

## 대규모 가스 센서 어레이에서 중복도의 제거와 확률신경회로망을 이용한 분류

김정도<sup>1</sup> · 임승주<sup>1</sup> · 박성대<sup>1</sup> · 변형기<sup>2</sup> · K.C. Persaud<sup>3</sup> · 김정주<sup>1,+</sup>

### The Classification Using Probabilistic Neural Network and Redundancy Reduction on Very Large Scaled Chemical Gas Sensor Array

Jeong-Do Kim<sup>1</sup>, Seung-Ju Lim<sup>1</sup>, Sung-Dae Park<sup>1</sup>, Hyung-Gi Byun<sup>2</sup>, K.C. Persaud<sup>3</sup>, and Jung-Ju Kim<sup>1,+</sup>

#### Abstract

The purpose of this paper is to classify VOC gases by emulating the characteristics found in biological olfaction. For this purpose, we propose new signal processing method based a polymeric chemical sensor array consisting of 4096 sensors which is created by NEUROCHEM project. To remove unstable sensors generated in the manufacturing process of very large scaled chemical sensor array, we used discrete wavelet transformation and cosine similarity. And, to remove the supernumerary redundancy, we proposed the method of selecting candidates of representative sensor representing sensors with similar features by Fuzzy c-means algorithm. In addition, we proposed an improved algorithm for selecting representative sensors among candidates of representative sensors to better enhance classification ability. However, Classification for very large scaled sensor array has a great deal of time in process of learning because many sensors are used for learning though a redundancy is removed. Throughout experimental trials for classification, we confirmed the proposed method have an outstanding classification ability, at transient state as well as steady state.

**Keywords :** Gas sensors, E-nose, Very large scaled chemical gas sensor array, Redundancy, PNN

#### 1. 서론

1982년 Persaud가 생물학적 후각 시스템을 모사하기 위해 전도성 고분자를 이용한 인공 후각 시스템을 제안한 이래로[1], 전도성 고분자(conducting polymer) 센서, QMB (Quartz Micro-balance) 센서 및 MOS (Metal Oxide Semiconductor) 센서 등 다양한 센서를 사용하는 전자코 시스템이 제안되고 연구된 바 있다 [2-4]. 이러한 전자코 시스템의 경우 수-수십 개의 개별적 센서를 어레이로 구성한 후, 퍼지 알고리즘 및 신경망 등의 신호처리 기술을 이용하여 냄새를 분류한다[5, 6].

생물학적 후각 시스템처럼 다양한 종류의 냄새를 구별하기 위해서는 냄새 분자에 따라 다른 전기적 신호를 발생시키는 많은 종류의 센서 물질들이 필요하며, 또한 많은 센서들의 중복성이 허용되어야 한다. 이러한 생물학적 후각 시스템의 특성을 반영하여 인공후각 시스템을 설계하기 위하여 NEUROCHEM 프로젝트에서는 수십 개의 전도성 고분자를 이용하여 2<sup>16</sup> (65,536)개의 센서 어레이를 작은 크기의 칩 속에 집적하였으며, 수십-수백 개의 센서들이 중복성을 가질 수 있는 구조로 센서 어레이 칩을 성공적으로 설계한 바 있다 [7, 8]. 그러나, 이러한 생물학적 후각 시스템을 모방한 대규모 가스 센서 어레이 시스템은 크게 두 가지의 문제점이 존재한다.

첫 번째, 작은 칩 속에 화학적 고분자 물질을 도핑하여 센서를 만드는 것은, 반도체 IC 기술과는 달리, 불안정한 센서들을 만들어 낸다는 것이다. 생물학적 후각 시스템의 경우 일부 후각 세포들이 불안정해지거나 죽더라도 크게 영향을 받지 않지만, 인공 후각 시스템에서는 이러한 불안정한 센서들이 상당한 영향을 줄 수 있다. 두 번째, 불안정한 센서들을 제거한다 하더라도 나머지 센서들 모두를 다 사용할 수 없다는 것이다. 나머지 센서 모두를 다 사용할 경우, 감당할 수 없는 계산 시간이 필요할 뿐만 아니라, 계산 오차의 누적에 의한 오류를 발생시킬 수 있다. 실제 생물학적 후각 시스템의 경우도, 중복성을 가지는 수많은 수용체가 하나의 사구체에 연결되어 한 개

<sup>1</sup>호서대학교 전자공학과(Department of Electronic Engineering, Hoseo University)

<sup>2</sup>강원대학교 정보통신공학과(School of Electronic, Information and Communication Engineering, Kangwon National University)

<sup>3</sup>맨체스터 대학교(SCEAC University of Manchester, UK)

<sup>+</sup>Corresponding author: ichromosome@nate.com

(Received : Mar. 20, 2013, Accepted : Mar. 26, 2013)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

의 신호만을 출력하게 되기 때문에 실제 냄새의 분석을 위해서 수백~수천 개의 사구체 신호만이 이용된다[9-12].

