Journal of the Society of Korea Industrial and Systems Engineering
Vol, 31, No. 4, pp.140－145, December 2008.

두 점과 분할 카디날리티가 주어진 퍼지 균등화조건을 갖는 퍼지분할

김경택*† ・김종수*・강성열**

*한남대학교 공과대학 산업경영공학과
**홍익대학교 상경대학 상경학부

Fuzzy Partitioning with Fuzzy Equalization Given Two Points and Partition Cardinality

Kyeongtaek Kim*† ・Chong Su Kim*・Sungyeol Kang**

*Department of Industrial and Management Engineering, Hannam University
**College of Business Administration, Hongik University

Fuzzy partition is a conceptual vehicle that encapsulates data into information granules. Fuzzy equalization concerns a process of building information granules that are semantically and experimentally meaningful. A few algorithms generating fuzzy partitions with fuzzy equalization have been suggested. Simulations and experiments have showed that fuzzy partition representing more characteristics of given input distribution usually produces meaningful results. In this paper, given two points and cardinality of fuzzy partition, we prove that it is not true that there always exists a fuzzy partition with fuzzy equalization in which two of points having peaks fall on the given two points. Then, we establish an algorithm that minimizes the maximum distance between given two points and adjacent points having peaks in the partition. A numerical example is presented to show the validity of the suggested algorithm.

Keywords : Fuzzy Partitions, Fuzzy Equalization

1. 서 론

데이터 마이닝은 통계학, 신경망이론, 퍼지집합, 데이터베이스, 데이터시각화 등 여러 분야의 기법들이 사용되고 있다. 사용 기법의 다양성에도 불구하고, 그 초점은 의미 있고 이해하기 쉬운 패턴의 발견에 있다. 주어진 데이터를 너무 정밀하게 모형화 하면, 그 의미를 잃게 되는 모순에 빠지게 된다[13]. 따라서 모형이 의미를 갖도록 정밀도를 조절하여야 한다. 정밀도의 조절은 정보 알갱이(Information Granule)란 개념으로 잘 설명된다. 정보 알갱이란 데이터의 유사성, 데이터 처리의 응집성(Cohesion) 또는 의미적 연관성에 따른 데이터의 모음을 가리킨다[5, 11, 12, 15]. 정보 알갱이를 크게 하면 모형의 정밀도가 낮아지며, 정보알갱이를 작게 하면 모형의 정밀도가 높아진다. 데이터를 정보 알갱이화 함으로써 보이지 않던 현상을 보다 잘 이해할 수 있게 되며, 처리가 보다 효율적으로 진행될 수 있다. 따라서 다수의 데이터를 처리해야 하는 데이터 마이닝에서 적절한 크

논문접수일 : 2008년 09월 04일 논문수정일(1차 : 11월 24일, 2차 : 12월 10일) 게재확정일 : 2008년 12월 11일
† 교신저자 kkim@hnu.kr
※ 본 논문은 2006년 한남대학교 학술연구조성비 지원에 의하여 연구되었음.

기로의 정보 알갱이화는 의미 있고 이해하기 쉬운 패턴 발견이라는 목적을 보다 빠른 시간 내에 달성하게 한다. 데이터를 정보 알갱이화를 하여 처리하는 개념적 수단으로 일반집합, 퍼지집합, 러프집합, 랜덤집합, 확률 등이 있다[5].

