

GA-FS를 이용한 GAS 식별 시스템 설계

방영근, 심재선, 변형기, 이철희
강원대학교

Design of GAS Identification System using GA-FS

Young-Keun Bang, Jae-Sun Shim, Hyung-Gi Byun, Chul-Heui Lee
Kangwon National University

Abstract - 퍼지이론은 비선형적 데이터의 취급에 있어 강력한 수단이 될 수 있으며, 본 논문에서는 퍼지시스템을 인간의 후각을 모사한 GAS 식별 시스템의 설계에 적용하였다. 먼저, 다수의 센서들의 특성 분석에 따라 GA를 이용하여 그룹화를 시킨 후 각각의 그룹들에서 나타나는 데이터들의 특성에 맞게 퍼지시스템을 설계하여, 센서들의 유동적 특성에 따른 조합과 퍼지시스템의 비선형 데이터에 대한 기술능력을 모두 수용할 수 있는 식별 시스템을 설계하였다. 마지막으로 성능 검증을 통해 하나의 퍼지시스템을 선택함으로써, 유동적 특성이 큰 센서들의 성능을 배제할 수 있도록 하여 보다 정확한 식별이 가능하도록 시스템을 설계 하였다.

1. 서 론

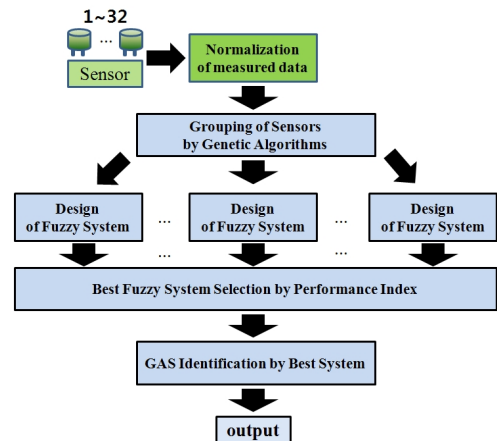
센서에 의해 검출된 신호들은 한 종류의 가스에 대해서도 같은 신호를 발생시키지 않는다. 또한, 같은 종의 센서일지라도 다른 신호를 발생시키기 때문에 센서에서 발생한 신호만으로 가스의 종류를 식별하는 것은 쉽지 않다. 뿐만 아니라 센서들은 시간이 지남에 따라 그 성능이 변화될 수 있기 때문에 센서로부터 측정된 신호들은 비선형적 특성을 포함하게 된다. 따라서 기존의 선형 기법들이나 클러스터링 기법들만으로는 가스 식별에 어려움이 존재할 수 있다 [1]. 그러므로 이러한 비선형 특성의 측정데이터들을 보다 효과적으로 취급할 수 있는 방법이 요구되며, 본 논문에서는 비선형적 데이터의 기술에 효과적인 퍼지시스템을 이용하여 가스의 종류를 판별하는 식별 시스템의 설계법을 다룬다. 먼저 본 논문은 32개의 센서들로부터 측정된 데이터들을 정규화하게 된다. 이는 측정 때마다, 또는 시간간 지나면서 생기는 측정값의 변화를 정규화하는 과정을 의미한다. 이렇게 정규화된 값들을 이용하여 서로 특성이 잘 드러날 수 있는 센서들의 결합으로 8개의 그룹을 생성하게 되며, 이때 유전알고리즘을 이용하게 된다. 이는 32개의 센서들에 대한 8개의 그룹을 결정할 경우 매우 많은 경우의 수가 발생되기 때문이며, 유전알고리즘은 이러한 부분에서 목적함수에 따라 빠른 수렴 특성과 최적해를 탐색함에 있어 강력한 수단이 될 수 있기 때문이다. 이러한 형태로 8개의 센서 그룹이 결정되면, 각각의 그룹에 대하여 센서들이 보여주는 비선형 특성의 측정된 데이터에 대하여 기술능력이 뛰어난 퍼지이론 [2]을 적용하여 8개의 퍼지 식별 시스템을 설계하게 된다. 퍼지 시스템의 설계에는 TSK 퍼지 모델을 사용하였으며, TSK 퍼지 모델의 전부분 규칙기반의 생성을 위해선 구조가 간단한 K-means 군집화 기법을 적용하고, 후반부의 선형식의 파라미터 추정에는 선형 추정기법으로 넓게 사용되는 최소자승기법을 사용하였다. 마지막으로 설계된 각각의 퍼지 식별 시스템들의 성능을 평가하고, 평가된 결과를 이동하여 최상의 동작을 하는 퍼지 식별 시스템의 출력력을 제한된 식별시스템의 최종 출력으로 사용함으로써 특성이 우수하지 않는 센서 조합으로부터 야기되는 성능저하를 피할 수 있도록 하였다.