불안정한 센서를 제거하고, 생물학적 시스템과 같이 비슷한 반응의 수용체를 사구체에서 통합하는 시스템을 모사하기 위해, 본 논문에서는 대규모 가스 센서 어레이 신호 중 불안정한 센서 신호들을 검출하고 제거하는 방법을 제안한다. 또한, 중복성을 보이는 센서들을 대표할 수 있는 대표 센서의 선정 방법을 제안한다. 불안정한 센서의 제거를 위해 코사인 유사도 알고리즘을 이용하여 센서간의 유사도 분석을 하였으며, 이를 통해 불안정 센서를 파악하고 제거하는 방법을 제안하였다. 또한, 안정된 센서만을 대상으로 하여 퍼지 클러스터링 알고리즘을 적용하여 중복성을 가지는 센서의 특성을 대표할 수 있는 대표 센서를 선택하는 방법을 제안하였다. 대표센서의 유용성을 입증하기 위하여 대표센서만을 이용한 PCA 분석이 이루어졌으며 매우 좋은 결과를 얻었다.

또한, 본 연구에서는 대표센서만을 사용했을 때의 분류성능을 파악하고, 최적의 분류능력을 얻기 위하여 확률 신경망(Probabilistic Neural Network: 이후 PNN)을 적용하였다. PNN은 Radial Basis Function Layer와 Competitive layer로 구성되어 있으며 최적의 분류기 중의 하나로 알려져 있다. 본 연구에서와 같이 전체 대규모 센서 어레이를 다 사용하지 않고 대표센서만을 사용한다 하더라도, 기존 연구에서 사용된 것 보다 매우 많은 수의 센서를 사용하게 된다. 기존의 전자코 시스템은 대부분 10개 이하의 센서 어레이를 사용하는 것이 보통이며, 많이 사용된다 하더라도 20개 이하인 경우가 대부분이다. 기존 전자코 시스템이 모두 가스가 주입된 후 일정 시간이 지난 정상상태에서 데이터를 추출하여 분석하는데 비해, 본 연구에서는 더 많은 센서 어레이가 사용됨에 따라 과도상태에서도 충분히 인식 및 분류가 가능할 것으로 미리 예측하였다.

이러한 예측을 입증하기 위하여, 가스 주입 후 정상상태에 도달하기 전인 과도상태에서 대표센서만의 데이터를 추출하여 분류능력을 실험하여 보았다. 실험결과, 초기 과도상태에서는 분류에 오류가 존재하지만, 후기 과도상태에서는 100%의 분류가 가능한 것을 확인하였다.

## 2. 생물학적 후각 시스템을 모방한 대규모 가스센서 어레이

유럽의 NEUROCHEM 프로젝트에서는 생물학적 후각 시스템을 모방한 감각기를 만들기 위해  $2^{16}$  (=65,536)개의 전도성 고분자를 갖춘 어레이를 구축한 바 있다. 65,536개의 센싱 요소는 각각  $64 \times 64$ 개로 구성된 16개의 모듈로 나뉘어져 있다. 각 모듈은 병렬로 데이터를 취득하기 때문에, 한 개 모듈이 센서 데이터를 취득하는 시간이 곧 전체 센서 어레이의 데이터 취득시간이 된다. 데이터 취득 시간은 약 1.24초이며, 한 개 센서 데이터의 샘플링 속도는 3.3 kHz이다.

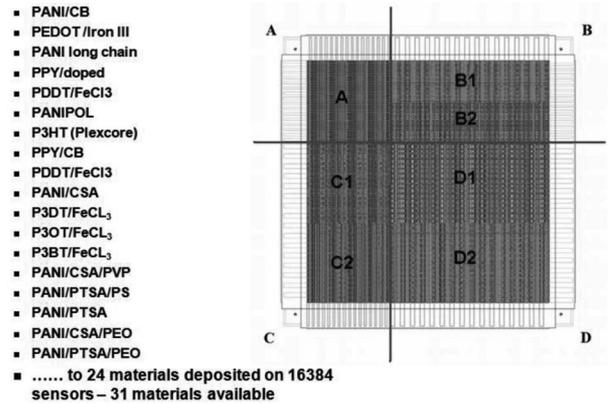


Fig. 1. Designs of the whole sensor die.