연속형 데이터를 보다 효율적으로 처리하기 위하여 데이터 마이닝에서 퍼지집합의 사용이 빈번해지고 있다. 이는 주어진 데이터를 여러 개의 퍼지집합으로 분할함으로써 정보 알갱이를 형성할 수 있기 때문이다. 이 경우, 소속함수가 퍼지모델의 성능에 심대한 영향을 끼칠 수 있기 때문에, 소속함수의 결정이 중요하게 된다[2, 14]. 정의역의 각 점에 대한 각 퍼지집합의 소속함수 값의 합이 1이 되는 관계가 있는 퍼지집합의 집합(Set of Fuzzy Sets)을 퍼지분할(Fuzzy Partition)이라 부른다. 퍼지분할에 속한 각 퍼지집합의 기대값이 동일할 때, 이를 퍼지균등화(Fuzzy Equalization)를 만족한다고 한다. Pedrycz[9]는 퍼지 균등화를 만족시키는 퍼지분할을 생성하는 알고리즘을 최초로 제시하였다. 이 퍼지분할에서 양쪽 끝 구간에 해당하는 퍼지집합은 사다리꼴(Trapezoidal) 퍼지숫자이며, 나머지 퍼지집합은 삼각(Triangular) 퍼지숫자의 형태를 취한다. 이 알고리즘은 복잡하지 않은 장점이 있는 반면, 동일한 입력 분포를 가지는 데이터에 대하여, 알고리즘에 의해 생성되는 퍼지분할은 항상 동일하다는 제한이 있다. 퍼지 분할이 데이터마이닝의 한 과정이라 할 때, 만일 결과로서 얻어진 퍼지분할이 더 이상 도움이 되지 않는다면 새로운 퍼지분할이 필요하게 된다. 김경택 외[1]는 동일한 분포를 가지는 데이터에 대하여 서로 다른 퍼지분할을 생성하는 알고리듬을 제시하여 이러한 단점을 해결하였다. 이 알고리듬은, 정의역내의 주어진 한 점과, 퍼지집합에서 피크치를 갖는 점이 오차의 범위 내에서 일치하는 퍼지분할을 생성할 수 있다.

본 논문에서는 이 알고리듬을 확장하여, 정의역 내의 두 점과 퍼지분할에 속한 퍼지집합의 수가 주어졌을 때, 주어진 두 점과 이 점에 가장 가까운 피크치를 갖는 퍼지분할을 생성하는 알고리듬을 제시하고, 그 적용 예를 제시한다.

2. 퍼지균등화

수치데이터로부터 퍼지집합을 이용하여 형성된 정보 알갱이는 의미론적(Semantical) 의미와, 실험적(Experimental) 의미를 가져야 한다[9]. 정보 알갱이를 표현하는 퍼지집합과 그들의 집합, 즉 퍼지집합들의 집합의 의미론적 의미에 관해서는 오랫동안 충분하게 논의되어 왔다[4, 5, 7, 8, 10, 11, 13, 15]. 퍼지집합의 실험적 의미는 이 퍼지집합과 관련된 데이터에 달려있다. 정의역 X에서 연속적으로 분포하는 확률밀도함수 $f(x)$를 갖는 데이터가 존재하고, 이 정의역에서 정의된 퍼지 집합 A_i가 있다고 가정하자. 그러면 퍼지집합의 지지구간에 대한 누적확률을 다음과 같이 정의된다.

$$P(A_i) = \int_x A_i(x) f(x) dx$$

이 $P(A_i)$는 퍼지 집합 A_i의 기대 값이다. 만일 퍼지집합의 확률 $P(A_i)$가 어떤 임계값과 같거나 크면, 그 퍼지집합은 의미 있는 것으로 볼 수 있다. 이러한 관계가 성립할 때, 퍼지집합 A_i는 실험적인 의미를 가진다[6].

퍼지집합이 응용되는 문제의 대부분에서는 하나의 퍼지집합이 사용되기보다는 퍼지집합들의 집합(Set of Fuzzy Sets)이 사용된다. 퍼지집합들의 집합 $\mathbf{A} = \{A_1, A_2, \cdots, A_n\}$는 동일한 X에서 정의 되는 퍼지집합의 모음이라 하자. 만일 X의 어떤 원소에 대하여서도 각 퍼지집합의 소속함수 값의 합이 1일 때, 즉,

$$\forall x \in X \quad \sum_{i=0}^{n} A_i(x) = 1$$

이 성립할 때, 퍼지집합들의 집합 $\mathbf{A}$를 퍼지분할이라 부른다. 그러면, 퍼지분할 $\mathbf{A}$에 대하여 다음의 관계가 성립한다.

$$\sum_{i=1}^{n} P(A_i) = \int A_1(x) f(x) dx + \cdots + \int A_n(x) f(x) dx \\ = 1$$

