工學碩士　學位論文

평균내부거리를 적용한 퍼지 클러스터링 알고리즘에 의한 영상분할

Image Segmentation Based on the Fuzzy Clustering Algorithm using Average Intracluster Distance

指導教授 趙 奭 濟

韓國海洋大學校 大學院

制 御 計 測 工 學 科

兪 鉉 在

목　　차

Image Segmentation Based on the Fuzzy Clustering Algorithm using Average Intracluster Distance

Hjyoun-jai, You

Dept. of Control & Instrumentation Engineering, Graduate School, Korea Maritime University

Abstract

Image segmentation is one of the important processes in the image information extraction for computer vision systems. The fuzzy clustering methods have been extensively used in the image segmentation because it extract feature information of the region. Most of fuzzy clustering methods have used the Fuzzy C-means(FCM) algorithm. This algorithm can be misclassified about the different size of cluster because the degree of membership depends on highly the distance between data and the centroids of the clusters.

This paper proposes a fuzzy clustering algorithm using the Average Intracluster Distance that classifies data uniformly without regard to the size of data sets. The Average Intracluster Distance take an average of the vector set belong to each cluster and increase in exact propotion to it's size and density. The experimental results demonstrate that the proposed approach has the good classication entropy and validity function.

제 1 장 서론

영상분할은 영상을 그 구성 성분 또는 물체별로 나누는 작업을 말하며 컴퓨터에 의한 영상 분석에서 반드시 다루어야 할 부분이다. 영상분할을 통한 물체의 모습이나 구성 성분은 다음 단계에서 이어질 영상해석이나 인지 또는 묘사에서 필수적으로 요구된다[1-3].

영상 분할 방법은 스펙트럼 해석(Spectrum analysis)방법, 이완화(Relaxation)방법, 문턱치(Thresholding)방법, 클러스터링(Clustering)방법 등이 있다. 스펙트럼 해석 방법은 영상 분할을 할 때 물체의 밝기정보에 대한 선지식(a *priori* knowledge)이 필요하다. 문턱치 방법은 영상 분할에 폭 넓게 사용되어온 방법으로 히스토그램의 적절한 문턱치를 구하여 영상을 분할한다. 클러스터링을 이용한 방법은 이완화 방법과 마찬가지로 계산량과 메모리가 많이 필요하고 클러스터의 개수를 미리 지정해 주어야 하는 문제점이 있으나 우수한 결과를 기대할 수 있어서 많이 연구되고 있다 [4-6].

클러스터링 방법은 단순 클러스터링(Hard clustering)과 FCM(Fuzzy C-Means), PCM(Possibilistic C-Means)등의 방법이 있다. 단순 클러스터링 기법은 주어진 데이터 상호간의 관계가 명확하다는 가정 하에 각 패턴을 분할하는 방식이다. 따라서 이 방법은 다루고자 하는 데이터의 경계가 명확하지 않을 경우 실제 데이터 상호간의 군집성을 묘사하기에 부적절할 뿐만 아니라, 주어진 데이터 분포의 성질을 손실하는 결과를 초래하게 된다. 이를 개선하기 위해, Bezdek[7]은 퍼지분할에 의한 FCM 알고리즘을 제안하였다. FCM 알고리즘은 각 패턴이 특정 클러스터의 중심에 속하는 소속정도를 부여하므로 보다 정확한 정보를 획득하도록 하였다. 그러나 이 알고리즘은 데이터로부터 각 클러스터들에 대한 소속정도의 합이 1

이 되는 확률적 제약조건(Probabilistic constraint)을 이용하므로 소속함수 값이 소속성(Belonging)이나 적합성(Compatibility) 정도의 직관적인 개념과 항상 일치하지는 않는다. Raghu[8]는 위의 문제점을 개선하기 위하여 PCM 알고리즘을 제안하였다. 이 방법은 소속정도의 값을 다른 클러스터와 관계가 없고, 각 패턴과 해당 클러스터 중심간의 절대 거리에만 의존하여 구하므로 FCM 알고리즘이 가지고 있었던 상대성 문제를 해결할 수가 있었다. 그러나 PCM 알고리즘은 각 영역에 해당하는 클러스터들의 크기와 초기 중심 값을 사전에 알아야함으로 Scale Space Filtering[9]과 같은 전처리 과정이 필요하다. FCM 방법은 어떤 패턴이 각 클러스터 중심까지의 거리가 같으면 각 클러스터에 대해 소속정도가 같다. 그러므로 중심탐색에 있어 클러스터의 크기가 다른 경우 문제점이 발생한다. 이러한 문제점을 해결하기 위해 평균내부거리를 이용한 클러스터링 방법[10]을 제시한다.

제안된 알고리즘은 각 패턴들로부터 내부클러스터까지의 거리에 의해 소속정도를 부여하고 중심을 탐색한다. 내부클러스터는 평균내부거리 안에 속하는 패턴들의 집합을 말하며 클러스터의 크기와 밀도에 비례하게 된다. 그 결과 소속정도를 클러스터의 크기에 상관없이 균일하게 부여하므로 중심 탐색 능력을 향상시킬 수 있었다. 또한 제안된 클러스터링 방법을 영상에 적용하여 영상분할을 시도하였고 분할된 영상의 성능을 평가하기 위하여 Bezdek[11]과 Xie, Beni[12]의 적합도 함수를 이용하였다 [13-16].

본 논문은 2장에서 클러스터링 방법과 FCM알고리즘에 대해 설명하고, 3장에서는 평균내부거리를 적용한 퍼지 클러스터링에 대해서 설명한다. 그리고 4장에서는 기존의 FCM 알고리즘과 평균내부거리를 적용한 퍼지 클러스터링에 의한 영상분할 방법을 실험하여 기존의 방법보다 성능이 우수함을 보인다. 그리고 5장에서 결론을 맺는다.

제 2 장 클러스터링 알고리즘

클러스터링 알고리즘의 주된 목적은 주어진 자료 집합을 특징이 유사한 클러스터들(Homogeneous clusters)로 나누는 것이다. 여기서 유사하다는 것은 그들간의 특징 벡터 공간상에서 같은 그룹내 모든 점들이 서로 근접하고 다른 그룹의 점들과 멀리 떨어져 있다는 것을 의미한다. 클러스터링 알고리즘은 특징 벡터들을 분류하거나 관련된 정보를 유지하도록 하면서 자료 집합의 크기를 줄이는데 사용된다. 어느 한 자료의 집합 내에는 서로 다른 클러스터들이 함께 존재하게 된다. 이러한 클러스터들을 분류에 대해 본 장에서 다루고자 한다.

2.1 유사도

데이터 군집을 정의하기 위해서는 특징 벡터를 특정한 군집 중심의 영역으로 배정하기 위한 유사도(Similarity measure)를 정의해야 한다. 먼저 거리가 가까울수록 유사도가 커지는 특징 벡터들의 유사도로서, 두 특징 벡터 x와 z간의 유클리드 거리를 다음과 같이 정의하였다.

$$D = \| x - z \| \tag{2.1}$$

그러나, 종종 유용하게 사용할 수 있는, 또 다른 의미 있는 거리 척도가 존재한다. 예를 들어, 특징 벡터 x로부터 모델 m까지의 Mahalanobis 거리는 통계적 특성을 명시적으로 고려할 때 유용한 유사도이다.

$$D = (x - m)' C^{-1} (x - m) \tag{2.2}$$

여기서 C는 특징 벡터 모집단의 공분산 행렬이고, m은 평균 벡터이며, x는 임의의 특징 벡터를 나타낸다.

$$s(x, z) = \frac{x' z}{||x|| ||z||} \tag{2.3}$$

유사도를 거리 척도로 제한할 필요는 없다. 예를 들어 두 특징 벡터 x와 z가 이루는 각의 코사인 값인 비계량(Nonmetric) 유사도 함수는 식 (2.3)과 같으며 x와 z가 원점에 대하여 같은 방향으로 위치했을 때 최대가 된다. 이 유사도는 군집 영역이 주축을 따라 나타나는 경향이 있을 때 유용한 척도이다. 그러나, 이 유사도는 군집 영역이 원점에 대하여 떨어져 있고, 또 군집 영역들 간에도 충분히 떨어져 있을 때 사용할 만하다.