각각의 모듈은 센싱 재료에 따라 크게 7개의 영역으로 크게 구분되어 있으며, 24~31개의 전도성 고분자물질들을 사용하여 생물학적 후각 시스템처럼 높은 중복성을 가지도록 설계하였다. 전도성 고분자는 휘발성 화합물을 감지 할 수 있으며, 폭넓게 가스를 감지 할 수 있다. 또한 빠른 응답시간과 낮은 전력, 좋은 감도를 보이기 때문에 전도성 고분자가 대규모 가스 센서 어레이를 위한 재료로 선정 되었다. 센서 어레이를 제작하기 위해서 산화 실리콘 웨이퍼를 사용하여 가공하였다. 센서 어레이는 포토 리소 그래피 기술에 의하여 처리되었다.

## 3. 대표센서 후보의 선정 방법

### 3.1 중복성을 제거한 대표센서 선정의 필요성

Fig. 2에서 알 수 있듯이, 생물학적 후각 시스템의 경우, 비슷한 특성을 가지는 수많은 수용체들이 한 개의 사구체에 연결되어 한 개의 신호만을 출력한다. 즉, 한 개의 사구체와 연결되는 수용체들의 신호는 중복성을 가진다는 것을 의미한다. 생물학적 후각 시스템에서는 많은 중복성을 허용하고, 얼마나 많은 수용체가 흥분하는가에 의해, 후각 수용구의 사구체의 출력이 결정된다. 생물학적 후각 시스템을 모사하고 있는 대규모 센서 어레이 시스템의 경우도, 수십 개의 고분자 물질을 이용하여 비슷한 중복성을 가지도록 설계되었다.

Fig. 3은 대규모 센서 어레이에서 에탄올이 주입되었을 때 108번 센서와 110번의 센서가 거의 비슷한 반응을 보이고 있다는 것을 보여주고 있다. 실제, 본 대규모 센서 어레이 시스템에서 108번 센서와 비슷한 반응을 보이는 센서는 약 30-100개 이상이다. 이러한 센서들은 서로 중복성을 가지고 있다고 할 수 있다. 어느 정도의 중복성 허용은 분류능력을 증대시킬 수 있지만, 지나친 중복성은 오차를 누적시키게 되어 다시 분류 능력을 떨어뜨릴 수 있다. 적절한 중복

성을 유지시키기 위해 지나치게 비슷한 특성을 가지는 센서들을 대표할 수 있는 대표센서를 선정하여 사용하여야 한다.

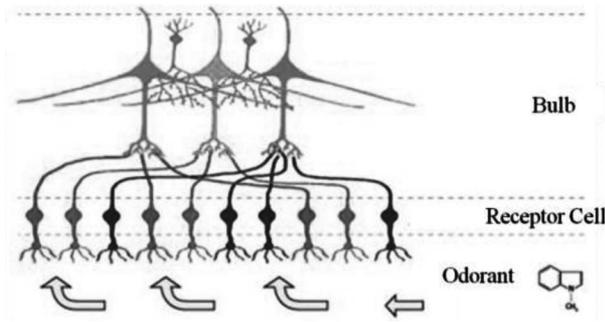


Fig. 2. Redundancy of biological receptors.

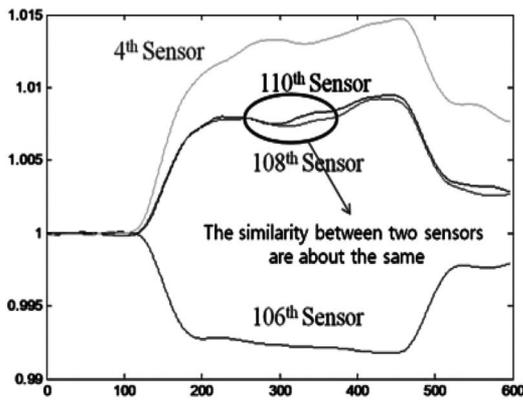


Fig. 3. Example of redundancy at response from ethanol.

