



## 퍼지 접속 연산자를 이용한 정보 융합 방법

전북대학교 전자공학과 이 준 환\*

### 목 차

I. 서 론	IV. 모의 실험
II. 퍼지 접속 연산자	V. 결 론
III. 학습 알고리즘	

### I. 서 론

의사 결정을 위한 불확실한 정보의 융합 방법은 불충분한 데이터로부터 제공되는 정보들로부터 의사 결정에 따르는 애매 모호성을 감소시키기 위한 목적으로 많은 이용 범위를 갖고 있다[1, 2]. 예를 들어 최근의 센서 기술의 발전으로 다양한 센서들이 싼 가격으로 공급되고 있기 때문에 다양한 센서들로부터 공급되는 불충분한 정보들을 총체화하고 이들을 단일 센서를 이용하는 경우의 한계를 극복하려는 시도가 컴퓨터 비전 시스템의 구성에 많이 적용되고 있다[13].

이러한 불확실한 정보 융합을 위해서는 반드시 융합을 위한 연산자를 필요로 한다. 따라서, 정보 융합에 사용되는 연산자의 종류가 정의된 방법론에 따라, 확률을 기본으로 하는 방법[1, 5, 12], 퍼지이론을 이용하는 방법들로 구분될 수 있다. 확률이론을 기본으로 하는 경우에는 Bayes 규칙이 융합을 위한 연산자로 사용되고, 믿음이론을 기본으로 하는 경우에는 Dempster의 규칙이 융합을 위한 연산자로 사용되고 있다. 그러나, Bayes 규칙이나, Dempster의 규칙을 적용하는 데는 확률론이나

믿음이론이 정의된 공리에 의한 융합 연산자의 적용에 몇가지 제약 조건이 가해진다. 예를 들면, Bayes 규칙의 apriori 확률이나, Dempster 규칙의 non-iterativeness 등이 그것이다.

퍼지이론을 이용하는 정보 융합 방법에는 퍼지 수(number)를 이용하는 방법[6], 퍼지 적분을 이용하는 방법[22], 퍼지 접속 연산자를 이용하는 방법[8, 9, 10, 11] 등 몇가지 형태가 발표되었다. 퍼지 수를 이용하는 방법은 각각의 정보원의 신뢰도에 따른 퍼지집합 형태의 가중치를 부여하고, 정보원으로부터 제공되는 정보를 퍼지집합 형태로 모형화하여, 이를 확장 원리(extension principle) 등의 퍼지 수 연산을 이용하고 융합의 결과는 퍼지집합이다. 퍼지 적분을 이용하는 방법[22]은 Sugeno의 g-measure를 이용하여 정보원 또는 평가기준의 중요성을 표현하고 평가 기준에 의한 평가치를 평가 기준으로 구성된 계층적인 의사 결정 모형에서 퍼지 적분을 이용하여 총체화하는 방법을 이용하고 있다. 이들 퍼지이론을 이용하는 두 방법에 있어서는 확률론이나, 믿음이론의 경우에서와 같은 융합 연산자 적용의 제약 조건이 없는 장점이 있으나, 정보원의 중요성을 부여하는 가중치들은 확률이론 적용시의 apriori 확률과 같이 미리 결정되어야 하는 단점이 있다.

\*정 회 원

본 논문에서는 퍼지 접속 연산자를 이용하는 새로운 정보 융합 방법을 소개한다. 퍼지이론에는 t-norm, co-norm, 일반화된 보상 연산을 기반으로 하는 수많은 접속 연산자[2, 3, 7, 14, 15]가 제공되고 있기 때문에, 정보의 총체화시 정책에 따라 다양한 연산자를 선택할 수 있을 뿐 아니라, 연산자 적용에 따르는 제약 요건도 없다. 본 논문에서 소개하는 방법에서는 학습을 통하여 총체화에 필요한 연산자와 연산자에 부가된 정보원의 중요성을 표현하는 파라미터 등을 자등으로 결정하기 때문에 가중치 결정에 선지식 등을 필요로 하지 않는다.

제안된 방법에서 정보 융합 문제는 다수의 평가 지수를 갖는 의사 결정 모형으로 가정하고, 일반적으로 이러한 평가 지수는 계층적인 구성이 가능하다. 예를 들어 은행의 대출을 요구하는 사람의 신용도[21]를 측정한다면, 신용도는 대출자의 경제력과 개인적 성질(personality)을 기초로 구할 수 있고, 경제력이란 평가 지수는 대출자가 가지고 있는 등산 또는 부동산의 정도를 근거로 하고 개인적 성질은 대출자의 성취도라든가 사업 수완 등을 바탕으로 결정될 수 있다. 즉 평가 지수들은 계층적인 망의 형태로 구성될 수 있고, 이러한 망에서, 어떤 평가 지수의 총체적인 만족도(degree of satisfaction)는 망의 하위 계층의 평가 지수들에서 제공되는 만족도들을 융합함으로써 계산될 수 있다. 본 논문에서 제안하고자 하는 정보 융합 방법은 이러한 총체적인 만족도를 구하는데 퍼지 접속 연산자를 사용하며, 정보 융합시의 비관도, 낙관도 또는 보상 정도를 조정할 수 있고, 이러한 성향은 학습 데이터를 이용하여 적응적으로 계산되어진다.

제안된 학습 방법은 전체적인 평가 지수들이 계층적인 신경회로망의 구조를 갖게 만들고, 신경회로망의 각 단(unit)들이 하나의 평가 지수에 해당되며, 신경회로망의 활성화 함수[17]를 일련의 퍼지 접속 연산자로 가정하고, 연산자에 부수되는 파라미터와 연산자를 역전과 학습에 의해 구하며, 이때 각 단에 입력되는 값은 하위 계층에서 제공되는 만족도들로 [0, 1]의 값을 갖는다.