2. 본 론

2.1 제한된 식별 시스템의 구조

그림 1은 제한된 가스 식별 시스템의 구조를 보여준다. 제한된 식별 시스템의 구조는 크게 데이터의 정규화, 유전알고리즘에 의한 센서들의 그룹화, 각 그룹에 상응하는 퍼지 식별 시스템의 설계, 마지막으로 성능 평가에 의한 베스트 퍼지 식별 시스템의 선택으로 이루어져 있다. 데이터의 정규화 과정은 가장 간단한 방법으로 32개의 센서들로부터 측정된 값의 합이 일정 값을 유지할 수 있도록 하는 구조로 되어 있으며, 센서들의 그룹화에는 실수코딩 유전알고리즘을 적용하였으며, 알고리즘의 목적함수로는 가스의 종류별 측정데이터의 평균에 대한 편차와 이에 대한 측정데이터들의 교차상관함수를 적용하여 강상관성과 역상관성의 특성에 따른 패턴 분석의 정확성을 유도할 수 있도록 하였다. 또한 퍼지 식별 시스템의 설계에는 TSK퍼지 모델, K-means클러스터링 그리고 최소

자승 파라미터 추정기법을 적용하여 구조가 간단하면서도 효율적인 시스템이 설계될 수 있도록 하였다.



〈그림 1〉 제한된 가스 식별 시스템의 구조

2.2 식별 시스템의 설계

본 논문의 전체 식별 시스템의 구현은 GA를 이용한 센서들의 그룹화 과정과 각 그룹에 상응하는 퍼지 식별시스템의 설계로 구성되어 있다.

2.2.1 GA를 이용한 센서 그룹 생성

32개의 센서들로부터 측정된 데이터들을 모두 퍼지 시스템의 구현을 위해 사용하게 된다면 이는 TSK 퍼지 모델의 전부분 규칙기반에 이용되는 입력데이터의 차원이 32차원이 되는 것을 의미하므로 엄청난 양의 규칙기반이 형성될 것이며, 또한 성능이 떨어지는 센서들의 측정값이 반영되어 TSK 퍼지모델의 파라미터 추정에 대한 부정확성을 초래할 수 있다. 따라서 4개의 센서들로 하나의 그룹을 결성하고 이를 이용하여 퍼지 시스템을 설계하게 되면 입력공간의 차원을 줄일 수 있게 된다. 이때 패턴의 구분이 명확할 수 있고 또한 서로 유사한 특성을 보여주는 센서들끼리 그룹을 형성할 수 있다면 후에 설계되는 퍼지 식별시스템 규칙의 정확도와 파라미터 추정 시 발생하는 부정확성을 줄일 수 있을 것이다. 따라서 본 논문에서는 그룹화를 위해 사용된 실수코딩 유전알고리즘의 목적함수로 다음과 같이 가스의 종류에 따라 측정된 데이터들의 평균에 대한 표준편차와 또한 그에 따른 교차 상관함수에 따른 상관성을 높일 수 있는 쪽으로 사용하였다.

$$\rho_{MN}^g = \frac{std(C_{MM}^g) C_{MN}^g}{\sqrt{C_{MM}^g} \sqrt{C_{NN}^g}} \quad (1)$$

여기서, ρ_{MN}^g 는 g 번째 센서 그룹에 해당하는 목적함수의 계수를 의미하고, C_{MN}^g 는 g 번째 가스 센서그룹에서 측정값들의 평균과 각각의 측정값들에 대한 교차 공분산을 의미한다. 또한, C_{MM}^g 와 C_{NN}^g 는 g 번째 가스 센서 그룹에서의 평균에 대한 공분산과 측정 데이터들의 각각의 공분산을 의미한다. 마지막으로 $std(C_{MM}^g)$ 은 평균들의 표준편차이다. 따라서 식 (1)에 의한 그룹화 과정을 거치게 되면 평균의 편차가 크고 여기에 강상관성을 보이거나 강한 역상관성을 보일 경우 보다 명확하게 데이터의 패턴을 분석할 수 있게 된다. 이러한 특성들은 추후 설계되는 퍼지 시스템의 정확성을 높일 수 있게 된다. 또한 본 논문에서는 수정 단순교배, 동적 돌연변이 연산자 그리고 엘리트 전략을 사용하여 알고리

즘의 수렴성과 최적해 탐색 성능을 개선할 수 있도록 하였다.

