고 융합은 교사 학습 방법이므로 비교사 학습 방법의 결과를 교사 학습을 위해 적용하는 것은 최적의 결과를 보장하기 어렵다. 따라서 이 논문에서는 기존에 널리 사용되는 클러스터링 알고리듬인 FCM에 분류 오류를 최소화하는 제약 조건을 첨가하여 WCD-FC을 제안하였다. 제안한 알고리듬은 분리된 클러스터, 즉, 문맥에서 FCM에 비해 분류 오류를 최소화할 수 있도록 보다 직관적인 클러스터를 형성함을 실험 결과를 통해 확인하였다.

제안한 WCD-FC는 교사 학습 방법의 일종으로 생각할 수 있으며 클래스 라벨을 활용하여 새로운 특징 공간을 형성하고 그 공간에서 클러스터링을 시행한다. 이 과정에서 특징 공간을 형성하는 방법은 클래스 라벨에 의한 분류 오류를 WCD-FC의 거리척도에 반영하는 방법에 따라 결정되며 이는 클러스터링을 위해 주어진 원래의 특징 공간을 바탕으로 결정되어야 할 것으로 생각된다. 현재 기존 특징 공간을 바탕으로 WCD-FC를 위한 특징 공간을 자동으로 결정하는 방법에 관해 연구 중에 있다.

참고문헌

〔1〕 L. A. Zadeh, "Fuzzy sets," Information and Control, Vol. 8, No. 3, pp. 338-353, June 1965.

〔2〕 E. H. Ruspini, "A new approach to clustering," Information and Control, Vol. 15, No. 1, pp. 22-32, July 1969.

〔3〕 Rui Xu and Donald Wunsch II, "Survey of Clustering Algorithms," IEEE Transactions on Neural Networks, Vol. 16, No. 3, pp. 645-678, May 2005.

〔4〕 J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well separated clusters," Journal of Cybernetics, Vol. 3, No. 3, pp. 32-57, Apr. 1974.

〔5〕 J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," Plenum, New York, 1981.

〔6〕 G. Heo, P. Gader and H. Frigui, "A Noise Robust Variant of Context Extraction for Local Fusion," Proceedings of 2010 IEEE International Conference on Fuzzy Systems, pp. 1-6, 2010.

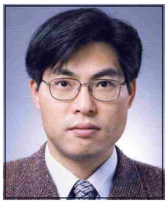
〔7〕 A. M. Bensaid, "Partially supervised clustering for image segmentation," Pattern Recognition, Vol. 29, No. 5, pp. 859-871, May 1996.

〔8〕 A. M. Martinez and A. C. Kak, "PCA versus LDA," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23 No. 2, pp. 228-233, Feb. 2001.

[9] D. E. Gustafson and W. C. Keller, "Fuzzy clustering with a fuzzy covariance matrix," Proceedings of the 1978 IEEE Conference on Decision and Control, pp. 761-766, 1979.
[10] R. Babuska, P. J. van der Veen and U. Kaymak, "Improved Covariance Estimation for Gustafson-Kessel Clustering," Proceedings of the 2002 IEEE International Conference on Fuzzy Systems, pp. 1081-1085, 2002.
[11] B. Feil and J. Abonyi, "Geodesic Distance Based Fuzzy Clustering," Advances in Soft Computing, Vol. 39, pp. 50-59, 2007.
[12] Gomes, Ryan, Andreas Krause, and Pietro Perona. "Discriminative clustering by regularized information maximization," Advances in Neural Information Processing Systems, Vol. 23 pp. 775-783, 2010.
[13] K. Zhang, I. W. Tsang and J. T. Kwok, "Maximum Margin Clustering Made Practical," IEEE Transactions on Neural Networks, Vol. 20, No. 4, pp. 583-596, Apr. 2009.

저 자 소 개

허 경 용
1994: 연세대학교
전자공학과 공학사.
1996: 연세대학교 대학원
전자공학과 공학석사.
2009: University of Florida
컴퓨터공학과 공학박사
현 재: 동의대학교 전자공학과 교수
관심분야: 인공지능, 패턴인식, 로봇공학
Email : hgycap@deu.ac.kr



김 성 훈
1988: 서강대학교
전자공학과 공학사.
1990: 연세대학교 대학원
전자공학과 공학석사.
1996: 연세대학교 대학원
전자공학과 공학박사
현 재: 경북대학교 컴퓨터정보학부 교수
관심분야: 인공지능, 패턴인식
Email : shkim1454@knu.ac.kr