일반적으로 퍼지분할은 전문가의 지식에 의거하여 이루어진다. 그러나 그러한 전문가가 없거나, 또는 전문지식이 아직 축적되지 않은 분야의 경우, 전문지식에 의한 퍼지분할이 불가능 하다. 이러한 경우, 퍼지분할은 임의로 이루어질 수 밖에 없다. 이때, 각 퍼지집합의 기대값이 동일하도록 퍼지분할을 하는 방법이 임의로 퍼지분할을 하는 방법보다 더 설득력을 가질 수 있다. 이처럼, 각 퍼지집합의 기대값이 동일하도록 퍼지분할을 할 때, 이를 퍼지균등화라 하며 다음 관계가 성립한다.

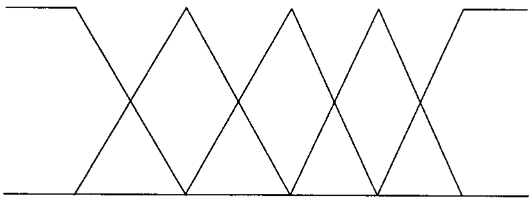
$$P(A_1) = P(A_2) = \cdots = P(A_n) = 1/n$$

이 조건을 만족하는 각 퍼지집합은 동일한 양의 실험적 증거(Experimental Evidence)를 가지고 있음을 나타낸다[9]. 또한, 이 조건을 만족하면, 확률밀도함수 값이 큰 구간에

서 퍼지집합의 지지 구간은 작아지고, 확률밀도함수 값이 작은 구간에서 퍼지집합의 지지 구간은 넓어진다.

3. 알고리듬

소속함수로 가장 일반적으로 사용되고 있는 퍼지 삼각 소속함수와 사다리꼴 소속함수는 다루기 쉽고 계산이 용이하다고 알려져 있다[2, 3]. 따라서 <그림 1>과 같이, 퍼지분할의 양 끝은 사다리꼴 퍼지함수이고, 나머지는 모두 삼각퍼지함수인 이단퍼지분할 형태를 취하기로 한다.



<그림 1> 이단 퍼지분할

본 논문에서는 입력데이터의 확률밀도함수 $f(x)$ (단, 모든 정의역에서 $f(x)>0$), P_1(정의역내의 한 점), P_2(정의역내의 한 점, 단 $P_1<P_2$), n(퍼지분할의 카디날리티)이 주어졌다고 가정한다. 그러면, 다음 관계가 성립한다.

$$\begin{aligned}\int_{a_0}^{a_j} f(x)ds &= \sum_{i=0}^{j-2}\int_{a_i}^{a_{i+2}} A_{i+1}(x)f(x)dx \\ &\quad + \int_{a_{j-1}}^{a_j} A_j(x)f(x)dx \\ &= \frac{j-1}{n} + \int_{a_{j-1}}^{a_j} A_j(x)f(x)dx\end{aligned}$$

따라서,

$$\frac{j-1}{n} < \int_{a_0}^{a_j} f(x)d(x) \le \frac{j}{n}$$

한편, $\frac{j-1}{n} < \int_{a_0}^{P_1} f(x)d(x) \le \frac{j}{n}$로부터, a_0, P_1, $f(x)$, n을 알고 있으므로, 정수 j를 구할 수 있다. 여기에서, 정수 j는, P_1에 일치할 수 있는 퍼지분할의 피크 (a_j)가 퍼지분할에 속한 몇 번째 퍼지 집합의 피크인지를 나타낸다. 동일한 방법으로 $\frac{k-1}{n} < \int_{a_0}^{P_2} f(x)d(x) \le \frac{k}{n}$로부터, a_0, P_2, $f(x)$, n을 알고 있으므로, 정수 k를 구할 수 있다. 여기에서, 정수 k는, P_2와 일치할 가능성을 가진 피크는 퍼지분할의 몇 번째 퍼지 집합의 피크인지를 나타낸다.

이제 $a_j=P_1$과 $a_k=P_2$가 되는 퍼지분할을 만들기로 하자. 먼저, $a_j=P_1$이 되는 퍼지 분할을 찾은 다음, a_j의 위치를 고정한 후, 다른 피크치들을 움직여, $a_k=P_2$도 만족 시키는 퍼지분할을 만들 수 있는 지를 살펴보기로 한다.