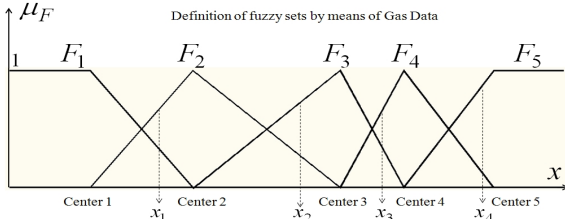
2.2.2 퍼지 식별 시스템의 설계

TSK 퍼지 모델의 j 번째 규칙의 일반식은 다음과 같다.

$$R^j: \text{if } x_1 \text{ is } F_1^j \text{ and } x_2 \text{ is } F_2^j \text{ and } \dots \text{ and } x_p \text{ is } F_p^j \quad (2)$$

$$\text{then } y^j(x) = c_0^j + c_1^j x_1 + c_2^j x_2 + \dots + c_p^j x_p$$

여기서, x 는 입력 데이터로 센서들의 측정값을 의미하게 되며, p 는 입력의 개수로 본 논문에서는 하나의 그룹에 4개의 센서 입력이 있으므로 4로 정의된다. 또한 c 는 추정되어야 할 파라미터이며, y 는 출력을 의미한다. 식(2)을 보면 TSK퍼지 규칙기반을 생성하기 위해 입력공간의 퍼지 분할이 요구되며, 본 논문에서는 아래와 같이 삼각형 소속함수를 이용하여 5개의 퍼지집합으로 입력 공간을 분할하였다.



〈그림 2〉 제안된 시스템의 퍼지 분할의 예

그림 2와 같은 퍼지 분할을 위해 본 논문에서는 K-means 군집화 기법을 사용하였으며, 초기 각 퍼지집합의 중심은 5종류의 가스들의 측정값의 평균값으로 주어지게 된다. 따라서 그림 2의 경우 아래와 같은 형태로 퍼지 규칙을 생성할 수 있게 된다.

$$R^1 : x_1 \text{ is } F_1 \text{ and } x_2 \text{ is } F_2 \text{ and } x_3 \text{ is } F_3 \text{ and } x_4 \text{ is } F_4 \quad (3)$$

$$R^2 : x_1 \text{ is } F_1 \text{ and } x_2 \text{ is } F_2 \text{ and } x_3 \text{ is } F_3 \text{ and } x_4 \text{ is } F_5$$

$$\vdots$$

$$R^j : x_1 \text{ is } F_2 \text{ and } x_2 \text{ is } F_2 \text{ and } x_3 \text{ is } F_4 \text{ and } x_4 \text{ is } F_4$$

$$\vdots$$

$$R^M : x_1 \text{ is } F_2 \text{ and } x_2 \text{ is } F_3 \text{ and } x_3 \text{ is } F_4 \text{ and } x_4 \text{ is } F_4$$

결국 하나의 입력 집합으로부터 1개에서 최대 16개의 퍼지 규칙을 생성할 수 있으며, 규칙의 생성과정에서 중복되는 규칙은 삭제하면서 모든 입력에 대한 퍼지규칙을 생성하게 된다. 퍼지 규칙 생성이 끝나면 각각의 퍼지 규칙의 전반부를 만족하는 모든 입력 데이터들을 이용하여 후반부의 파라미터를 식별하게 된다. 먼저 j 번째 퍼지 규칙 R^j 를 만족한 입력 데이터들이 n 개라면 다음과 같이 n 개의 연립방정식을 표현할 수 있다.

$$y_1^j = c_0^j + c_1^j x_{11} + c_2^j x_{12} + c_3^j x_{13} + c_4^j x_{14} \quad (4)$$

$$y_2^j = c_0^j + c_1^j x_{21} + c_2^j x_{22} + c_3^j x_{23} + c_4^j x_{24}$$

$$\vdots$$

$$y_n^j = c_0^j + c_1^j x_{n1} + c_2^j x_{n2} + c_3^j x_{n3} + c_4^j x_{n4}$$

여기서 y_n^j 는 각 규칙들이 식별하고자 하는 가스의 종류에 해당하므로 그림 2의 퍼지집합들의 중심값들이 적용되며, 따라서 파라미터들은 최소자승법을 이용하여 다음과 같이 추정할 수 있게 된다.

$$\hat{C}^j = ((X^j)^T X^j)^{-1} (X^j)^T Y^j \quad (5)$$

여기서 \hat{C}^j 는 추정되어야 할 파라미터 벡터, X 는 입력벡터 그리고 Y 는 출력벡터를 각각 의미한다. 이렇게 파라미터 추정이 끝나면 각 입력에 대한 가스의 종류를 식별하게 되는데, 이때 식별을 위해 추정되는 출력은 입력데이터가 만족하는 퍼지 규칙의 소속함수 f^j 와 그때의 출력값 y^j 의 가중합으로부터 유도될 수 있다. 따라서 식 (2)와 같이 하나의 입력은 여러 개의 퍼지 규칙을 생성할 수 있으므로 이 입력이 M 개의 퍼지 규칙을 만족하였다면, 그때의 출력은 다음과 같이 정의된다.