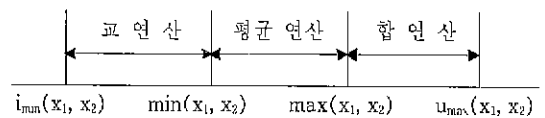
제2장에서는 퍼지 접속 연산자들을 요약하고, 제3장에서는 정보 융합의 성향을 결정하는 파라미터를 찾아내기 위한 학습 방법과, 제4장에서는 몇가지 데이터들을 이용한 모의 실험 결과와 제5장에서는 결론을 맺고자 한다.

## II. 퍼지 접속 연산자

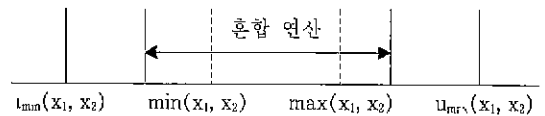
퍼지 집합 이론에는 소속 함수 값을 총체화(aggregation)하는 수많은 접속 연산자가 정의되어 있다. 이들 접속 연산자들은 총체화의 성향(attitude) 또는 영역(range)에 따라 합(union), 교(intersection) 및 보상(compensative) 접속 연산자 등의 세가지 종류로 태별되고, 보상 연산자는 평균(mean) 연산자와 혼합(hybrid) 연산자들로 구분된다. (그림 1)에는 이들 연산자들의 영역을 표시하였다. (그림 1)에서  $i_{\min}(x_1, x_2)$ 과  $u_{\max}(x_1, x_2)$  연산자는 각각 다음과 같이 정의되는 가장 비관적이고 낙관적인 연산자이다[7].

$$i_{\min}(x_1, x_2) = \begin{cases} 0 & (x_1 \neq 1 \text{이고 } x_2 \neq 1 \text{일 경우}) \\ x_1 & (x_2 = 1 \text{일 경우}) \\ x_2 & (x_1 = 1 \text{일 경우}) \end{cases} \quad (1)$$

$$u_{\max}(x_1, x_2) = \begin{cases} 0 & (x_1 \neq 0 \text{이고 } x_2 \neq 0 \text{일 경우}) \\ x_1 & (x_2 = 0 \text{일 경우}) \\ x_2 & (x_1 = 0 \text{일 경우}) \end{cases} \quad (2)$$



(가) 합, 교 및 평균 연산자의 영역



(나) 혼합 연산자의 영역

(그림 1) 접속 연산자의 영역

### 2.1 합 접속 연산자

합 접속 연산자는 연산자의 입력중 값이 가장 큰 것보다 연산 결과가 작지 않은 성질을 갖는 연산자로서, 영역은  $[\max(x_1, x_2), u_{\max}(x_1, x_2)]$ 의 범위를 갖고, 정보 융합시 낙관적인 성향을 나타내는 연산자이다. 가장 대표적인 "max" 연산은 합 접속 연산자 중에 가장 비관적인 연산자이고, 합 연산자의 정의는 co-norm의 정의와

동일하기 때문에 발생 함수(generating function)에 따라 무수히 많은 합 연산자의 정의가 가능하다. 또한 이들 합 접속 연산자에는 파라미터를 가지고 있기 때문에 입력의 최대치와 1 사이의 영역에서 출력 값을 조정할 수 있는 일군(family)의 연산자가 정의되기도 한다. Yager의 일반화된 합 연산자[19]도 이들 중 하나로서, n개의 입력에 대해

$$u(x_1, \dots, x_n) = \min(1, (x_1^p + \dots + x_n^p)^{1/p}) \quad (3)$$

로 정의된다. 식 (3)의 합 연산자는 파라미터  $p \in (0, \infty)$  값에 따라 단조 감소하고,  $[\max(x_1, x_2, \dots, x_n), 1]$  영역을 갖기 때문에 p 값을 조정함에 따라, 총체화에 따르는 낙관도를 조정할 수 있게 된다.

### 2.2 교 접속 연산자

교 접속 연산자는 총체화된 값이 모든 입력 중 가장 작은 입력보다 크지 않은 성질을 갖는 연산자로서,  $[\min(x_1, x_2), \min(x_1, x_2)]$ 의 영역을 갖는다. 따라서, 정보 융합시 비판적인 성향을 나타내는 연산자로 볼 수 있고, 대표적인 "min" 연산자는 가장 낙관적인 교 접속 연산자이다. 합 연산자와 마찬가지로, 교 연산자는 t-norm과 같은 방법으로 정의되기 때문에 발생함수로부터 무수히 많은 연산자를 정의할 수 있고, 이들 중에는 파라미터를 가지고 비판도를 표현할 수 있는 일군의 연산자들이 있다.

예를 들어 Yager의 교 연산자[19]는

$$l(x_1, \dots, x_n) = 1 - \min[1, ((1-x_1)^p + \dots + (1-x_n)^p)^{1/p}] \quad (4)$$

과 같이 정의되고, 파라미터  $p \in (-\infty, 0)$  값에 따라 단조 감소하는 함수이고,  $[0, \min(x_1, x_2, \dots, x_n, \dots, x_n)]$ 의 영역에서 비판도를 조정할 수 있다.

### 2.2 보상 연산자

많은 의사 결정 문제에 있어서는 합이나 교 연산자들의 극단적인 연산자보다는 보상 연산자를 이용하는 경우가 많이 있다. 특히, 서로 다른 개념으로부터 제공되는 정보의 융합에는 일련의 보상 연산자가 적합하다는 사실이 실험을 통하여 입증된 바 있다[3]. 퍼지 연산자에는 무수히 많은 종류의 보상 연산자들이 있으며, 이들은 정의되는 형태에 따라 평균 연산자와 혼합 연산자의 크게 두가지 부류로 분류될 수 있다. 평균 연산자의 경

우는 일련의 공리를 근거로 정의되는 반면, 혼합 연산자는 일반적인 합이나 교 접속 연산자의 산술 또는 기하 평균으로 정의된다.