2.2 클러스터 분류 기준

특징 벡터의 유사도를 채택한 다음에는 주어진 데이터를 군집 영역으로 분할하기 위한 절차가 남아있다. 사용될 군집화 기준(Clustering criterion)은 휴리스틱 기법으로 표현하거나 특정한 성능 지표의 최소(또는 최대)화에 바탕을 둘 수 있다.

휴리스틱 접근방법은 경험과 직관에 의한 것으로, 특징 벡터를 군집 영역에 배정하기 위해 선택한 유사도를 이용하는 규칙들로 구성된다. 그러나, 두 특징 벡터간의 근접은 상대적인 유사도이기 때문에, 보통 군집 탐색 과정에서 수용 가능한 유사도를 정의하기 위한 임계값을 설정할 필요가 있다.

성능 지표 접근 방법은 선택된 성능 지표를 최대화하거나 최소화하는 절차이며, 가장 자주 사용되는 지표는 제곱 오차의 합 지표로서 다음과 같다.

$$J = \sum_{j=1}^{N_c} \sum_{x \in S_j} ||\boldsymbol{x} - m_j||^2 \qquad (2.4)$$

여기서, N_c는 군집 영역의 개수이고, S_j는 j번째 군집 영역에 속하는 표본 집합이며, m_j는 집합 S_j에 속한 표본의 평균 벡터로 다음과 같다.

$$m_j = \frac{1}{N_j} \sum_{x \in S_j} \boldsymbol{x} \qquad (2.5)$$

식 (2.5)에서 N_j는 집합 S_j에 속하는 특징 벡터의 개수를 의미한다. 식 (2.4)의 지표는 군집 영역에 속하는 표본과 평균 벡터간의 차를 제곱하여 합산한 것이다.

위에 주어진 것 이외에도 다른 많은 성능 지표가 있는데 표본간의 평균 제곱 거리, 다른 군집 영역에 속하는 표본간의 평균 제곱 거리, 산포 행렬의 개념에 기반을 둔 지표, 최대 최소 분산 지표 등이 여기에 해당된다.

2.3 단순 클러스터링 알고리즘

단순 클러스터링 알고리즘은 어떤 제한을 갖고 있기는 하지만, 주어진 데이터 집합에 대하여 대략적인 특성을 만드는 단순하면서도 빠른 방법이다. 또한 특정한 T값으로 군집 영역을 설정하기 위하여 표본 집합에 대

하여 단지 한번의 계산 과정만을 필요로 하기 때문에 계산상으로 매우 효과적인 방법이다. 그러나 표본의 기하학적 분포를 제대로 분석하기 위해서는 다양한 임계값과 여러 시작점에 대하여 충분히 실험해야 할 필요가 있다. 일반적으로 표본들은 결과를 시각적으로 해석할 수 없는 고차원의 데이터이기 때문에 이런 정보는 데이터에 대하여 각 과정을 거친 결과의 영역에 포함되는 표본들의 개수 및 군집 중심간의 거리를 비교하여 얻을 수 있다. 또한 각 영역에 대하여 군집 중심으로부터의 최근접 점과 가장 멀리 떨어진 점, 그리고 군집 영역내의 분산 등은 유용한 매개변수이다. 데이터에 대하여 각 과정을 통해 얻어진 이와 같은 정보는 다음 단계에 적용할 새로운 T값과 시작점을 선택할 때 유용하게 사용될 수 있다. 이 과정은 데이터가 임계값으로 매우 잘 분리되는 성질을 가질 때 유용하다.

N개의 표본 특징 벡터 집합을 $X = \{x_{1,}\ x_2,\ \cdots,\ x_N\}$라 하고 클러스터 중심들의 집합을 $V = \{v_1, v_2, \cdots, v_k\}$라 할 때 단순 클러스터링 알고리즘은 다음과 같은 과정으로 동작한다. 표본 특징 벡터 중의 하나를 첫 번째 군집 중심 z_1로 놓고, 임의의 음수가 아닌 임계값 T를 선택하자. 편의상 x_1을 z_1으로 선택하자. 그리고 x_2와 z_1간의 거리 D_{21}를 계산한다. 이 거리가 T를 초과하면 새로운 군집 중심 $z_2 = x_2$를 생성하고, 그렇지 않으면, 군집 중심 z_1의 영역에 x_2를 배정한다. 두 번째 군집 중심 z_2를 설정하기 위해 $D_{21} > T$라고 가정하자. 다음 단계에서 x_3에서 z_1및 z_2간의 거리 D_{31}및 D_{32}를 계산한다. D_{31}과 D_{32}가 모두 T보다 크면, 새로운 군집 중심 $z_3 = x_3$를 생성하고, 그렇지 않으면 최근접하는 군집 중심 영역에 x_3을 배정한다. 마찬가지 방법으로, 각각의 새로운 특징 벡터에 대하여 모든 군집간의 거리를 계산하고 임계값을 넘는지를 검사하여, 거리가 모두 T를 넘으면 새로운 군집 중심을 생성한다. 그렇지 않으면 최근접하는 군집

중심 영역에 그 특징 벡터를 배정한다. 이 과정의 결과는 첫 번째 선택된 군집 중심과, 특징 벡터가 고려되는 순서, T의 값, 데이터의 기하학적 특성 등에 의존한다.

2.4 Fuzzy C-Means 알고리즘

주어진 데이터의 집합을 균일한 요소들로 단순히 분할하는 것 뿐 아니라, 그 분할된 각 클러스터들을 잘 대표하는 원소를 찾기를 원하는 경우가 있다. 이를 위해 구현된 알고리즘이 C-means 알고리즘이다. 주어진 n개의 자료들 $x_1, x_2, \cdots, x_n$이 있다고 하고 이 데이터들의 집합을 X라 하자. 여기서 x_j $(1 \le j \le n)$는 d차원의 데이터 벡터이고, $X = \{x_1, x_2, \cdots, x_n\}$는 데이터 벡터 집합이다. 이것을 c개의 클러스터로 분할하는 경우에 대해 알아보면, 먼저 j 번째 데이터가 i 번째 클러스터에 속하는 소속정도를 u_{ij}로 나타낸다. 이때 단순 클러스터링에서는 u_{ij}가 0과 1의 두 가지 값을 갖지만, 퍼지 클러스터링(fuzzy clustering)에서는 0에서 1 사이의 적당한 값을 가질 수가 있다. 즉 퍼지 클러스터링은 1개의 데이터가 2가지 이상의 클러스터에 속하는 것을 허용하고 있음을 알 수 있다. 그림 2.1은 단순 클러스터링과 퍼지 클러스터링의 차이점을 보여 주고 있다. (a)는 경계영역이 애매한 데이터 집합을 보여주고, (b)는 단순 클러스터링에 의 해 경계영역이 0과 1로 명확하게 분할된 경우를 나타낸다. 그리고 (c)는 경계영역에 대해 0과 1사이의 소속정도를 가진 퍼지 클러스터링을 보여주고 있다.