$a_j=P_1$이며, 퍼지균등화 조건을 만족시키는 이단퍼지분할을 $\boldsymbol{A}=\{A_1, A_2, \cdots, A_n\}$라 하자. $b_j=P_1$이고 $b_k=P_2$이며, 퍼지균등화 조건을 만족시키는 이단퍼지분할을 $\boldsymbol{B}=\{B_1, B_2, \cdots, B_n\}$라 하자. 그러면, 다음 두 식이 성립한다.

$$\begin{aligned}\int_{a_0}^{a_j} f(x)dx &= \frac{j-1}{n} + \int_{a_{j-1}}^{a_j} A_j(x)f(x)dx \\ &= \frac{j-1}{n} + \int_{a_{j-1}}^{a_j} \frac{x-a_{j-1}}{a_j-a_{j-1}} f(x)dx\end{aligned}$$

$$\begin{aligned}\int_{b_0}^{b_j} f(x)dx &= \frac{j-1}{n} + \int_{b_{j-1}}^{b_j} B_j(x)f(x)dx \\ &= \frac{j-1}{n} + \int_{b_{j-1}}^{b_j} \frac{x-b_{j-1}}{b_j-b_{j-1}} f(x)dx\end{aligned}$$

$a_j=b_j=P_1$이므로, 위 두식은 동일한 값을 가져야 한다. 그러기 위해서는, $a_{j-1}=b_{j-1}$이 된다. 동일한 방법을 적용하면, 모든 i에 대하여 $a_i=b_i$가 성립함을 쉽게 증명할 수 있다. 그러므로 P_1과 일치하는 피크치를 가지며, 퍼지균등화 조건을 갖는 이단퍼지분할은 오직 하나만 존재함을 알 수 있다. 따라서 $a_j=P_1$이며, 퍼지균등화 조건을 갖는 이단퍼지분할 $\boldsymbol{A}$에서 $a_k=P_2$가 성립하지 않는다면, $a_j=P_1$과 $a_k=P_2$를 동시에 만족시키는 퍼지균등화 조건을 갖는 이단퍼지분할은 존재하지 않는다.

퍼지균등화 조건을 유지하면서, 주어진 두 피크치와 가장 가까운 피크치를 갖는 퍼지분할을 찾아보자. 이 경우, "가장 가깝다"는 여러 가지 방법에 의해 정의 될 수 있으며,

- 거리의 합의 최소
- 거리의 제곱의 합의 최소
- 거리의 최대치의 최소

등이 포함될 수 있다. 이들 방법사이의 우열은 없으며, 본 논문에서는 계산상의 편의를 고려하여 거리의 최대치를 최소화하는 방법을 택하였다.

동일 확률밀도함수에서 정의된, 퍼지균등화 조건을 갖는 이단퍼지분할 $\boldsymbol{A}=\{A_1, A_2, \cdots, A_n\}$, $\boldsymbol{B}=\{B_1, B_2, \cdots, B_n\}$는 다음의 관계를 갖는다[1].

$$\int_{a_{i-1}}^{a_i} A_i(x)f(x)dx - \int_{b_{i-1}}^{b_i} B_i(x)f(x)dx$$
$$= \int_{a_{i-1}}^{a_i} f(x)dx - \int_{b_{i-1}}^{b_i} f(x)dx$$
$$+ \int_{a_{i-2}}^{a_{i-1}} A_{i-1}(x)f(x)dx - \int_{b_{i-2}}^{b_{i-1}} B_{i-1}(x)f(x)dx$$
$$\int_{a_0}^{a_1} A_1(x)f(x)dx - \int_{b_0}^{b_1} B_1(x)f(x)dx = \int_{b_1}^{a_1} f(x)dx$$

이므로, 위 식은 다음과 같이 단순화 된다.

$$\int_{a_{i-1}}^{a_i} A_i(x)f(x)dx - \int_{b_{i-1}}^{b_i} B_i(x)f(x)dx = \int_{b_i}^{a_i} f(x)dx$$

이 식으로부터 각 퍼지분할에 속한 A_i와 B_i는 상대적 위치에 있어서 다음과 같은 특성을 갖는다[1].