정의 1 : 평균 연산자 m은

- ①  $m(a, b) \geq m(c, d)$  if  $a \geq c$  and  $b \geq d$  (monotonicity)
- ②  $\min(a, b) \leq m(a, b) \leq \max(a, b)$

의 성질을 만족하는  $m: [0, 1] \times [0, 1] \rightarrow [0, 1]$ 의 사상(mapping)으로 정의된다.

이러한 성질을 만족하는 연산자에는 산술 평균(arithmetic mean), 조화 평균(harmonic mean), 기하 평균(geometric mean), 파라미터를 이용하는 일반화된 평균(generalized mean) 등이 있다. Dyckhoff and Pedrycz에 의해 제안된 일반화된 평균[4]은

$$g(x_1, x_2, \dots, x_n; p, w_1, w_2, \dots, w_n) = \left( \sum_{i=1}^n w_i x_i^p \right)^{1/p} \quad (5)$$

와 같이 정의되고, 여기서

$$w_1 + w_2 + \dots + w_n = 1 \quad (6)$$

이고, 파라미터 p 값이  $-\infty$ 와  $+\infty$  사이에서 단조 증가 함수이며, min과 max 사이의 값을 갖는다. 따라서, 파라미터 p 값을 조정함에 의해 융합시 보상의 정도를 결정할 수 있고,  $w_i$  값들에 의해 정보원의 중요성을 표현할 수 있다. 식 (5)에서 모든  $w_i$  값들이  $1/n$ 이고,  $p=1$ 이라면 산술 평균,  $p=0$ 이라면 기하 평균,  $p=-1$ 이라면 조화 평균과 같다.

혼합 연산자란 일반적인 합과 교 연산자의 산술 또는 기하 평균으로 정의된다. 따라서, 무수히 많은 합과 교 연산자의 조합을 생각할 수 있으며, 영역은 혼합 연산자 정의에 사용된 교 연산자 값에서 합 연산자 값 사이의 범위를 가지고, 보상적인 성질을 갖게 된다. Zimmermann과 Zysno에 의해 정의된 r-모델은 이러한 연산자의 일종으로 식 (7)과 같이 정의된다[18, 20, 21].

$$y = \left( \prod_{i=1}^n x_i \delta_i \right) \left( 1 - \prod_{i=1}^n (1-x_i) \delta_i \right)^r \quad (7)$$

여기서  $\sum_{i=1}^n \delta_i = n$ 이고  $0 \leq r \leq 1$ 이다.

식 (7)에서 정의된 r-모델은 교 연산자의 일종인 대수적 적(product) 연산과 합 연산자의 일종인 대수적 합(sum) 연산 두 부분으로 되어 있고,  $\delta_i$ 는  $x_i$ 의 가중치이며, r는 보상의 정도를 표현하는 파라미터이다. 식 (7)에서 r 값이 0이면,

$$y = h_1(x_1, x_2, \dots, x_n) = \prod_{i=1}^n x_i \delta_i \quad (8)$$

입력이 가중된 대수적 곱 연산과 같고, 1이면

$$y = h_1(x_1, x_2, \dots, x_n) = \prod_{i=1}^n (1 - x_i) \delta_i \quad (9)$$

가중된 대수적 합과 같다.

정리 1:  $r$ -model은  $r$  값에 따라 단조 증가(monotonically nondecreasin)하고 영역은 대수적 곱으로부터 대수적 합에 이르는 범위를 갖는다.

증명:  $r$ -model의  $r$  값에 따른 편미분은

$$\frac{\partial y}{\partial r} = y \cdot \ln\left(\frac{h_2}{h_1}\right) \quad (10)$$

과 같다. 식 (10)에서  $h_1$ 과  $h_2$ 는 식 (9)에 주어졌고 만약, 동일한 입력에 대해  $h_2 > h_1$ 이면,  $y$ 가 항상 0 이상이기 때문에 식 (8)의 편미분은 항상 0 이상이고 단조 증가 함수이다. 식 (7)에서  $\delta_i$ 는 항상 0보다 작지 않아야 하기 때문에  $\delta_i = \alpha_i^2$  치환할 수 있다.  $p = h_2 - h_1$ 이라 놓자.  $p$ 의 최대 최소를 결정하기 위해 Lagrange Multiplier를 도입하여,

$$Y = p + \lambda \cdot (\sum_i \alpha_i^2 - n) \quad (11)$$

라 놓고, 이를 모든  $\alpha_i$ 에 대하여 편미분하면, 식 (12)와 같은 형태의  $n$ 개의 방정식을 얻는다.

$$-2 \cdot \alpha_i \cdot ((1 - h_2) \cdot \ln(1 - x_i) + h_1 \cdot \ln(x_i)) + 2 \cdot \alpha_i \lambda = 0 \quad (12)$$

식 (12)의  $n$ 개의 방정식은

$$\alpha_i = 0 \text{ 또는 } -(1 - h_2) \cdot \ln(1 - x_i) - h_1 \cdot \ln(x_i) + \lambda = 0 \quad (13)$$

일 경우 만족한다. 만약  $\alpha_i$ 가 모두 0이 아니라면, 일반적으로  $x_i$  값들은 서로 다르기 때문에 유일한  $\lambda$  값을 가질 수 없다. 따라서, 식 (12)의 방정식을 만족시키는 경우는 오직 하나의  $\alpha_i$  값만이 0이 아닐 때 유일한  $\lambda$  값을 가질 수 있다. 만약  $\alpha_i \neq 0$ 이고,  $i = k$ 를 제외한 모든  $\alpha_i = 0$  라면,  $\delta_k = \alpha_k^2 = n$ 이 되고,

$$p = 1 - (1 - x_k)^n - x_k^n \quad (14)$$

이다. 식 (14)에서  $p$ 는  $n=1$ 일 경우 0이고,  $n > 1$ 에서  $p$ 의  $n$ 에 대한 편미분은 항상 양이기 때문에

$$p = h_2 - h_1 \geq 0$$

이다.