퍼지 클러스터링 방법중 가장 많이 알려져 있고 폭넓게 사용되고 있는 클러스터링 알고리즘으로는 Bezdek[7]에 의해 제안되어진 FCM(Fuzzy C-Means)알고리즘이다. 이 알고리즘은 식 (2.6)과 같이 각 데이터와 각

(a) 원그림

(b) 단순 클러스터링 (c) 퍼지 클러스터링

$u_{ij} \in \{0, 1\}$ 단순 클러스터링의 소속 정도

$u_{ij} \in [0, 1]$ 퍼지 클러스터링의 소속 정도

단, 모든 i 에 대해, $0 < \sum_{j=0}^{n} u_{ij} < n$ 이고

모든 j 에 대해, $\sum_{i=0}^{c} u_{ij} = 1$

그림 2.1. 단순 클러스터링과 퍼지 클러스터링의 비교

Fig. 2.1. Comparison between hard clustering and fuzzy clustering

클러스터 중심과의 거리를 고려한 유사도 측정에 기초한 목적함수를 최소화할 수 있도록 데이터 집합을 분할하는 방법이다.

$$J_m(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} (u_{ij})^m \| x_j - v_i \|^2 \qquad (2.6)$$

여기서, 모든 j 에 대해 $\sum_{i=1}^{c} u_{ij} = 1$

$$u_{ij} = \{ \sum_{k=1}^{c} (\frac{d_{ij}}{d_{kj}})^{\frac{2}{(m-1)}} \}^{-1} \qquad (2.7)$$

여기서, $d_{ij} = \| v_i - x_j \|$, $(i = 1, 2, \cdots, c \ ; j = 1, 2, \cdots, n)$

$$v_i = \frac{\sum_{j=0}^{n} (u_{ij})^m x_j}{\sum_{j=0}^{n} (u_{ij})^m} \qquad (2.8)$$

식 (2.8)에서 v_i $(1 \le i \le c)$는 i번째 클러스터의 중심이고, $V = \{v_1, v_2, \cdots, v_c\}$는 클러스터 중심들의 집합을 나타낸다. $m \in [1, \infty)$은 퍼지 정도를 나타내는 가중치 지수이고, m이 1인 경우에는 k-means 방법과 같고, 그 값이 클수록 퍼지 정도가 강하게 된다. u_{ij}는 j 번째 데이터가 i 번째 클러스터에 속하는 소속정도를 나타내고 $c \times n$인 행렬로 나타나는 퍼지 분할 U로 나타낼 수가 있다. 여기서 m이 1보다 큰 경우에 모든 i, j에 대해서 $v_i \neq x_j$를 만족한다고 가정하면 위의 식을 만족할 때만 (U, V)가 J_m의 최소화를 가능하게 한다. FCM 알고리즘은 식(2.7)와 식(2.8)의 과정을 반복하므로 J_m은 어떤 정해진 값으로 수렴하게 된다.

FCM알고리즘의 수행과정은 다음과 같다.

단계 1 : 초기 클러스터의 개수(c)와 퍼지 가중치 정도(m)를 결정한다.

$$(2 \le c < n, \ 1 < m < \infty)$$

단계 2 : 중심값과 소속정도를 임의로 초기화한다.

단계 3 : 새로운 중심값을 갱신한다. ($l = 0, 1, 2, \cdots$)

$$v_i = \frac{\sum_{j=0}^{n} u_{ij}^{(l)m} x_k}{\sum_{j=0}^{n} u_{ij}^{(l)m}}$$

단계 4 : 각 데이터로부터 모든 클러스터 중심에 대한 소속정도를 구한다.

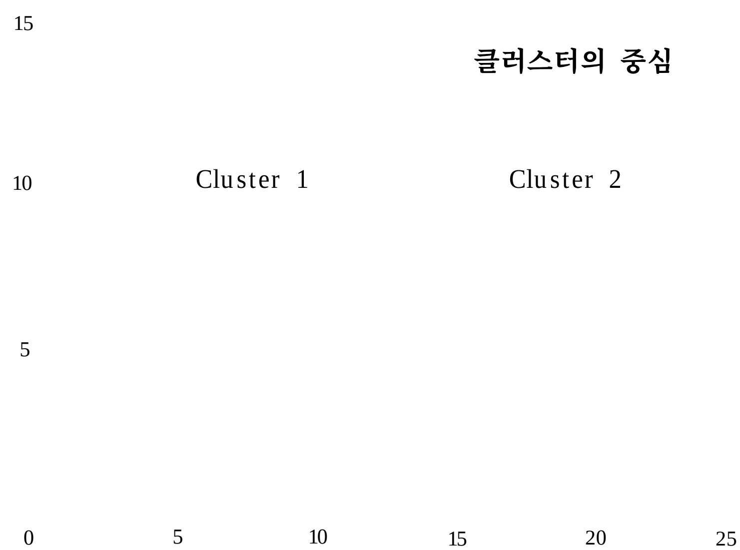
$$u_{ij}^{(l+1)} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\| x_j - v_i \|}{\| x_j - v_k \|} \right)^{\frac{2}{(m-1)}}}$$

단계 5 : 모든 소속정도의 변화가 충분히 작다면 끝내고 그렇지 않으면 l을 1증가하고 단계 3으로 이동한다.

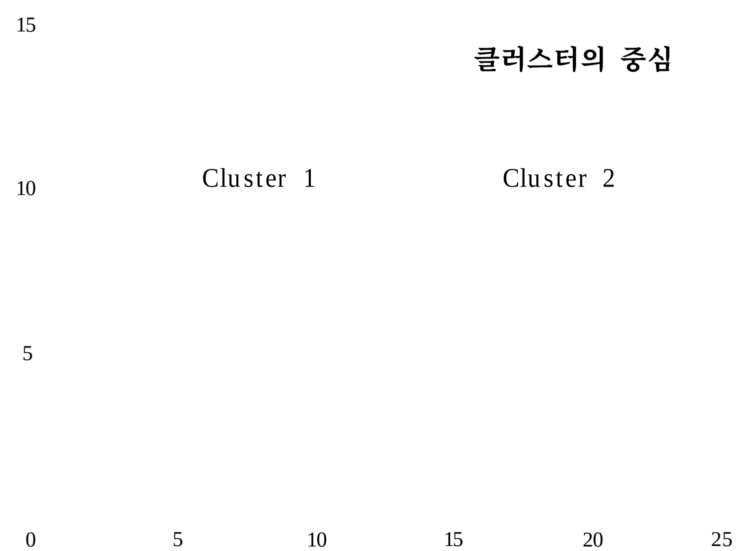
$\| U(l) - U(l-1) \| < \delta$, δ 는 0과 1 사이의 아주 작은 값

FCM 알고리즘은 각 데이터로부터 클러스터 중심까지의 거리에 의해 소속정도를 부여하게되므로 클러스터의 크기가 다른 경우 잘못된 소속정도를 가지게 된다. 이러한 문제점을 다음 예에서 보여주고 있다. 그림 2.2는 두 개의 클러스터로 이루어진 2차원 데이터 집합을 보여주고 있다. 데

이터 집합 A는 Cluster 1과 Cluster 2의 크기가 같고, B는 Cluster 1이 상대적으로 작은 데이터 집합이다. FCM 알고리즘은 데이터로부터 클러스터 중심에 이르는 거리의 제곱 합을 최소화하는 조건에 의해 소속정도를 결정한다. 이렇게 구해진 소속정도에 의해 중심 탐색이 이루어진다. 그리고 그림 2.2는 각 데이터 집합에서 구해진 중심점을 보여주고 있다. 데이터 집합 B에서 알 수 있듯이 FCM 알고리즘은 클러스터의 크기가 다른 경우에는 잘못된 중심점을 찾고 데이터를 오분류한다. 표 2.1과 표 2.2는 경계 영역에 속한 데이터의 소속정도를 보여주고 있다. 여기서 클러스터의 크기가 같은 경우에는 군집화가 잘 이루어지고 있지만 다른 경우에는 작은 클러스터에 많은 소속정도를 주게 되어 잘못된 군집화가 이루어지고 있음을 알 수가 있다.