오른쪽 위치의 특성

B_i의 지지의 최소값(Lower Bound)이 A_i의 지지의 최소값 보다 크면(오른쪽에 위치하면)다음이 성립한다.

특성 1 : B_{i+1}의 지지의 최소값은 A_{i+1}의 지지의 최소값 보다 작다.

특성 2 : B_{i-1}의 지지의 최소값은 A_{i-1}의 최소값 보다 작으며, A_{i-2}의 지지의 최소값 보다 크다.

왼쪽 위치의 특성

B_i의 지지의 최소값이 A_i의 지지의 최소값 보다 작으면(왼쪽에 위치하면) 다음이 성립한다.

특성 3 : B_{i+1}의 지지의 최소값은 A_{i+1}의 지지의 최소값보다 크며, A_{i+2}의 지지의 최소값 보다 작다.

특성 4 : B_{i-1}의 지지의 최소값은 A_{i-1} 지지의 최소값 보다 크다.

본 논문에서 택한, 주어진 두 점과 가장 가까운 피크치 사이의 거리의 최대값을 최소화하는 방법은, 곧, $\min\max(|a_j-P_1|, |a_k-P_2|)$인 a_j 및 a_k를 찾는 방법이다. 단, 허용 오차는 δ로 하며, 퍼지균등화 조건을 갖는 이단퍼지분할은 김경택 외[1]가 제시한 알고리듬을 이용하여 생성한다. 앞에서 제시된 퍼지균등화 조건을 갖는 이단퍼지분할의 상대적 위치에 따른 특성에 의하면, $j+k$가 홀수이고, $(a_j-P_1)\times(a_k-P_2)\geq 0$이면, a_j (또는 a_k)가 증감할 때, a_k(또는 a_j)는 그 반대방향으로 이동하게 되어, $||a_j-P_1|-|a_k-P_2||\leq\delta$이면, a_j(또는 a_k)의 증감이, $\min\max(|a_j-P_1|, |a_k-P_2|)$ 값을 감소시킬 수 없게 된다. $j+k$가 홀수이고, $(a_j-P_1)\times(a_k-P_2)<0$이면, $|a_j-P_1|$ (또는 $|a_k-P_2|$)의 감소가, $\min\max(|a_j-P_1|, |a_k-P_2|)$ 값의 감소를 가져올 수 있다. $j+k$가 짝수일 때의 a_j (또는 a_k)의 증가 또는 감소에 따른 $\min\max(|a_j-P_1|, |a_k-P_2|)$ 값의 영향도 쉽게 알 수 있다. 이러한 특성을 근거로 다음과 같은 알고리듬을 제시한다.

1) $a_j=P_1$인 퍼지분할 $\boldsymbol{A}=\{A_1, A_2, \cdots, A_n\}$를 만든다. 만일 $|a_k-P_2|\leq\delta$ 이면, $\boldsymbol{A}$는 최적해 이므로 멈춘다.
2) $a_k=P_2$인 퍼지분할 $\boldsymbol{B}=\{B_1, B_2, \cdots, B_n\}$를 만든다. 만일 $|a_j-P_1|\leq\delta$ 이면, $\boldsymbol{B}$는 최적해 이므로 멈춘다.
3) $c_1=\dfrac{a_1+b_1}{2}$인 퍼지분할 $\boldsymbol{C}=\{C_1, C_2, \cdots, C_n\}$을 만든다.
4) 만일
   $|c_j-P_1|\leq\delta$ 이고, $|c_k-P_2|\leq\delta$이거나
   $j+k$가 홀수이고, $(c_j-P_1)\times(c_k-P_2)\geq 0$이며,
   $||c_j-P_1|-|c_k-P_2||\leq\delta$이거나,
   $j+k$가 짝수이고, $(c_j-P_1)\times(c_k-P_2)\leq 0$이며,
   $||c_j-P_1|-|c_k-P_2||\leq\delta$이면,
   $\boldsymbol{C}$는 최적해 이므로 멈춘다.
5) $d_j=c_j+\dfrac{||c_j-P_1|-|c_k-P_2||}{|c_j-P_1|+|c_k-P_2|}\times\dfrac{|c_j-P_1|}{P_1-c_j}$인 퍼지분할 $\boldsymbol{D}=\{D_1, D_2, \cdots, D_n\}$를 만든다.
6) 만일 $\max(|d_j-P_1|, |d_k-P_2|)<\max(|c_j-P_1|, |c_k-P_2|)$이면, 모든 i에 대하여 $C_i=D_i$가 되도록 C_i를 바꾼 후, 4)로 간다.
7) $d_j=c_j+\dfrac{1}{2}\dfrac{||c_j-P_1|-|c_k-P_2||}{|c_j-P_1|+|c_k-P_2|}\times\dfrac{|c_j-P_1|}{P_1-c_j}$인 퍼지분할 $\boldsymbol{D}=\{D_1, D_2, \cdots, D_n\}$를 만든 다음, 6)으로 간다.