### III. 학습 알고리즘

평가 지수들이 계층적 체제로 구성된 정보 융합망에서 각 평가 지수의 부분적인 만족도를 총체화하는 문제는, 만약 구체적인 망의 형태와 각 단(node)에서 사용되는 접속 연산을 알고 있다면 어려운 문제는 아니다. 그러나, 종종 망의 개략적인 형태만을 알고, 구체적인 접속 연산자는 알 수 없을 경우가 많다. 본 논문에서는 이러한 형태의 기정에서 출발하여, 망의 구체적인 구조와 각 단(node)에 필요한 연산 형태를 학습에 의해 구하는 방법을 소개하고자 한다.

본 논문에서 학습을 위해 필요한 데이터는 최하위 계층의 각 평가 지수에 대한 만족도들과 이러한 만족도들이 입력되었을 때의 최상위 계층의 총체화된 평가 지수로 가정하였다. 이러한 평가 지수들로 구성된 학습 데이터를 이용하여, 각 단의 연산자와 파라미터 결정은 가능한 연산자의 형태를 각 단에 가정하고 이들 연산자를 계층적 신경회로망의 활성화 함수라 가정하고, 역전파 학습과 같은 gradient descent 방법을 이용하였다.

만약, 어떤 단에  $n$ 개의 입력이 가해진다고 하자.  $N$ 가지의 학습 데이터  $\{x_{1k}, \dots, x_{nk}; Y_k\}$ 가 주어졌다고 가정하자. 여기서  $k=1, \dots, N$ 이고  $Y_k$ 는  $k$ 번째 입력에 대한 원하는 출력이다. 문제는 실제 총체화된 값과 원하는 출력과의 불일치가 가장 작아진다는 관점에서 최적의 연산자와 파라미터 값을 결정하는 것이다. 여기서, 불일치 정도는

$$E = \sum_{k=1}^N (f_k - Y_k)^2 \quad (15)$$

로 정의하였고, 식 (15)에서  $f_k$ 는 입력  $(x_{1k}, \dots, x_{nk})$ 에 대한 접속 연산자를 사용하였을 경우의 총체화된 값이다.

학습 알고리즘은 식 (15)에서,  $E$  값의 연산자  $f_k$ 의 파라미터에 대한 negative gradient를 구하고 이를 작은 양의 실수 값(learning constant)과 곱하여, 현재의 파라미터 값에 더하여, 모든 망의 파라미터 값이 더 이상 변하지 않을 때까지 반복 수행하는 역전파 학습 방법을 이용하였다.

본 논문에서는 망에 대한 선지식의 정도에 따라 다음과 같이 세가지의 알고리즘을 생각할 수 있다.

#### 3.1 알고리즘 1

첫째, 망의 각 단의 접속 형태, 즉 연산자의 형태는

알고 있으나, 접속된 연산자의 파라미터 값을 모른다고 가정한다. 이 경우 망의 접속 형태가 비판적인 융합이라면 Yager의 교 연산자, 낙관적이라면 Yager의 합 연산자, 보상적이라면 일반화된 평균 연산자를 가정하였다. 이들 세 연산자는 동일한 발생 함수  $x^n$ 로부터 정의될 수 있다. 망의 각 접속단에 이들 세 연산자를 활성 함수로 고정시키고, Yager의 합 또는 교 연산자일 경우는 낙관도와 비관도를 표현하는 파라미터  $p$  값을 찾아내고, 보상 연산자일 경우는 일반화된 평균의 보상정도를 표현하는  $p$  값과 입력의 가중치에 해당되는  $w_i$  값들을 결정한다. 이때, 초기의 파라미터 값들은 Yager의 합 연산일 경우는  $p=1$ 에서, 교 연산일 경우는  $p=-1$ 에서, 일반화된 평균의 경우는  $p=1$ , 모든  $w_i$ 는  $1/n$ , 즉 산술 평균에서 시작한다.

그러나, 식 (5)에서 알 수 있듯이 입력의 가중치  $w_i$ 에 는 제약 요건이 있기 때문에 일반적인 신경 회로망의 제약 조건이 없는 망의 강도(network strength)를 구하는 방식을 그대로 적용할 수는 없다. 본 논문에서는 이러한 제약 조건을 우회시키는 다음과 같은 방법을 고안하였다. 식 (5)와 (6)에서, 함수  $g$ 는 식 (16)과 같은 방법으로 고쳐 쓸 수 있다.

$$g = \left( \frac{w_1^2}{\sum w_i^2} x_1^p + \dots + \left( \frac{w_n^2}{\sum w_i^2} x_n^p \right)^{1/p} \quad (16)$$

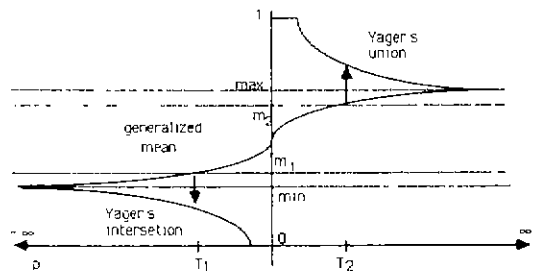
식 (16)에서  $w_i$ 들은 식 (5)의 가중치와는 다르게 제약 조건이 없고, 식 (16)의 새로운 가중치들  $w_i^* = w_i^2 / (\sum w_i^2)$ 로 정의하면 식 (6)의 제약 조건을 만족시키게 된다. 즉 알고리즘에서 식 (16)의  $w_i$ 를 제약 조건 없이 개선하여 수렴시키고 수렴된 값에서  $w_i^*$ 를 찾아내면, 가중치의 제약 조건을 우회하여 제충적 망의 파라미터를 찾아내는데 역전파 학습법을 이용할 수 있다. 실제 알고리즘의 구현에 있어서  $p$ 의 값이 무한히 증가할 수는 없기 때문에,  $-15 < p < 15$ 로 제한하였다.