### 3.2 대표센서 선정을 위한 실험에 사용된 Data Sets

본 연구의 유효성 입증을 위해 두 종류의 가스와 4가지 농도 그리고 각 농도에 대한 4번의 실험 데이터로 실험을 하였다. 전체 65,536개의 센서 어레이는 16개의 모듈로 나누어져 있다. 본 논문의 실험에서는 16개 모듈의 데이터(L1, L2, L3, ..., L16)군 중 L2의 4,096개의 센서 값을 사용하여 실험을 하였다. 가스는 에탄올과 부탄원을 사용하였다. 에탄올의 경우 1,716, 3,433, 6,867과 13,734 ppm에서 실험이 이루어졌고, 부탄원의 경우 2,240, 4,880, 9,760과 19,520 ppm에서 실험이 이루어졌다. 각 농도별로 10번의 반복 실험이 이루어졌으나, 에탄올은 4차, 5차, 6차, 8차 실험 데이터를 이용하였으며, 부탄원의 경우 3차, 4차, 5차, 7차 실험 데이터를 이용하였다.

### 3.3 신호 측정 잡음 제거를 위한 전 처리 과정

대규모 가스 센서 어레이의 원 신호를 측정할 때 많은 신호잡음이 포함될 수 있다. 이러한 잡음을 제거하기 위하여 디지털 필터인

IIR 필터를 사용하였다. 사용된 IIR 필터는 3차 버터워스 저역통과 필터(BWLPF)이며 0.006 Hz의 차단 주파수가 사용되었다.

Fig. 4의 (a)는 4,096개 센서 중 5번 센서 데이터를 필터 통과 전 데이터, (b)는 필터를 통과한 후의 5번 센서 데이터이며, (c)는 4,096개 센서 데이터 전체가 필터를 통과 한 후의 데이터이다. 그림에서 x축으로 표현된 프레임(frame)은 시간 축을 나타내며, 1프레임은 1.24초를 나타낸다.

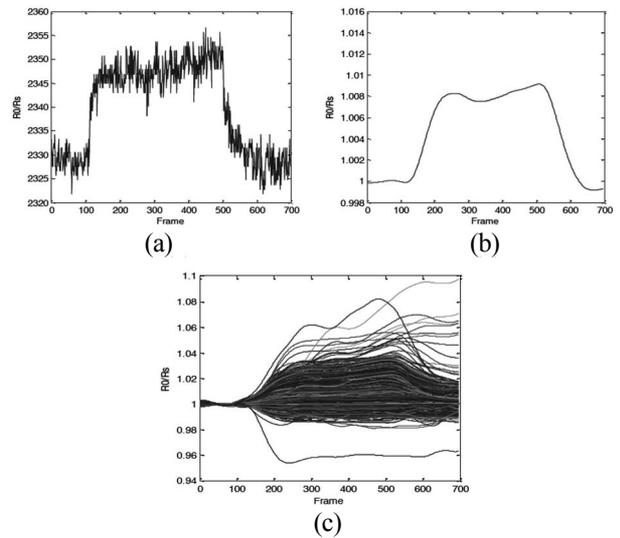


Fig. 4. The result after 3rd order Butterworth filtering; (a) 5th sensor, (b) 5th sensor after 3rd order BWLPF, and (c) 4,096 sensors after 3rd order BWLPF.

### 3.4 대표 센서 후보의 선정

특정 농도의 가스에서 측정된 대규모 센서 데이터를 이용하여, 중복성이 제거된 대표센서후보(The candidates of representative sensor : CRS)를 선정하기 위하여, 본 연구에서는 에탄올 1,760 ppm의 데이터셋 중에서 4번째 데이터셋을 사용하였다. 먼저 cosine 유사도를 이용하여 이 데이터셋에서 존재하는 불안정한 센서를 제거하였다. 안정한 센서만을 대상으로 하여 FCM 알고리즘을 이용하여, 비슷한 중복도를 가지는 센서들을 대표할 수 있는 대표센서의 후보를 선정하였다[18].

#### 3.4.1 코사인 유사도를 이용한 불안정한 센서의 제거

고분자 화학물질이 작은 칩 속에 도핑되어야 하는 대규모 가스 센서 어레이 시스템의 경우, 모든 센서가 실질적으로 안정한 동작을 하는 것은 거의 불가능하다. 65,535개의 센서 어레이가 사용될 경우, 수십~수백 개 이상의 센서가 불안정하게 동작할 가능성이 존재한다.