(a) **클러스터의 크기가 같은 경우**(**데이터 집합** A)



(b) **클러스터의 크기가 다른 경우**(**데이터 집합** B)

그림 2.2. **두 개의 클러스터를 가지는 데이터 집합**

Fig. 2.2. Data sets with two clusters

표 2.1. **경계영역 데이터가 각 클러스터에 속할 소속정도(집합** A)

Table 2.1. Membership grades belong to the each clusters of the boundary data(Data set A)

	좌 표	Cluster 1에 속할 소속정도(u_{ij})	Cluster 2에 속할 소속정도(u_{ij})
경계영역에 속한 데이터(x_j)	X (15,3)	0.3646	0.6354
	X (15,4)	0.3439	0.6561
	X (15,5)	0.3278	0.6723
	X (15,6)	0.3214	0.6786
	X (15,7)	0.3278	0.6722
	X (15,8)	0.3442	0.6558
	X (15,9)	0.3650	0.6350

표 2.2. **경계영역 데이터가 각 클러스터에 속할 소속정도(집합** B)

Table 2.2. Membership grades belong to the each clusters of the boundary data(Data set B)

	좌 표	Cluster 1에 속할 소속정도(u_{ij})	Cluster 2에 속할 소속정도(u_{ij})
경계영역에 속한 데이터(x_j)	X (15,3)	0.5152	0.4839
	X (15,4)	0.5473	0.4527
	X (15,5)	0.5568	0.4432
	X (15,6)	0.5610	0.4390
	X (15,7)	0.5573	0.4426
	X (15,8)	0.5482	0.4518
	X (15,9)	0.5381	0.4619

제 3 장 평균내부거리를 적용한 퍼지 클러스터링 알고리즘에 의한 영상 분할

평균내부거리를 적용한 클러스터링 방법은 기존의 FCM 알고리즘이 속성 데이터로부터 클러스터의 중심 사이의 거리를 가능한 작게 하는 방향으로 소속정도를 부여하므로 클러스터의 크기가 다른 경우 잘못된 소속정

C1 X1 C2

a b

d1 d2

(a) FCM **알고리즘**(d1=d2)

C1 X2 C2

a b

d1 d2

(b) **제안된 알고리즘**(d1-a=d2-b)

그림 3.1. **두 개의 클러스터를 가지는 데이터 집합 동일한 소속정도를 가지는** FCM**과 제안된 알고리즘**

Fig. 3.1. FCM and proposed algorithm with the same membership grade in data sets with two clusters

도를 가지는 문제점을 개선하고자 하였다. 먼저 평균내부거리 안쪽에 속하는 데이터들의 집합을 내부클러스터라 정의한다. 그리고 각 데이터로부터 내부클러스터까지의 거리에 의존하여 소속정도를 구하므로 클러스터의 크기에 상관없이 균일한 소속정도를 부여하고 각 클러스터의 대표값을 정확하게 찾을 수 있다. 그림 3.1는 이것을 잘 보여주고 있다. C1과 C2는 각 클러스터의 중심, a와 b는 각 클러스터의 평균내부거리, 그리고 d1과 d2는 데이터로부터 클러스터 중심까지의 거리를 나타낸다. 그리고 x1, x2는 동일한 소속정도를 가지는 위치를 보여주고 있다. 그림 3.1의 FCM 알고리즘에서 x1은 큰 클러스터에 가깝게 위치하고 있지만 동일한 소속정도를 부여한다. 이것은 FCM 알고리즘이 특징 벡터로부터 중심까지의 거리에 의해 소속정도를 부여하기 때문이다. 하지만 제안된 알고리즘에서는 평균내부클러스터의 크기를 고려함으로 클러스터의 크기가 다르더라도 균일한 소속정도를 부여할 수 있었다. 제안된 알고리즘의 목적함수는 평균내부거리를 고려한 각 특징 벡터로부터 클러스터 중심까지의 거리를 제곱한 함수로 중심으로부터 데이터에 이르는 분산 정도를 최소화하도록 하는 소속정도를 부여한다. 제안된 목적함수와 평균내부거리는 다음과 같다.

3.1 목적함수

제안된 알고리즘의 목적함수는 각 데이터로부터 각 내부클러스터까지의 거리를 최소로 하는 응집성을 고려하여 데이터 집합을 분할하도록 설계한다. 정의된 목적함수는 식 (3.1)과 같다.

$$J_m(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} (u_{ij})^m \| x_j - v_i \|^2 - \sum_{i=1}^{c} \eta_i \sum_{j=1}^{n} (u_{ij})^m \qquad (3.1)$$

여기서, 모든 j 에 대해 $\sum_{i=1}^{c} u_{ij} = 1$

여기서 η_i는 적당한 양의 정수로 내부클러스터의 크기를 나타내는 값이다. 이에 대한 내용은 다음 절에서 언급하도록 하겠다. 위의 식에서 첫 번째 항은 측정 데이터로부터 클러스터 중심사이의 거리를 가능한 작게 유지하도록 하고, 두 번째 항은 각 내부클러스터의 크기만큼 보정해 주도록 한다. 그러므로 목적함수를 최소화하는 조건을 만족하는 소속정도 u_{ij}를 클러스터의 크기에 따라 균일하게 부여할 수 있도록 하였다. $X = \{x_1, x_2, \cdots, x_n\}$는 특징 벡터의 집합, $V = \{v_1, v_2, \cdots, v_c\}$는 클러스터의 중심 벡터의 집합, $U = [u_{ij}]$는 특징 벡터와 클러스터 중심 사이의 유사성을 나타내는 $c \times n$ 행렬의 소속정도를 나타낸다. 그리고 m은 퍼지 가중치 정도를 나타내며 m이 무한대로 갈수록 소속정도는 최대로 퍼지화된 값을 가지게 된다. m이 1보다 큰 경우에 모든 i, j에 대해서 $v_i \neq x_j$를 만족한다고 가정하면, J_m을 최소화하도록 하는 조건 (U, V)는 다음과 같다.

$$u_{ij} = \{ \sum_{k=1}^{c} (\frac{d_{ij} - \eta_i}{d_{kj} - \eta_k})^{\frac{2}{(m-1)}} \}^{-1} \tag{3.2}$$

여기서, $d_{ij} = \| v_i - x_j \|$, $(\ i = 1, 2, \cdots, c\ ;\ j = 1, 2, \cdots, n\ )$

$$v_i = \frac{\sum_{j=0}^{n} (u_{ij})^m x_j}{\sum_{j=0}^{n} (u_{ij})^m} \tag{3.3}$$

제안된 알고리즘은 식 (3.2)와 식 (3.3)의 과정을 반복하므로 J_m은 어떤 정해진 값에 수렴하게 된다.