### 3.2 알고리즘 2

둘째, 망의 각 단의 접속 형태를 모르고, 단지 망의 개략적인 구조만을 안다고 가정하자. 이러한 경우는 가능한 각 단의 접속 연산자의 형태는 Yager의 합, 교 또는 일반화된 평균중의 하나라고 가정한다. 2장에서 언급된 바와 같이 Yager의 합 연산자는 입력들 중 최대치와 1 사이의 융합 범위를 가지고, 교 연산자는 최소치와 0 사이의 범위, 일반화된 평균은 최대치와 최소치의 범위

를 가지기 때문에 진반적인  $[0, 1]$ 의 결정 범위를 모두 포함시킬 수 있다. (그림 2)는 이들 연산자의 영역을 도시하였다.

본 알고리즘에서 최초의 모든 단들은 일반화된 평균 연산자라고 가정하고,  $p=1$ 과 모든  $w_i$ 들은  $1/n$ , 즉 산술 평균에서 학습을 시작한다. 학습 도중 만약 어떤 단의 일반화된 평균에서  $p$  값이 점점 증가하여 정해진 문턱값 ( $p=2$ )을 넘어서면 그 단의 연산자를 Yager의 합 연산자로 적합한 형태의 연산자(활성화 함수)를 교체한다. 즉, 일반화된 평균에 있어서,  $p$  값이 커지면 합 연산자와 유사한 성질을 갖는 성질을 이용하였다. 또한, 일반화된 평균에서  $p$  값이 작아지면 교 연산자와 유사한 형태가 되기 때문에, 만약 알고리즘 수행 도중 최초 가정된 일반화된 평균의  $p$  값이 정해진 문턱값( $p=2$ )보다 작아지면 Yager의 교 연산자로 교체한다. 접속 연산자를 Yager의 교 또는 합 연산자로 교체할 경우 일반화된 평균에서 가정되었던  $w_i$  값들은 모두 1로 하고 비관도나 낙관도를 표현하는  $p$  값을 찾기 위해 학습을 계속한다. 단, 이 알고리즘의 문제점은  $p > 2$ 이거나  $p < 2$ 의 일반화된 평균값은 도달할 수 없는 단점이 있다. 즉 (그림 2)에서  $[m_2, \max]$ 와  $[\min, m_1]$  영역은 도달할 수 없다.



(그림 2) Yager의 연산자와 일반화된 평균의 영역과 알고리즘 2

### 3.3 알고리즘 3

알고리즘 3에서는 접속 가능한 연산자의 형태를 모두  $r$ -model로 가정한다. 실제  $r$ -model의 영역은 대수적 교 연산자와 합 연산자 사이의 값을 갖기 때문에 (그림 1-(나))에서와 같이 하한이 모든 입력들 중 최소치보다 작고, 상한은 최대치보다 크다. 즉,  $r$ -model의 영역이 합이나 교 연산 영역까지 침투되어 있기 때문에  $r$ -model 만으로도 거의 모든  $[0, 1]$ 의 영역을 포함시킬 수 있다.

또한  $r$ -model은  $r$  값에 따라 단조 증가하기 때문에 학습 결과  $r$  값이 작다면, 바람직한 연산은 교 연산이고,  $r$  값이 크다면 합 연산이며,  $r$  값이 0.5 근방이라면 보 상 짐속 연산자라는 것을 알 수 있다. 알고리즘 시작시  $r$ 는 0.5에서 시작하고,  $\delta_i$  값들은 모두 1로 동일한 가중치를 갖고 시작한다.

그러나, 식 (7)에서 보는 바와 같이  $r$ -model의  $r$  값과  $\delta_i$  값들에는 제약 조건이 있기 때문에, 역전파 학습에 의해 파라미터를 개선해 가는데 어려운 점이 있다. 본 논문에서는 이러한 제약 조건을 다음과 같이 우회하였다. 식 (7)에서  $r$ 와  $\delta_i$ 를 각각

$$r = \frac{a^2}{a^2+b^2} \text{ 와 } \delta_i = \frac{n\delta_i^{*2}}{\sum \delta_k^{*2}} \quad (17)$$

로 치환하자. 식 (17)에서,  $a$ 와  $b$  또는  $\delta_i^*$ 에 의해 구한  $r$ 와  $\delta_i$ 는 식 (7)에 제약 조건을 만족시키기 때문에 알고리즘에서는  $a, b$ 와  $\delta_i^*$ 를 개선해 나갈 수 있다.

#### IV. 모의 실험

본 장에서는 3장에서 소개된 알고리즘들을 이용하여 모의 실험을 수행한 결과를 제시한다.