실제 생물학적 후각 시스템의 경우에서도, 죽어 있거나 이상 동작하는 수용체가 존재하기는 하지만 충분한 중복성의 허용 때문에

사구체의 출력이 영향 받지 않는다. 하지만 아직 생물학적 후각 시스템을 제대로 모사하지 못하는 인공 후각 시스템의 경우에서, 이러한 불안정 센서는 시스템의 특성을 혼란시키는 중요한 요인이 될 수 있다. Fig. 5는 대규모 가스 센서 어레이에 존재하는 불안정 센서의 예를 보여주고 있다. 약 100 프레임 부근에서 가스를 인가하고 500 프레임 부근에서 클리닝을 시작하였을 때, 7번째 센서의 경우 상당히 안정한 동작을 보이고 있으나 40번째 센서의 경우 불안정한 동작을 보임을 알 수 있다.

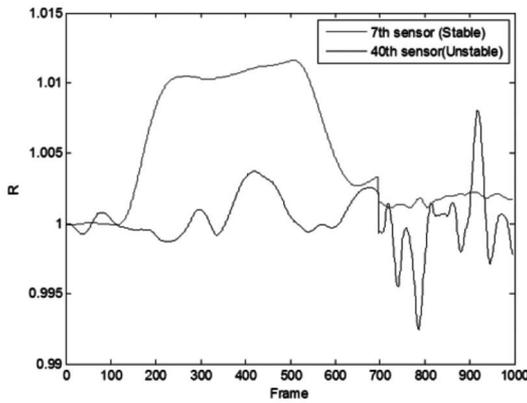


Fig. 5. Examples of unstable sensor.

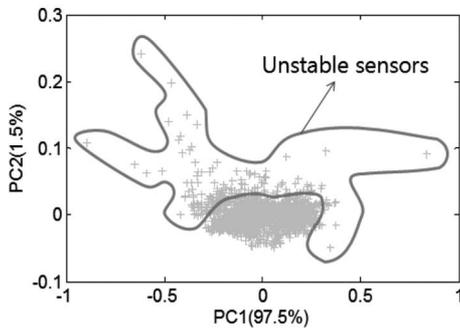


Fig. 6. Unstable sensors of 4096 sensors.

대규모 센서 어레이에는 약 수십~수백 개의 불안정 센서가 존재할 수 있기 때문에 불안정한 센서를 파악하고, 이를 분석과정에서 제거하는 것은 매우 중요하다.

Fig. 6은 PCA 분석을 통해 2차원으로 축소된 불안정한 센서를 보여주고 있다. 중첩성을 보이지 않고 별도로 떨어져 있는 특성을 보이는 센서들의 경우, 대부분 불안정한 센서들이며, 이는 실험을 통해 확인하였다. 본 연구에서는, 4,096개의 센서 중에서 외곽에 존재하는 불안정한 센서의 검출을 위해 이산 웨이블릿 변환과 코사인 유사도 분석을 결합한 방법을 이용한다[13-15].

불안정한 센서를 제거하는 방법은 다음과 같다.

1) 6-step DWT (Discrete Wavelet Transformation)를 통해

전체 데이터를 압축시킨다. 6-step DWT를 통해 얻어진 유사 계수(approximation)는 저주파 특성을 간직하게 된다. 또한, 4096개 센서 각각이 700 프레임이면 10 프레임으로 압축되며, 다음 식으로 표현된다.

$$A_i = (a_{i,1}, a_{i,2}, \dots, a_{i,n}), i = 1, 2, \dots, M, \quad (1)$$

여기서  $M$  은 4,096,  $n$  은 압축된 프레임 수

2)  $CNT = [cnt_1, cnt_2, cnt_i, \dots, cnt_M]$ 를 모두 0으로 설정한다.

여기서  $cnt_i$ 는  $i$ 번째 센서와 유사한 특성을 가지는 센서의 개수를 의미한다.

3)  $i = 1$ 로 놓는다.