3.2 평균내부거리

평균내부거리 η_i는 해당 클러스터의 크기와 밀도에 관한 정보로 클러스터의 크기에 따라 다른 값을 가지게 된다. Krishnapuram[8]은 η_i를 식 (3.4)와 식 (3.5)를 이용하여 추정하였다. 그리고 η_i는 퍼지분할을 먼저 실행하고 퍼지분할로부터 얻은 정보를 가지고 구할 수 있다.

$$\eta_i = K \frac{\sum_{j=1}^{n} (u_{ij})^m d_{ij}^2}{\sum_{j=1}^{n} (u_{ij})^m} \tag{3.4}$$

여기서 K는 보통 1보다 작은 값을 선택하게 된다. 영상은 데이터 개수가 많기 때문에 η_i를 크게 설정하면 발산할 위험이 있다.

$$\eta_i = \frac{\sum_{x_j \in (\Pi_i)_\alpha}^{n} d_{ij}^2}{|(\Pi_i)_\alpha|} \tag{3.5}$$

여기서 Π_i는 특징 벡터가 i 번째 클러스터에 속할 소속정도의 집합을 나타내고 $(\Pi_i)_\alpha$는 적절한 수준의 α- cut이다. 즉 α를 높게 설정함으로 i 번째 클러스터에 속할 가능성이 높은 특징 벡터들의 평균거리가 될 것이다. 그리고 d_{ij}는 i 번째 클러스터의 중심으로부터 j 번째 특징벡터 사이의 거리이다. 식 (3.5)는 적절한 수준의 역치값에 의해 평균내부거리를 구하기 때문에 데이터 집합이 변할 때마다 새로운 역치값을 구해주어야 어려움이 있다. 그래서 본 논문에서는 식 (3.4)를 이용하여 평균내부거리 η_i를 구한다.

3.3 제안된 알고리즘의 수행 절차

제안된 알고리즘의 목적함수를 최소화 하기 위한 조건을 만족하는 중심들의 집합 V와 퍼지분할 U는 다음과 같은 과정에 의해 구해진다.

단계 1: 초기 클러스터 수(c)와 퍼지 가중치 정도(m)를 결정한다.

$$(2 \le c < n \ , \ 1 < m < \infty)$$

단계 2: 다음 조건을 만족하는 퍼지 c 분할 $U^{(0)}$을 초기화한다.

$$\sum_{i=1}^{c} u_{ij} = 1 \ , \ 0 < \sum_{j=1}^{n} u_{ij} < n \ ; \ 1 \le i \le c, 1 \le j \le n$$

단계 3: 각 클러스터에 대한 중심값 $v_i^{(l)}$을 구한다. $(l = 0, 1, 2, \cdots)$

$$v_i^{(l)} = \frac{\sum_{j=0}^{n} u_{ij}^{(l)m} x_k}{\sum_{j=0}^{n} u_{ij}^{(l)m}}$$

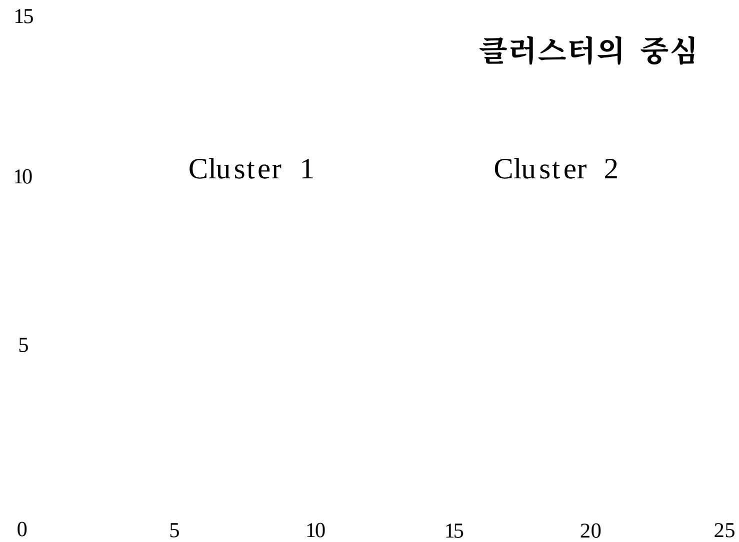
단계 4: 평균내부거리 η_i를 구한다.

$$\eta_i = K \frac{\sum_{j=1}^{n} (u_{ij})^m d_{ij}^2}{\sum_{j=1}^{n} (u_{ij})^m}$$

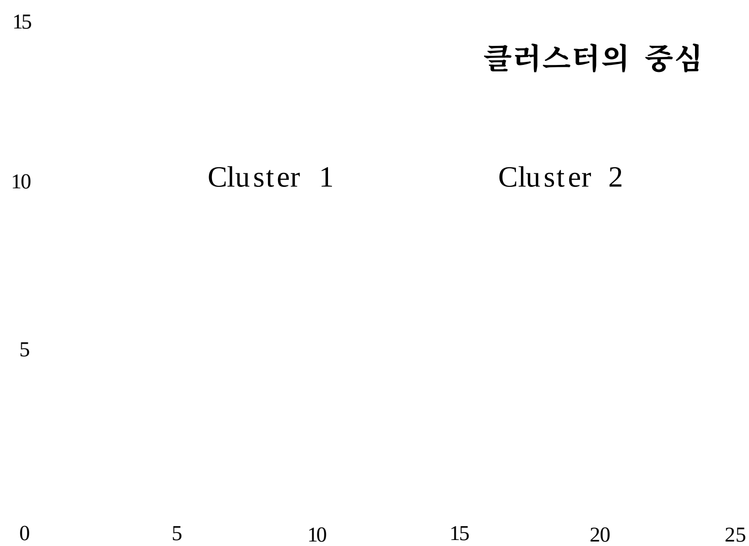
단계 5: 퍼지 c 분할 $U^{(l+1)}$ 을 계산한다. 단, $\eta_i \neq d_{ij}$ 일 때

단계 6: 만약 $\| U^{(l+1)} - U^{(l)} \| \le \varepsilon$ 이면 알고리즘을 종료하고 그렇지 않으면 단계 3으로 이동하여 반복 수행한다.

제안된 알고리즘이 클러스터의 크기가 다른 경우 FCM보다 우수한 성능을 가지고 있음을 보이기 위해 FCM에서 사용한 예제와 동일한 데이터 집합에 대해 적용한다. 그림 3.2은 제안된 알고리즘에 의한 중심탐색 결과를 보여주고 있다. FCM 알고리즘은 데이터로부터 클러스터 중심에 이르는 거리의 제곱 합을 최소화하는 조건에 의해 소속정도를 결정하였다. 그러나 제안된 알고리즘에서는 데이터로부터 내부클러스터까지의 거리에 의해 소속정도를 부여하므로 크기에 상관없이 균일한 소속정도를 부여하고 각 클러스터의 대표값을 잘 찾는다. 그림 3.2에서 Cluster 1은 큰 클러스터의 영향을 받지 않고 중심값을 잘 찾고 있음을 알 수 있다. 또한 FCM에서는 경계영역에 있는 데이터에 대해 오분류를 하고 있었지만 제안된 알고리즘에서는 각각의 클러스터를 잘 분류하고 있음을 알 수가 있다. 표 3.1은 경계 영역에 속한 데이터의 소속정도를 보여주고 있다. FCM에서는 경계영역에 속한 데이터를 Cluster 1에 많은 소속정도를 주었지만 제안된 알고리즘에서는 크기에 상관없이 균일하게 소속정도를 부여하고 있음을 알 수가 있다.