[예제 1]: 이 예제에서는 3장에서 제안된 알고리즘의 정당성을 실험적으로 판단하기 위하여 단순한 합 연산에 적합하게 학습 데이터를 만들었다. <표 1>에서 원하는 출력은 항상 두 입력의 최대치보다 크게 하였고, 실제 학습 결과는 이러한 원하는 출력과 매우 근접함을 알 수 있다. 알고리즘 1에서는 Yager의 합 연산자를 가정하였고, 알고리즘 2에서는 일반화된 평균 연산자로 시작하여  $p$  값이 2를 넘어서 Yager의 합 연산자로 적합한 연산의 형태가 교체된 후 알고리즘 1과 같은  $p$  값으로 수렴하였다. 알고리즘 1과 2의 수렴 후  $p$  값은 1.30이었다. 알고리즘 3을 이용하였을 때는  $\delta_1=0.97, \delta_2=1.93, r$

<표 1> 합 연산자의 실험

입 력		출 력	학 습 후 출 력	
$x_1$	$x_2$	$y_k$	알고리즘 1과 2	알고리즘 3
0.343	0.123	0.412	0.412	0.419
0.111	0.999	0.111	0.111	0.121
0.037	0.222	0.003	0.006	0.009
0.900	0.200	0.150	0.200	0.174

=0.99로 수렴하였다. 즉 예상하였던 바와 같이 1에 가까운  $r$  값으로 수렴하므로 합 연산자에 가까워짐을 확인할 수 있었다.

[예제 2]: 본 예제에서는 [예제 1]과 같은 입력 데이터에 대해 원하는 출력을 항상 두 입력의 최소치보다 작게 하여, 적합한 짐속 연산자의 형태가 교 연산이 되도록 설정하였다. 알고리즘 1에서는 Yager의 교 연산자로 정의한 후 학습을 시도하였고, 알고리즘 2에서는 일반화된 평균 연산자로 시작하여 Yager의 교 연산자로 바뀐 다음 동일한  $p$  값 -6.73으로 수렴하였다. 알고리즘 3에서는  $\delta_1=0.97, \delta_2=1.03, r=0.01$ 로 수렴하여, 교 연산자의 형태를 나타내었다.

<표 2> 교 연산자의 실험

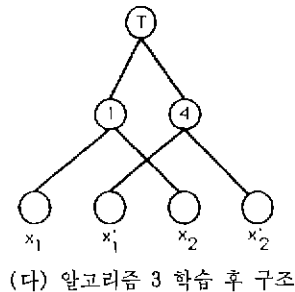
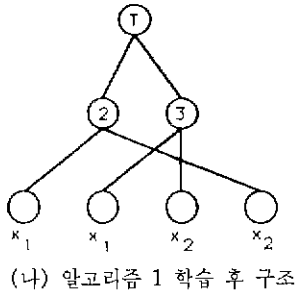
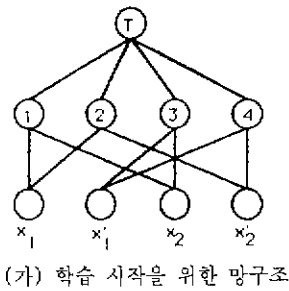
입 력		출 력	학 습 후 출 력	
$x_1$	$x_2$	$y_k$	알고리즘 1과 2	알고리즘 3
0.343	0.123	0.123	0.105	0.042
0.111	0.999	0.111	0.111	0.121
0.037	0.222	0.003	0.006	0.009
0.900	0.200	0.150	0.200	0.174

[예제 3]: (EXOR 문제) 본 예제에서는 EXOR 문제에 대한 학습 결과를 알아보기 위해 (그림 3)과 같은 구조의 망을 구성하였고, <표 3>과 같은 입력을 가정하였다. 알고리즘 1을 이용하는 학습에서는 각 단의 연산자를 일반화된 평균 연산이라 가정하였다. 학습이 진행되면서 최상위 단 T에 연결된 단 1과 단 4의 가중치는 0으로 수렴하고, 단 2와 단 3에 연결된 가중치는 0.5로 수렴하였다. 또한 단 2와 단 3에 연결된 입력들에 대한 가중치는 모두 0.5로 수렴하였다. 각 단의 파라미터  $p$  값을 보면 단 T의 경우  $p=5.202$ , 단 2와 단 3의 경우는 -1.322로 수렴하였다. 일반화된 평균이 큰  $p$  값에서는 합 연산자에 가까워지고, 낮은  $p$  값에서는 교 연산자에 가까워지는 현상을 고려하면, 전반적인 학습 결과는  $x_1x_2' + x_1'x_2$ 의 형태로 근사화될 수 있고, 이는 학습에 의한 EXOR의 논리함수를 찾아낼 수 있음을 입증한다.

알고리즘 3을 학습에 적용하였을 경우는 단 T와 연결된 단 1과 단 4의 가중치  $\delta$  값들은 2.0으로 수렴하고, 단 2와 단 3에 연결된 값들은 0으로 수렴하였다. 또한 단 1과 단 4에 연결된 입력들의 가중치  $\delta$  값들은 모두

2.0으로 수렴하였고, 각 단의  $r$  값들은 단 T의 경우 0.377, 단 1과 단 4의 경우 모두 1.0으로 수렴하였다.  $r$ -model이 큰  $r$  값에서는 합 연산자와 유사하고 작은  $r$  값에서는 교 연산자와 같이 작용하는 성질을 고려하면, 학습 결과는  $(x_1+x_2)(x_1'+x_2')$ 으로 근사화될 수 있으며, 학습에 의해 EXOR의 논리함수를 찾아낼 수 있음이 실험적으로 입증되었다.

본 예제를 통하여 제안된 학습 알고리즘은 적응적으로 논리식(adaptive logic)을 근사화할 수 있음이 실험을 통해 입증되었고 아울러, 일반화된 평균이나  $r$ -model 등을 이용할 경우 학습을 통하여, 각 단의 총체화된 평가치에 불필요한 입력으로부터의 가중치를 0으로 만들어 감에 의해 의사 결정시의 redundant한 입력을 제거할 수 있음을 알 수 있다.