4. 예 제

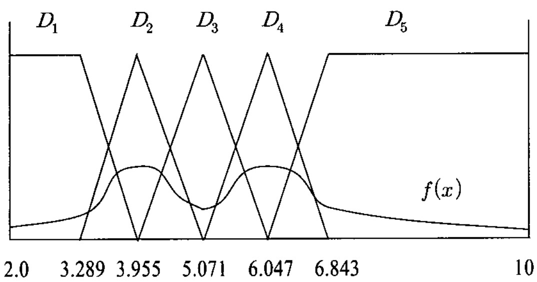
다음과 같은 확률밀도함수에 대하여 5개의 퍼지집합으로 이루어진 퍼지균등화 조건을 갖는 2단 퍼지분할을 만든다고 가정하자.

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}}exp\left(-\frac{(x-4)^2}{2}\right), && \text{if } 2\leq x\leq 5 \\ &= \frac{1}{\sqrt{2\pi}}exp\left(-\frac{(x-6)^2}{2}\right), && \text{if } 5\leq x\leq 10 \\ &= 0, && otherwise \end{aligned}$$

주어진 두 점은 입력분포에서 피크를 이루는 4, 6이고, $\delta=0.002$라 하자. 제안된 알고리듬을 적용하면, $\boldsymbol{A}$는 2, 3.235, 4.002, 5.014, 6.990, 6.788, 10에서 피크가 된다.

B는 2, 3.336, 3.908, 5.123, 6.001, 6.894, 10에서 피크를 이룬다. 알고리듬의 3)에서 **C**는 2, 3.286, 3.955, 5.069, 6.051, 6.842, 10 에서 피크가 되며, $\max(|c_j - P_1|, |c_k - P_2|) = 0.051$이다. 알고리듬 5)에서 **D**는 2, 3.289, 3.952, 5.071, 6.047, 6.843, 10에서 피크 값을 갖는다(<그림 2>). $\max(|d_j - P_1|, |d_k - P_2|) < \max(|c_j - P_1|, |c_k - P_2|)$이므로, 알고리듬 4)로 이동하여 $j + k = 6$, $(c_j - P_1) \times (c_k - P_2) \le 0$이고, $\| c_j - P_1 | - | c_k - P_2 \| = 0.001$이 $\delta = 0.002$ 보다 작게 되어 D가 최적해가 된다.

본 예제에서는 계산상의 용이성과 결과의 시각적 효과를 강조하기 위하여 두 개의 피크를 갖는 입력분포를 사용하였다. 그러나 다수의 피크를 갖는 임의의 분포에 대해서도, 두 점을 지정하면, 제시된 알고리듬에 의해 퍼지분할을 생성할 수 있다.



<그림 2> 예제에서 주어진 분포에 대하여 퍼지균등화 조건을 만족하는 퍼지분할

5. 결 론

데이터 마이닝에서 적절한 크기로의 정보 알갱이화는 의미 있고 이해하기 쉬운 패턴 발견이라는 목적을 보다 빠른 시간 내에 다다르게 하는 중요한 수단중의 하나이다. 이때 정보 알갱이화는 정보알갱이를 어떻게 표현할 것인가 하는 문제와 정보알갱이와 이를 뒷받침하는 데이터와 어떠한 관계를 가져야 하는가 하는 문제를 포함한다. 본 논문에서는 퍼지분할을 사용하여 데이터를 정보알갱이로 표현하였다. 퍼지분할에 속한 퍼지집합으로 표현되는 정보알갱이들과 이에 연관된 데이터와의 관계는 퍼지 균등화 조건으로 규정하였다.