$$y_{TSK,1}(X) = \frac{\sum_{j=1}^M f^j(X) (\hat{c}_0^j + \hat{c}_1^j x_1 + \hat{c}_2^j x_2 + \hat{c}_3^j x_3 + \hat{c}_4^j x_4)}{\sum_{j=1}^M f^j(X)} \quad (6)$$

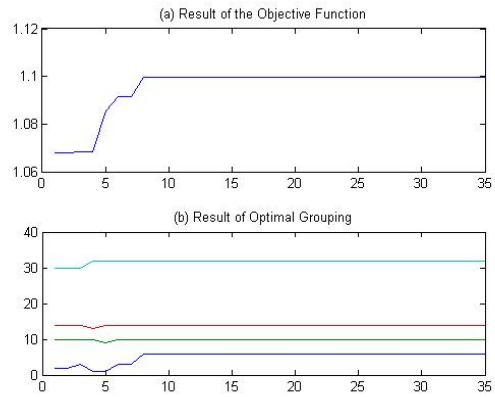
마지막으로 추정된 출력과 각각의 퍼지집합과의 거리를 연산하여 가

장 근접한 퍼지집합 중심에 해당하는 가스를 식별된 값으로 출력하게 된다.

이러한 출력의 결과들은 설계된 퍼지 예측시스템이 8개가 되므로 8개의 판별 성능을 가지게 될 것이며, 마지막으로 퍼지 식별 시스템의 훈련과정에서 가장 우수한 식별 성능을 보인 그룹에 해당하는 퍼지 시스템을 제안된 시스템의 최종 출력 값으로 선택하여 가스의 종류를 식별하게 된다.

3. 시뮬레이션

그림 3은 최적 센서조합의 첫 번째 그룹화 과정에서 35세대까지 진화시킨 유전알고리즘의 목적함수의 변화와 센서조합의 변화를 보여준다.



〈그림 3〉 GA를 이용한 센서그룹화 과정

표 1은 설계된 퍼지 식별시스템에 대하여 학습데이터를 이용하여 가스의 종류를 식별한 결과를 보여준다.

〈표 1〉 퍼지 식별 시스템의 성능과 최종 출력

식별 단계	가스 종류	퍼지 시스템(성능 : %)							
		FS1	FS2	FS3	FS4	FS5	FS6	FS7	FS8
퍼지 시스템 성능	아세톤	100	36.37	56.67	60	80	83.3	86.67	100
	부탄	100	100	100	96.67	100	100	100	100
	메탄	100	36.37	100	90	66.67	93.3	100	93.3
	프로판	96.67	96.67	96.67	86.67	96.67	73.3	93.3	96.67
	물	100	100	100	100	60	63.3	100	100
그룹별 평균		99.3	74	90.67	86.67	80.67	82.67	96	98
시스템의 최종 성능		아세톤	부탄	메탄	프로판	물			
		100	100	100	96.67	100			

표 1을 살펴보면 첫 번째 퍼지시스템, 여덟 번째 퍼지시스템, 일곱 번째 퍼지시스템 순으로 성능이 우수하게 나타났으며, 제안된 식별 시스템의 최종 결과는 첫 번째 퍼지시스템(FS1)의 출력이 된다. 또한, 시스템의 최종 성능을 살펴보면 프로판 가스를 제외한 다른 4종의 가스에 대해선 100% 검출하였음을 알 수 있다. 이는 측정시 마다 조금씩 다르게 나타나는 센서들의 특성을 제안된 시스템이 잘 수용한 결과로 볼 수 있다.

3. 결 론

본 논문은 GA와 퍼지 시스템을 이용한 가스 식별 시스템의 설계법을 다루었다. 센서에 의해 측정되는 가스별 신호의 유동적 특성에 잘 대처하기 위해 비선형 데이터에 대한 기술 능력이 우수한 퍼지시스템과 또한 센서들의 특성에 따른 조합을 통해 시뮬레이션에 나타나듯이 좋은 식별 결과를 얻을 수 있었다.

[참 고 문 헌]

- [1] Ricardo Gutierrez-Osuna, "Pattern Analysis for Machine Olfaction: A Review", IEEE SENSORS JOURNAL, vol. 2, no. 3, pp. 189-202, 2002.
- [2] J. M. Mendel, "Uncertainty, Fuzzy Logic and Signal Processing", Signal Processing, vol. 80, pp. 913-933, 2000.