(그림 3) EXOR 실험의 망 구조

〈표 3〉 EXOR 데이터

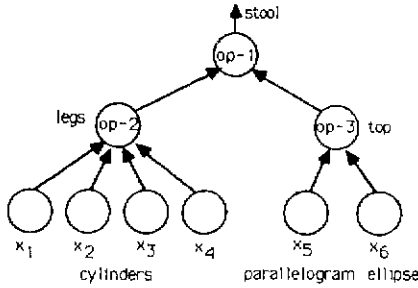
입 력		출 력	학 습 후 출 력	
$x_1$	$x_2$	$y_k$	알고리즘 1	알고리즘 3
0.100	0.100	0.150	0.160	0.180
0.100	0.900	0.850	0.780	0.810
0.900	0.100	0.850	0.780	0.810
0.900	0.900	0.150	0.260	0.180

[예제 4]: (Stool 문제) 본 예제에서는 가상적인 의자를 가정하여 학습 데이터를 다음과 같은 규칙에 의해 구성하였다.

“민약 어떤 물체가 4개의 다리중 2개 이상의 다리가 존재하고, 다리들에 의해 반어진 등글거나, 사각형의 판이 존재하면, 의자이다.

(그림 4)에는 의자를 판단하기 위한 평가 지수들로 구성된 계층적인 망을 도시하였고, 학습 데이터의 학습 결과는 〈표 4〉에 정리하였다. 〈표 4〉의 학습 데이터는 전체 학습 데이터를 열거한 것은 아니고, 표를 간단하게 만들기 위해 규칙의 대칭성에 의해 유추할 수 있는 학습 데이터는 생략하였다. 예를 들어 4번째 행의 데이터는 실제  $x_4$  위치뿐 아니라  $x_1$ 부터  $x_4$ 까지 0.9가 하나 입력되고,  $x_5$ 와  $x_6$  중 하나의 입력이 0.9인 8개의 입력을 대표해서 표현한 것이다. (그림 4)와 〈표 4〉에서 각 단의 점속 연산자는 다리 2개 이상 존재해야 하기 때문에 op-1은 보상 연산자, 등글거나 사각형의 판이 존재하면 되기 때문에 op-2는 합 연산자. op-1에서는 op-2와 op-3의 결과가 모두 만족되어야 의자로 판명되기 때문에 교 연산자가 되어야 합을 예상할 수 있다.

본 예제를 위해서는 알고리즘 2를 사용하여 학습시켰다. 즉 모든 단을 초기에 일반화된 평균이라 가정하고, 학습 도중 일반화된 평균의  $p$  값이  $-2$ 보다 작아지면 Yager의 교 연산자,  $+2$ 보다 크면 Yager의 합 연산자로 전환하여 학습하게 하였다. 학습 결과 예상했던 바와 같이 op-1은 Yager의 교 연산자로 전환되어  $p=-3.44$ 로 수렴하였고, op-2는 Yager의 합이나 교 연산자로 전환되지 않고 일반화된 평균 영역에서  $p=1.83$ 로, 모든 입력의 가중치  $w_i=0.25$ 로 수렴하였고, op-3는 Yager의 합 연산자로  $p=3.65$ 로 수렴하였다. 즉 학습 결과는 예상했던 점속 연산자를 찾고, 이러한 형태의 문제는 규칙 기반을 이용하는 컴퓨터 비전 문제에 제안된 방법을 사용할 수 있음을 알 수 있다.



(그림 4) 의자 문제의 망 구조

(표 4) 의자 데이터

입력						출력	알고리즘1 학습 후
x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>	x <sub>6</sub>	d	
0.1	0.1	0.1	0.1	0.1	0.1	0.01	0.04
0.1	0.1	0.1	0.1	0.1	0.9	0.05	0.17
0.1	0.1	0.1	0.9	0.1	0.1	0.05	0.10
0.1	0.1	0.1	0.9	0.1	0.9	0.05	0.50
0.1	0.1	0.9	0.9	0.1	0.1	0.05	0.13
0.1	0.1	0.9	0.9	0.1	0.9	0.90	0.62
0.1	0.9	0.9	0.9	0.1	0.1	0.01	0.15
0.1	0.9	0.9	0.9	0.1	0.9	0.95	0.74
0.9	0.9	0.9	0.9	0.1	0.1	0.05	0.18
0.9	0.9	0.9	0.9	0.1	0.9	0.98	0.88

V. 결 론

본 논문에서는 퍼지 접속 연산자를 이용한 새로운 정보 융합 방법을 제안하였다. 제안된 방법에서는 평가 지수들의 계층적 체계를 이용하여 [0, 1] 사이의 평가 지수의 만족도들을 전파할 수 있는 망을 구성하고, 역전파 학습을 통하여 망의 각 단계에서 필요한 접속 연산의 형태를 접속 연산자에 부수되는 파라미터를 결정하였다.

즉, Yager의 합 또는 교 연산자와 일반화된 평균 연산자, Zimmermann의  $\gamma$ -model을 계층적 신경 회로망의 활성화 함수를 가정하고 역전파 학습에 의해 필요한 연산자와 파라미터 값을 조정함으로써, 정보 융합시 나타나는 비관적, 낙관적 또는 보수적 성향을 자동적으로 결정한다. 따라서, 퍼지 적분이나 퍼지수 등을 이용하는 정보 융합 방법들에 있어서 임기 응변적으로 가중치를 부가하는 등의 단점이 없으며, Bayes 확률에 의한 방법이나, 믿음 이론의 Dempster의 규칙 적용 등과 같은 제약 조건이 없다.