(a) 클러스터의 크기가 같은 경우(데이터 집합 A)



(b) 클러스터의 크기가 다른 경우(데이터 집합 B)

그림 3.2. 두 개의 클러스터를 가지는 데이터 집합

Fig. 3.2. Data sets with two clusters

표 3.1. **경계영역에 속한 데이터가 각 클러스터에 속할 소속정도(집합** A)

Table 3.1. Membership grades belong to the each clusters of the boundary data(Data set A)

	좌 표	Cluster 1에 속할 소속정도(u_{ij})	Cluster 2에 속할 소속정도(u_{ij})
경계영역에 속한 데이터(x_j)	X (15,3)	0.3412	0.6588
	X (15,4)	0.3153	0.6847
	X (15,5)	0.2949	0.7051
	X (15,6)	0.2869	0.7131
	X (15,7)	0.2949	0.7051
	X (15,8)	0.3153	0.6847
	X (15,9)	0.3412	0.6588

표 3.2. **경계영역에 속한 데이터가 각 클러스터에 속할 소속정도(집합** B)

Table 3.2. Membership grades belong to the each clusters of the boundary data(Data set B)

	좌 표	Cluster 1에 속할 소속정도(u_{ij})	Cluster 2에 속할 소속정도(u_{ij})
경계영역에 속한 데이터(x_j)	X (15,3)	0.4143	0.5857
	X (15,4)	0.3979	0.6021
	X (15,5)	0.3846	0.6154
	X (15,6)	0.3793	0.6207
	X (15,7)	0.3846	0.6154
	X (15,8)	0.3979	0.6021
	X (15,9)	0.4143	0.5857

제 4 장 실험 및 성능평가

본 논문에서는 실험을 위하여 디지털 카메라로부터 획득한 100×100의 크기를 가지는 영상들을 이용하였다. 그리고 각 클러스터의 크기가 서로 다른 데이터 집합 C을 만들어 평균내부거리를 이용한 클러스터링 방법이 클러스터의 크기가 다른 경우 기존의 FCM 알고리즘보다 우수함을 성능평가 함수로 보였다.

퍼지 클러스터링의 성능을 평가하기 위하여 퍼지분할 후 각 클러스터와 중심과의 관계를 나타내는 분류 엔트로피(Classification Entropy)와 중심들간의 관계를 나타내는 응집성(Compactness)과 분할성(Separation)에 관한 적합도 함수를 이용하였다. 분류 엔트로피는 각 특징벡터들이 클러스터의 중심에 잘 분류될수록, 즉 애매한 특징벡터의 개수가 적을수록 낮은 값을 가지게 된다. 그리고 응집성과 분할성에 관한 적합도 함수는 중심으로부터 특징벡터까지의 거리를 작게 하여 응집성을 높이고, 중심들간의 거리를 크게 하여 분리성을 높이도록 설계하므로 클러스터의 분류가 잘 될수록 낮은 값을 가지게 된다. 그러므로 최적의 퍼지분할은 적합도 함수의 값이 최소일 때이고 그때의 분할 개수가 최적의 분할 개수가 될 것이다. 하지만 퍼지 클러스터링이 가지는 확률적 제약조건 때문에 명확하게 최적의 분할 개수라고 할 수는 없을 것이다. 식 (4.1)은 Bezdek[11]에 의해 제안된 분류 엔트로피(CE)이고 식 (4.2)는 Xie[12]에 의해 제안된 응집성과 분할성에 관한 적합도 함수(S)이다.

$$CE = -\frac{1}{k}\left[\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{i=1}^{c}\{u_{mn}^{i}\log(u_{mn}^{i})\}\right] \tag{4.1}$$

$$S = \frac{\sum_{i=1}^{c}\sum_{m=1}^{M}\sum_{n=1}^{N}(u^{i}_{mn})^{2}\| x_{mn} - v_i \|^{2}}{k[\min_{i \neq j}(\| v_i - v_j \|)]} \qquad (4.2)$$

여기서 k은 특징벡터의 개수 $M \times N$이고, u^{i}_{mn}는 퍼지분할 후의 소속정도이다. 그리고 $\frac{1}{k}\sum_{i=1}^{c}\sum_{m=1}^{M}\sum_{n=1}^{N}(u^{i}_{mn})^{2}\| x_{mn} - v_i \|^{2}$는 응집성(Compactness)을 나타내고, $(\min_{i \neq j}[\| v_i - v_j \|])$는 분할성(Separation)을 나타낸다. 이러한 적합도 함수는 이웃한 점들에 대한 균일성(Uniformity)과 동질성(Homogeneity)을 보여주기도 한다.

4.1 퍼지 클러스터링에 의한 데이터 집합 분류

데이터 집합 C는 2차원 벡터이고 3개의 클러스터로 이루어진 집합이다. 그림 4.1은 응집성만을 고려하여 설계된 FCM의 퍼지분할을 보여주고 있다. 여기서 상대적으로 작은 첫 번째 클러스터를 두 번째 클러스터로 분류하고 세 번째 클러스터를 두 개로 나누어 중심 탐색을 하고 있다. 그러므로 FCM 알고리즘은 각 클러스터의 대표값을 잘못 찾을 뿐 아니라 데이터를 오분류하고 있다. 그림 4.2는 응집성과 클러스터의 크기에 따라 소속 정도를 보정해 줄 수 있도록 설계된 제안된 알고리즘의 퍼지분할 결과를 보여주고 있다. 그림에서 알 수 있듯이 각 클러스터의 대표값을 잘 찾을 뿐 아니라 데이터 분류도 명확하게 하고 있다.

표 4.1은 데이터 집합 C를 제안된 방법과 기존의 클러스터링 방법과의 성능을 비교하고 있다. 애매한 데이터의 개수가 적을수록 작은 값을 가지는 분류 엔트로피와 분류된 데이터 집합의 응집성과 분할성에 관한 적합도 함수의 값이 기존의 알고리즘보다 제안된 알고리즘이 더 적게 나옴을 보여주고 있다.

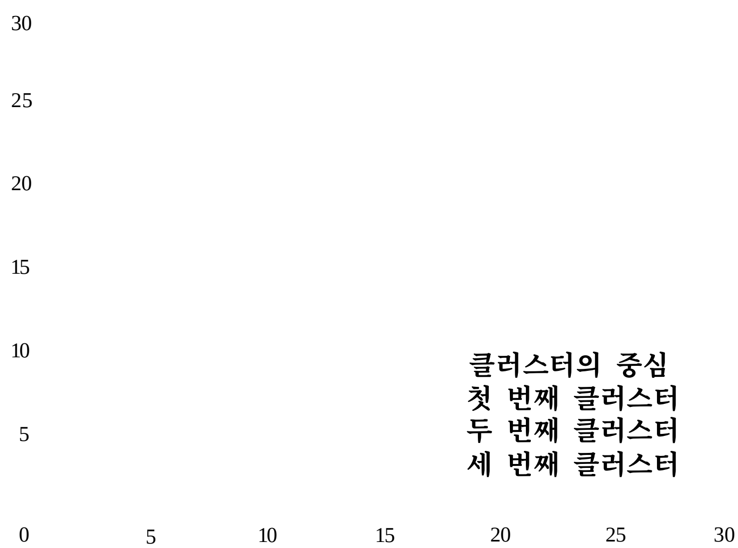


그림 4.1. FCM **방법에 의한 실험 데이터 집합**C**의 중심**

Fig. 4.1. The center of experimental data set C by FCM method

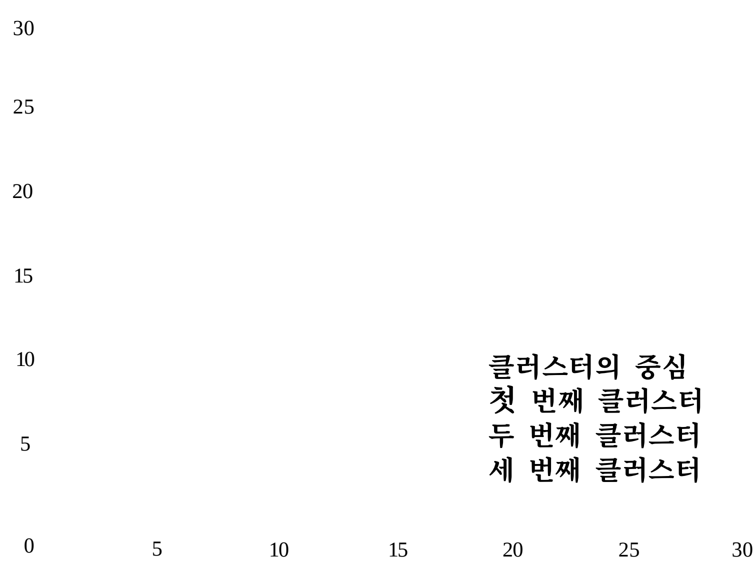


그림 4.2. **제안된 방법에 의한 실험 데이터 집합**C**의 중심**

Fig. 4.2. The center of experimental data set C by the proposed method

표 4.1. 각 클러스터링 기법의 성능 비교(집합 C)
Table 4.1. Comparison of criteria for the each clusters(Data set C)