4)  $j = i + 1$ 로 설정한다

5)  $A_i$  와  $A_j$  사이에 코사인 유사도 분석을 행한다. 여기서  $j$  는  $j = i + 1$ 이다. 코사인 유사도 분석 알고리즘의 표현 식은 다음과 같다.

$$\begin{aligned} \cos(A_i, A_j) &= \frac{A_i \cdot A_j}{\|A_i\| \|A_j\|} = \frac{(a_{i,1} \cdot a_{j,1} + a_{i,2} \cdot a_{j,2} + \dots + a_{i,N} \cdot a_{j,N})}{\sqrt{a_{i,1}^2 + a_{i,2}^2 + \dots + a_{i,N}^2} \sqrt{a_{j,1}^2 + a_{j,2}^2 + \dots + a_{j,N}^2}} \\ &= \frac{\sum_{k=1}^N (a_{i,k} a_{j,k})}{\sqrt{\sum_{k=1}^N a_{i,k}^2} \sqrt{\sum_{k=1}^N a_{j,k}^2}} \end{aligned} \quad (2)$$

코사인 유사도는 두 개의 데이터가 일치할수록 0에 가까워지며 불일치할수록 1에 가까워진다.

6)  $\cos(a_i, a_j) < S_{threshold}$  이면 8)번으로 이동하고, 그렇지 않으면 7)번으로 이동한다.

7)  $i$  번째 센서와 유사한 특성을 가지는 센서의 개수  $cnt_i$ 와  $j$  번째 센서와 유사한 특성을 가지는 센서의 개수  $cnt_j$ 를 다음과 같이 설정한다.

$$cnt_i = cnt_i + 1, cnt_j = cnt_j + 1$$

8)  $j = j + 1$ 로 설정한 후에 만약  $j \leq M$  이면 5)번으로 이동하고 그렇지 않으면 9)번으로 이동

9) 만약  $cnt_i < SRN$  이면  $i$  번째 센서를 불안정한 센서로 설정하고, 그렇지 않으면 안정한 센서로 설정한다. 여기서 SRN은 안정한 센서인지 불안정한 센서인지를 결정할 기준중복회수 (standard redundant number)로서 사용자에게 의해 결정할 수 있다. SRN이 크다면 불안정한 센서의 수가 많아지게 되며, 사용자에게 의해 실험적으로 설정된다.

10)  $i = i + 1$ 로 설정한 후에 만약  $j = M$  이면 11)번으로 이동하고 그렇지 않으면 4)번으로 이동

11) 안정한 센서의 번호와 안정한 센서의 수를 저장하고, 알고리즘을 완료한다.

불안정한 센서를 결정하기 위한 알고리즘의 첫 번째 단계에서 6-step DWT를 취하고 있는데, 이는 계산 시간을 줄이기 위해서

사용되었다. 센서의 수가 많기 때문에, 센서 특성을 유지하면서도 프레임 수를 줄여 전체 계산시간을 줄일 필요가 있다.

SRN의 값에 따라 불안정 센서가 결정되기 때문에 신중한 선택이 요구된다. 따라서 본 논문에서는 SRN의 값을 변화를 주면서 적절한 값을 정하는 실험을 진행 하였으며, SRN에 의해 불안정 센서가 제거된 가스 데이터를 PCA로 표현하여 Fig. 7에서 나타내었다.

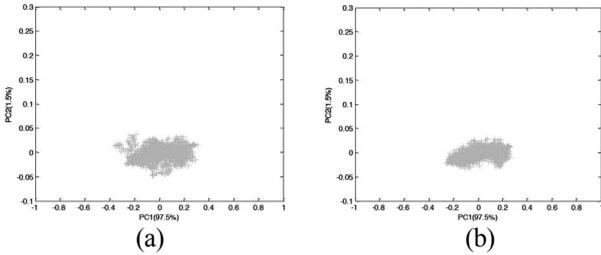


Fig. 7. Removal of unstable sensors according to SRN; (a) SRN=10 and (b) SRN=30.