한편, EXOR 예제에서 보았듯이 제안된 학습 방법은 일반화된 평균이나  $\gamma$ -model과 같은 가중치가 부수되어 있는 연산자를 이용하면 불필요한 정보원으로부터의 가중치를 0으로 수렴시킴으로써 의사 결정에 반영시키지 않는 특성이 있다. 즉 정보원들이 패턴 인식에 있어서 속성들이라면 속성의 수를 감소(feature reduction)시키는 성질을 가지고 있다. 여기서 불필요한 정보원의 제거란 redundancy를 증가시켜서 의사 결정의 강인성을 증가시키는 정보원을 포함시키지는 않는다.

또한, 본 논문에서는 Zimmermann의  $\gamma$ -model이  $\gamma$  값에 따라 단조 증가 함수라는 사실을 증명하였고, 식 (15)와 (16)에서의 같이 가중치 등의 제약 조건있는 최적화 문제를 계층적 망에서 해결하는 간단한 방법을 제시하였다. 제안된 방법의 문제점으로는 일반적인 gradient를 이용하는 역전파 학습으로 인한 극부 최소치 문제와 느린 수렴속도 등을 들 수 있다.

참 고 문 헌

1. Andress, K. M., and Kak, A. C., "Evidence Accumulation and Flow of Control in a Hierarchical Spatial Reasoning System," *Artificial Intelligence Magazine*, pp. 75~94, Summer 1988.
2. Dubois, D., and Prade, H., "A Class of Fuzzy Measure Based on Triangular Norms," *International Journal of General Systems*, Vol. 8, No. 1, pp. 43~61, 1982.
3. Dubois, D., and Prade, H., "A Review of Fuzzy Set Aggregation Connectives," *Information Sciences*, Vol. 36, No. 1&2, pp. 85~121, 1985.
4. Dyckhoff, H., and Pedrycz, W., "Generalized Means as a Model of Compensation Connectives," *Fuzzy Sets and Systems*, Vol. 14, No. 2, pp. 143~154, 1984.
5. Gordon, J., and Shortliffe, H., "A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space," *Artificial Intelligence*, Vol. 26, pp. 323~357, 1985.
6. Keller, J., Hobson, G., Wotton, J., Nafarieh,



A., and Luetkemeyer, K., "Fuzzy Confidence Measures in Midlevel Vision," *IEEE Trans. on SMC*, Vol. 17, No. 4, pp. 676~683, 1987.

7. Kirl, G. J., and Folger, T. A., *Fuzzy Sets, Uncertainty and Information*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

8. Krishnapuram, R., and Lee, J., "Fuzzy-Connective-Based Hierarchical Aggregation Networks for Decision Making," to be appeared at *Fuzzy Sets and Systems*.

9. Krishnapuram, R., and Lee, J., "Propagation of Uncertainty in Neural Networks," *Proceedings of the SPIE Conference on Intelligent Robot and Computer Vision*, Cambridge MA, pp. 377~383, November 1988.

10. Krishnapuram, R., and Lee, J., "Determinating the Structure of Uncertainty Management Networks," *Proceedings of the SPIE Conference on Robotics and Computer Vision*, Philadelphia, 1989.

11. Lee, J. *Fuzzy-Connective-Based Information Fusion Network and Their Application to Computer Vision*, Ph. D. Dissertation, Dept. of Electrical and Computer Engineering, University of Missouri in Columbia, Aug. 1990.

12. Li, z., "Uncertainty Management in a Pyramid Vision System," *International Journal of Approximate Reasoning*, Vol. 3, No. 1, pp. 59~85, 1989.

13. Mithiche, A., and Aggarwal, J. K., "Multiple Sensor Integration/Fusion through Image Processing: A Review," *Optical Engineering*, Vol. 25, No. 3, pp. 380~386, 1986.

14. Mizumoto, M., "Pictorial Representations of Fuzzy Connectives, Part I: Cases of  $t_{\text{norm}}$ ,  $t_{\text{conorm}}$ , and Averaging Operators," *Fuzzy Sets and Systems*, Vol. 31, pp. 217~242, 1989.

16. Pearl, J., "Fusion, Propagation, and Structuring in Belief Networks." *Artificial Intelligence*, Vol. 29, No. 3, pp. 241~288, 1986.

17. Rumelhart, D. E., McClelland, J. M., and the PDP Research Group. *Parallel Distributed Processing*, Chapter 8, Vol. 1, MIT Press, 1986.

18. Thole, U., and Zimmermann, H. -J., "On the Suitability of Minimum and Product Operations for the Intersection of Fuzzy Sets," *Fuzzy Sets and Systems*, Vol. 2, pp. 167~180, 1979.

19. Yager, R. P., "On a General Class of Fuzzy Connectives," *Fuzzy Sets Systems*, Vol. 4, pp. 235~242, 1980.

20. Zimmermann, H. -J., and Zysno, P., "Latent Connectives in Human Decision Making," *Fuzzy Sets and Systems*, vol. 4, No. 1, pp. 37~51, 1980.

21. Zimmermann, H. -J., and Zysno, P., "Decisions and Evaluations by Hierarchical Aggregation of Information," *Fuzzy Sets and Systems*, Vol. 10, No. 3, pp. 243~260, 1983.

22. H. Tahani & J. M. Keller, "Information Fusion in Computer Vision Using the Fuzzy Integral," *IEEE Trans. on SMC*. Vol. 20, No. 3, May/June 1990.



이 준 환

1980년 2월 한양대학교 전자공학과 졸업(공학사)  
 1982년 2월 한국과학기술원 전기 및 전자공학과 졸업(산업전자 전공, 공학석사)  
 1984년 4월~1987년 8월 전북대학교 전자공학과 전임강사  
 1987년 8월~1990년 8월 미국 미주리대학교(콜럼비아) 전기 및 컴퓨터공학과 졸업(공학박사)  
 1990년 10월~현재 전북대학교 전자공학과 조교수  
 관심분야: 불확실성 관리, 영상처리 및 컴퓨터 시각, 패턴인식 등