본 논문에서는 주어진 두 점에서 동시에 피크를 이루는 퍼지분할이 항상 존재하지는 않음을 입증하였다. 아울러, 이러한 경우, 그들 사이의 거리의 최대값을 최소화하는 퍼지분할을 찾는 알고리듬을 제시하였다. 주어지는 두 점은, 사용자가 지정하는 것이므로, 입력분포의 특성을 반영하여 입력분포에서 피크를 이루는 점들 중에서 선택하는 것이 자연스럽지만, 꼭 그 점들 중에서 선택해야한다는 제한조건은 없다. 본 논문은 추후 3개 이상의 점에 대하여서도 적용 가능한 알고리듬을 개발하는 등, 입력분포 특성을 보다 더 반영하는 연구로의 확장이 가능할 것으로 사료된다.

참고문헌

[1] 김경택, 김종수; "퍼지 균등화조건을 갖는 2단 퍼지분할", 산업경영시스템학회지, 25(6) : 54-59, 2002.

[2] Au, W. H., Chan, K. C. C., and Wong A. K. C.; "A Fuzzy Approach to Partitioning Contonuous Attributes for Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, 18(5) : 715-719, 2006.

[3] Foulloy, L. and Benoit, E.; "Building a Class of Fuzzy Equivalence Relations," *Fuzzy Sets and Systems*, 15 7 : 1417-1437, 2006.

[4] Guillaume, S. and Magdalena, S.; "Expert Guided Integration of Induced Knowledge into a Fuzzy Knowledge base," *Soft Computing*, 10 : 773-784, 2006.

[5] Hirota, K. and Pedrycz, W.; "Fuzzy Computing for Data Mining," *Proceeding of the IEEE*, 87(9) : 1575-1600, 1999.

[6] Pedrycz, W.; "Fuzzy Sets Framework for development of Perception Perspective," *Fuzzy Sets and Systems*, 37 : 123-137, 1990.

[7] Pedrycz, W.; "Selected Issues of Frame of Knowledge Representation Realized by Means of Linguistic Labels," *International Journal of Intelligent System*, 7 : 155-170, 1992.

[8] Pedrycz, W. and Gomide, F.; An Introduction to Fuzzy Sets: Analysis and Design, MIT Press, Cambridge, MA, 1998.

[9] Pedrycz, W.; "Fuzzy Equalization in the Construction of Fuzzy Sets," *Fuzzy Sets and Systems*, 119 : 329-335, 2001.

[10] Pedrycz, W. and Bargiela, A.; "Granular Clustering: A Granular Signature of Data," *IEEE Transactions on Systems, Man, and Cybernetics*, 32(2) : 212-224, 2002.

[11] Zadeh, L. A.; "The Concept of a Linguistic Variable and its Application to Approximate Reasoning," *Information Science*, 8(3) : 199-249, 1975.

[12] Zadeh, L. A.; "Fuzzy Sets and Information Granularity," Advances in Fuzzy Set Theory and Applications, (ed.) M.M. Gupta, R. K. Ragade, and R. R. Yager, North Holland, Amsterdam, 3-18, 1979.

[13] Zadeh, L. A.; "Toward a Theory of Fuzzy Information Granulation and its Uncertainty in Human Reasoning

and Fuzzy Logic," *Fuzzy Sets and Systems*, 90 : 11-127, 1997.

[14] Zhuang, H. and Wu, X.; "Membership Function Modification of Fuzzy Logic Controllers with Histogram Equalization," *IEEE Transactions on Systems, Man, and Cybernetics*, 31(1) : 125-132, 2001.

[15] Zimmermann, H. J.; Fuzzy Set Theory and Its Applications, Kluwer, Boston, MA, 1991.