### 3.4.2 FCM 알고리즘을 이용한 대표센서 후보 선정

Fig. 7에 나타난 것처럼 안정된 센서는 많은 중복성이 있다. 이 중복은 생물학적 후각 시스템에서도 존재하며, 사구체에서 하나만이 출력으로 결정된다. 이러한 중복성을 보이는 특성 중 하나의 특성을 선택하여 출력으로 나타낼 필요가 있다. 이를 위해 안정한 센서 N개를 M개의 그룹으로 클러스터링하여, 각 클러스터의 중심과 가장 가까운 센서를 대표센서 후보로 선정하였다. 각각의 클러스터에 존재하는 센서는 서로 비슷한 특성을 가지고 있기 때문에 중복성을 가지고 있게 되며, 대표센서 후보는 중복성을 가지는 센서들의 대표 특성을 보유하게 된다. 클러스터 센터를 너무 많이 설정한다면 각 클러스터 센터 사이에 중복이 다시 존재하게 된다. 또한, 클러스터의 수를 너무 적게 설정한다면, 센서의 특성이 다른 센서를 하나의 클러스터로 설정 할 수 있기 때문에 결과에 안 좋은 영향을 준다. 클러스터를 분할하는 클러스터링 알고리즘은 여러 가지가 존재하지만, 본 논문에서는 Fuzzy c-means algorithm (FCMA)을 사용하였다. FCM 알고리즘은 클러스터 센터와 인접 센서간의 소속도를 구할 수 있기 때문에 다른 알고리즘보다 쉽게 대표센서 후보를 선정할 수 있다[16, 17]. FCM을 이용하여 대표센서 후보를 선정하는 알고리즘은 다음과 같다.

- 1) 대표센서후보 선정 알고리즘을 통해 구해진 안정한 센서번호 S\_NUM = {s\_1, s\_2, ..., s\_i, ..., s\_R}을 불러온다. 여기서 s\_i는 R개의 안정한 센서 중에서 i번째 안정한 센서의 센서 번호를 의미한다. 만약, R이 3,000이면 총 3,000개의 안정한 센서가 존재하는 것을 의미하며, 이중에서 s\_100 = 3950은 100번째 안정한 센서가 3,950번 센서임을 의미한다.
- 2) 안정한 센서의 데이터를 불러온다.

$$\text{즉, } \begin{bmatrix} A_{s_1} \\ A_{s_2} \\ \vdots \\ A_{s_i} \\ \vdots \\ A_{s_R} \end{bmatrix} = \begin{bmatrix} a_{s_1,1} & a_{s_1,2} & \cdots & a_{s_1,j} & \cdots & a_{s_1,N} \\ a_{s_2,1} & a_{s_2,2} & \cdots & a_{s_2,j} & \cdots & a_{s_2,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{s_i,1} & a_{s_i,2} & \cdots & a_{s_i,j} & \cdots & a_{s_i,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{s_R,1} & a_{s_R,2} & \cdots & a_{s_R,j} & \cdots & a_{s_R,N} \end{bmatrix} \quad (3)$$

여기서  $A_{s_i}$ 는 4,096 센서 중  $s_i$ 번째 센서 데이터를 의미하며, DWT에 의해 압축된 N개의 요소를 가진다.  $A_{s_i}$ 의  $a_{s_i,j}$ 는  $s_i$ 번째 센서 데이터 중에서 j번째 x축 프레임의 요소를 의미한다.

- 3)  $A_{s_i} = [a_{s_i,1}, a_{s_i,2}, \dots, a_{s_i,N}]$ 에서 가스가 주입된 후, 정상상태의 프레임 4개를 선정한 후, R개의 센서데이터를 P개로 군집화하는 FCM알고리즘을 시행하여 클러스터 센터  $C = [c_1, c_2, \dots, c_i, \dots, c_p]$ 를 구한다.
- 4) FCM에 의해 결정된 각각의 클러스터센터  $c_i$ 와의 소속도를 검사하여 소속도가 가장 큰 센서를 대표센서 후보  $R = [r_1, r_2, \dots, r_i, \dots, r_p]$ 로 결정한다. 여기서  $r_i$ 는 P개의 대표센서후보 중 i번째 대표센서후보의 센서번호를 의미한다.

Fig. 8은 FCM 알고리즘을 사용하여 불안정 센서가 제거 된 가스 센서의 데이터에서 클러스터의 중심이 되는 센서를 선택하여 나타내었다. 선택된 대표 센서는 비슷한 특성을 보이는 센서 데이터의 중복성을 하나의 대표 특성을 통해 나타낼 수 있다.

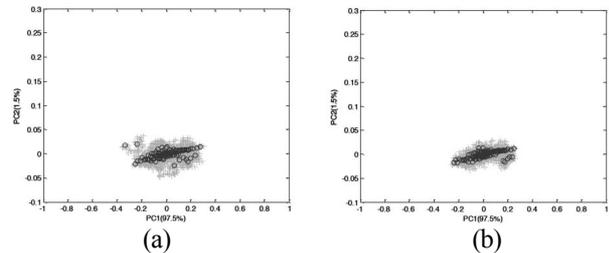


Fig. 8. Extraction of CRS using FCM Algorithm (NCRS=100); (a) SRN=10 and (b) SRN=30.