	FCM	제안된 방법
클러스터1의 중심	(6.0428,21.1668)	(6.0891,9.0729)
클러스터2의 중심	(16.9771,18.3205)	(6.4278,21.9989)
클러스터3의 중심	(20.9085,18.7888)	(20.1533,18.9993)
CE	0.2776	0.1913
S	0.3798	0.2745

4.2 퍼지 클러스터링에 의한 영상분할

영상 분할 결과를 평가하기 위한 균일한 평가 기준이 없기 때문에 대부분의 영상 분할 결과는 사람의 시각에 의해 평가된다. 그리고 영상분할의 결과들은 다음과 같은 조건에 의해 평가되어 진다. 첫째, 분할되어진 영상의 영역들은 명도나 질감(Texture) 같은 속성에 관해 균일성과 동질성을 가져야 한다. 둘째, 영역들은 간단해야 되고, 영역 내부에 많은 구멍(hole)들이 포함되지 말아야 된다. 셋째, 인접한 영역들은 현저히 다른 특성을 가져야 한다. 넷째, 각 영역의 경계들은 간단해야 되고 매끄러워야 한다. 그러나 위의 여러 가지 조건을 만족하는 평가 척도를 얻는 것은 어렵다. 따라서 영상 분할에 있어서 중대한 문제점은 분할 결과를 평가할 수 있는 균일한 평가 기준이 없다는 점이다. 물론, 수작업에 의해 분할된 결과에 근거해 오분류률을 계산하기 위한 여러 시도가 있었다. 그러나 이 방법은 수작업에 의해 분할된 결과가 정확하다는 확신을 가질 수가 없다. 그러므로 이러한 방법은 영상의 영역 정보를 알 수가 있는 합성된 영상에서만 가능하고 실영상에서는 이용하기 어려운 문제점이 있다.

본 논문에서는 분할된 영상의 결과를 평가하기 위하여 분류 성능을 나

타내는 CE와 분할된 영역의 응집성과 분할성을 나타내는 S-function을 사용하여 각 클러스터링 방법을 평가하였다. 그리고 CE와 S-function은 모두 작은 값을 가질수록 군집화 성능이 우수하므로 두 평가함수의 곱으로 나타나는 새로운 적합도 함수 $F = CE \times S$[3]를 정의하였다.

그림 4.3은 디지털 카메라로 획득한 영상 1을 제안된 알고리즘과 기존의 FCM 알고리즘에 적용한 적합도 함수의 값을 보여주고 있다. 초기 중심값은 임의로 선택하였으며 퍼지가중치 m은 2로 하고 클러스터의 개수는 2개에서 7개까지 변화하면서 F 함수를 구하였다. 그림에서 알 수 있듯이 적합도 함수의 값은 제안된 알고리즘이 기존의 FCM 알고리즘보다 적게 나온다. 이것은 제안된 방법이 영상분할을 잘하고 있음을 말해준다. 그리고 그림 4.4는 제안된 알고리즘에 의해 영상 1을 영상분할한 결과이다. (a)는 원영상을, (b)는 히스토그램을, (c), (d), (e)는 순서대로 3개, 4개, 5개로 영상분할된 결과를 보여주고 있다. 그리고 F 함수가 최소값을 가지는 클러스터의 개수가 3일 때 분할된 영역이 사람의 시각적 평가 조건에 가장 일치함을 알 수 있었다.

그림 4.5는 영상 2에 대해 F 함수의 실험 결과를 보여주고 있다. 여기서 FCM 알고리즘은 3개에서 최적의 분할 개수를 가지고 제안된 방법은 4개에서 최적의 분할 개수를 가짐을 알 수 있었다. 그림 4.6은 영상 2를 각 알고리즘에 대해 최적의 분할 개수로 영상분할한 결과를 보여주고 있다. (a)는 원영상이고 (b)는 3개의 영역으로 분할된 FCM 알고리즘의 결과이다. 그리고 (c)는 4개의 영역으로 분할된 제안된 알고리즘의 결과이다. 제안된 방법이 기존의 방법보다 최적의 분할 개수를 잘 찾고 있음을 알 수가 있다.

그림 4.7은 명확하게 4개의 영역으로 구분되면서도 한 영역이 다른 영역에 비해 아주 작은 영상 3에 대한 영상분할을 보여주고 있다. (c)는

FCM에 의해 영상분할된 결과로 상대적으로 작은 영역을 찾지 못하고 있다. 그러나 제안된 알고리즘에서는 4개의 영역에 대한 분할이 잘 이루어지고 있음을 알 수 있다.

그림 4.8과 그림 4.9는 각각 영상 2과 가우시안 잡음이 15% 추가된 영상 2에 대한 영상분할 결과를 보여주고 있다. 그림 4.8 (c)의 FCM 알고리즘에 의해 분할된 영상을 보면 상대적으로 작은 정보를 가진 영역을 손실하고 있다. 이것은 각 특징 벡터로부터 클러스터 중심까지의 거리를 최소로 하는 목적함수에 의해 영상분할이 이루어지기 때문이다. 하지만 (d)의 제안된 알고리즘은 각 클러스터의 크기를 보정해 주는 목적함수를 가지고 영상분할을 하기 때문에 아주 작은 특징 정보에 대해서도 영상분할을 잘 하고 있다. 또한 그림 4.9에서 제안된 알고리즘은 잡음이 추가된 경우에도 작은 영역에 대해 기존의 알고리즘보다 영상분할 성능이 우수함을 알 수 있다. 그리고 그림 4.10에서 알 수 있듯이 잡음이 추가된 경우에는 적합도 함수 값이 원영상에서 보다 높아지게 된다. 즉 영상분할 성능이 저하됨을 알 수 있었다.

7
6
5
4
3
2
1
0

2 3 4 5 6 7 8

FCM 알고리즘
제안된 알고리즘

그림 4.3. **영상** 1**의** F-**함수**

Fig. 4.3. F-function of image 1

(a) **영상** 1 (b) **영상** 1**의 히스토그램**

(c) 3**개의 영역** (d) 4**개의 영역** (e) 5**개의 영역**

그림 4.4. **제안된 알고리즘에 의한 영상** 1**의 영상분할**

Fig. 4.4. Segmentation of image 1 for the proposed algorithm

2.5

FCM 알고리즘
제안된 알고리즘

2

1.5

1

0.5

0

2 3 4 5 6 7 8

그림 4.5. 영상 2의 F-함수

Fig. 4.5. F-function of image 2

(a) 영상 2

(b) 히스토그램

(c) FCM 방법에 의해 분할된 영상(3개의 영역)

d) 제안된 방법에 의해 분할된 영상(4개의영역)

그림 4.6. 최적의 분할 개수에 의한 영상 2의 영상 분할

Fig. 4.6. Segmentation of image 2 with compatible partition number

(a) **영상** 3

(b) **영상** 3**의 히스토그램**

(c) FCM **방법에 의해 분할된 결과**

(d) FCM **방법에 의해 분할된 결과의 히스토그램**

(e) **제안된 방법에 의해 분할된 결과**

(f) **제안된 방법에 의해 분할된 결과의 히스토그램**

그림 4.7. 4**개의 영역을 가진 영상** 3**의 영상 분할**

Fig. 4.7. Segmentation of image 3 with 4-partitions

(a) **영상** 2

(b) **영상** 2**의 히스토그램**

(c) FCM **방법에 의해 분할된 결과**

(d) **제안된 방법에 의해 분할된 결과**

그림 4.8. 4**개의 영역으로 분할된 영상** 2

Fig. 4.8. Segmentation of image 2 with 4-partitions

(a) 잡음이 추가된 영상 2

(b) 잡음이 추가된 영상 2의 히스토그램

(c) FCM 방법에 의해 분할된 결과

(d) 제안된 방법에 의해 분할된 결과

그림 4.9. 4개의 영역으로 분할된 잡음을 추가한 영상 2의 영상 분할 (15%의 가우시안 잡음 추가)

Fig. 4.9. Segmentation of noisy image 2 with 4-partitions (15% gaussian noise)

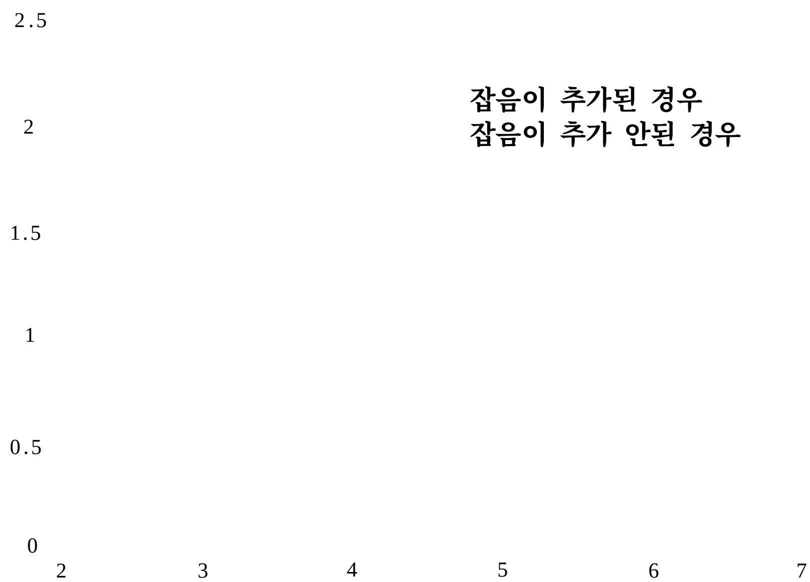


그림 4.10. **제안된 알고리즘에 대한 영상** 2**와 잡음이 추가된 영상** 2**의** F **함수**

Fig. 4.10. F-function of image 2 and noisy image 2 for the proposed method(15% gaussian noise)

제 5 장 결론

본 논문에서는 평균내부거리를 이용한 퍼지 클러스터링 알고리즘을 제안하였다. 평균내부거리는 각 특징 벡터들로부터 해당 클러스터 중심까지의 거리를 평균한 값으로 클러스터의 크기와 밀도에 비례한다. 제안된 알고리즘은 특징 벡터로부터 클러스터 중심까지의 거리가 평균내부거리보다 작은 데이터들의 집합을 내부클러스터라 정의하고, 특징 벡터로부터 내부클러스터까지의 거리를 최소로 하는 목적함수를 설계하므로 클러스터의 크기에 상관없이 균일한 군집화를 할 수가 있다.

기존의 FCM 알고리즘은 각 특징 벡터로부터 클러스터 중심까지의 거리을 최소로 하는 목적함수를 가지기 때문에 클러스터의 크기가 다른 경우 잘못된 소속정도를 가진다. 그러므로 클러스터를 대표하는 중심값을 잘못 찾고 특징 벡터들을 오분류하게 된다. 제안된 알고리즘은 이러한 문제점을 개선할 뿐 아니라

제안된 알고리즘을 클러스터의 크기가 다른 2차원 데이터 집합과 영상에 적용하였다. 그리고 작은 값을 가질수록 분류성능이 우수한 적합도 함수를 이용하여 기존의 알고리즘과 비교해 보았을 때 우수함을 보였다. 하지만 퍼지 클러스터링 방법은 화소들이 가지는 특징 정보만을 이용하여 특징 공간상에서 클러스터링을 하기 때문에 잡음이 추가된 경우 분할 성능이 저하됨을 알 수 있었다. 이러한 문제점은 영역 확장법등의 후처리 과정을 통해 해결할 수 있을 것이다.

참고 문헌

[1] Y. W. Lim and S. U. Lee, "On the Color Image Segmentation Algorithm Based on the Thresholding and the Fuzzy C-Means Techniques," *Pattern Recognition*, Vol. 23, No. 9, pp.935-952, 1990.

[2] H. Rhee and K. Oh, "A Design and Analysis of Objective Function Based Unsupervised Neural Networks of Fuzzy Clustering," *Neural Processing Letters*, Vol. 4, pp. 83-95, 1996.

[3] D. N. Chun, "A Method of Clustering and Image Segmentation Based on Fuzzy Genetic Algorithm," Department of Computer Science. KAIST, pp. 26-34, 1996.

[4] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, Addison Wesley, pp. 443-457. 1992.

[5] J. C. Bezdek and M. M. Rrivedi, "Low Level Segmentation of Aerial Images with Fuzzy Clustering," *IEEE Transaction on System*, Vol. SMC-16, No. 4, pp. 589-598, 1986.

[6] R. Krishnapuram and J. M. Keller, "A Possibilistic Approach to Clustering," *IEEE Transaction on Fuzzy Systems*, Vol. 1, No. 2, pp. 98-110, 1993.

[7] James C. Bezdek, "A Convergence Theorem for the Fuzzy ISODATA clustering algorithms," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 1, pp. 1-8, 1980.

[8] R. Krishnapuram, H. Frigui and O. Nasraoui, " Fuzzy and

Possibilistic Shell Clustering Algorithms and Their Application to Boundary Detection and Surface Approximation," *IEEE Transaction on Fuzzy Systems*, Vol. 3, No. 1, pp. 29-60. 1995.

[9] A. P. Witkin, "Scale-Space Filtering," Processing IJCAI-83, pp. 1019-1022, 1983.

[10] **유 현재**, **조 석제**, " **평균내부거리를 적용한 퍼지 클러스터링**," 99**년 추계 정보처리 학회**, pp. 172-176, 1999.

[11] N. Pal and J. Bezdek, "On Cluster Validity for the Fuzzy C-Means Model," *IEEE Transation on Fuzzy System*, Vol. 3, No. 3, 1995.

[12] X. Xie and G. Beni, "A Validity Measure for Fuzzy Clustering," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 8, pp. 841-847, 1991.

[13] G. P. Babu and M. N. Murty, "Clustering with Evolution Strategies," *Pattern Recognition*, Vol. 27, No. 2, pp. 321-329, 1994.

[14] H. D. Cheng and J. R. Chen, "Threshold Selection based on Fuzzy C-Partition Entropy Approach," *Pattern Recognition*, Vol. 31, No. 7, pp. 857-870, 1998.

[15] P. K. Sahoo and D. W. Slaaf, "Threshold Selection using a Minimal Histogram Entropy Difference," *Society of Photo-Optical Instrumentation Engineers*, Vol. 36, No. 7, 1997.

本 論文을 兪鉉在의 工學碩士 學位論文으로 認准함

委員長　　工學博士　劉　永　昊　　印

委　員　　工學博士　陳　康　奎　　印

委　員　　工學博士　趙　爽　濟　　印

1999 年　12 月 29 日

韓國海洋大學校　大學院

制御計測工學科　兪 鉉 在

감사의 글

영상처리 연구실에서 공부한 시간들의 결실로 이렇게 졸업논문을 내놓게 되었습니다.