FCM알고리즘을 사용할 경우, 만약 선정된 대표 센서가 일정시간이 지난 후 불안정해지거나 죽을 경우 기존 대표센서와 소속도가 높은 인접센서로 대표센서를 재선정할 수 있다. 다양한 클러스터링 알고리즘 중에서 FCM알고리즘이 사용되어야 하는 이유가 여기에 있다. 대부분의 화학적 센서는 염소와 같은 물질에 반응하거나 사용시간이 지나면 불안정해지는 경향이 있기 때문에 인접센서로 대표센서를 재지정하는 것은 매우 중요한 문제일 수 있다.

Fig. 9의 (a)는 에탄올 3,433 ppm의 4차 실험 데이터로서 4,096개의 센서를 6차 이산 웨이블릿 변환을 사용하여 시간 축을 10 프레임으로 압축해 놓은 그림이다. (b)는 그림(a)에서 SRN을 100으로 하여 불안정한 센서를 제거한 후, FCM을 통해서 100개의 대표 센서를 추출한 그림이다.

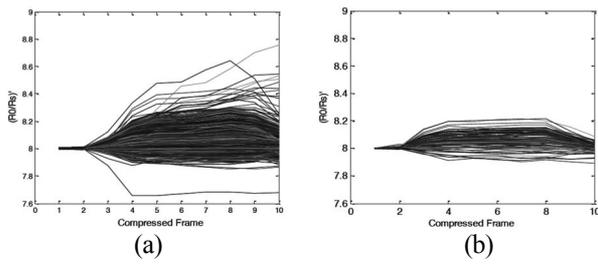


Fig. 9. The comparison on time base (1/64 Compression); (a) 4,096 sensors and (b) Representative sensors.

### 3.4.3 대표센서 후보를 사용했을 때의 유효성 검증

에탄올 3,433 ppm의 4차 실험 데이터를 이용하여 대표센서 후보를 선정할 후, 다른 데이터셋 역시 모두 선정된 대표센서 후보만을 사용하여 유효성을 검증하였다.

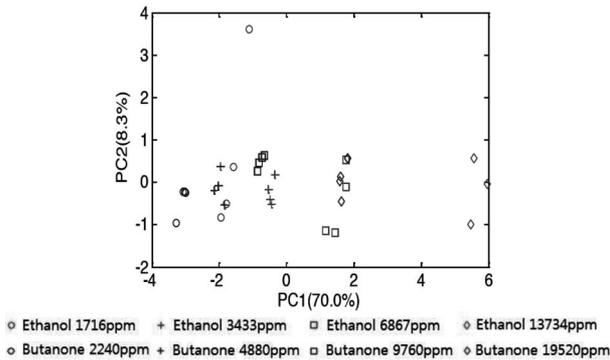


Fig. 10. The result when all sensors (4096 sensors) are used for PCA analysis.

Fig. 10은 4,096개의 센서 전부를 사용하였을 때의 분류 능력을 알아보기 위한 실험 결과이다. 거의 분류가 불가능한 상태를 확인할 수 있다. Figs. 11, 12, 13은 대표센서후보만을 사용하여 실험한 결과로서, 중복도와 대표센서 수가 가스 농도의 분류에 미치는 영향을 나타내었다.

그림은 각각 50, 150 200의 SRN을 적용한 결과이며, 각 그림의 (a), (b), (c)와 (d)는 대표센서후보의 수(The number of candidates for representative sensor: 이하 NCRS)를 각각 50, 100, 200, 300으로 놓고 실험을 한 결과이다.

Figs. 11, 12과 13로부터 대표 센서후보의 수가 결과에 큰 영향을 주었다는 것을 알 수 있다. 각 그림의 (a), (b), (c), (d)를 보면 NCRS가 100 이상 선택되었을 때 좋은 결과를 얻었다. 하지만 NCRS가 50에서는 제한한 방법을 적용하기 전과 차이가 없이 가스 별 분류가 어려웠다. 따라서 대표 센서의 선정은 매우 중요하다. 대표 센서가 적절하게 선택되었을 때는 매우 만족스러운 결과를 얻었다. 그리고 농도가 짙어 질수록 각 가스 별로 뚜렷하게 나누어지는 것을 볼

수 있다. 하지만 몇 가지의 문제점이 발견되었는데, 이는 3.5절에서 설명한다.

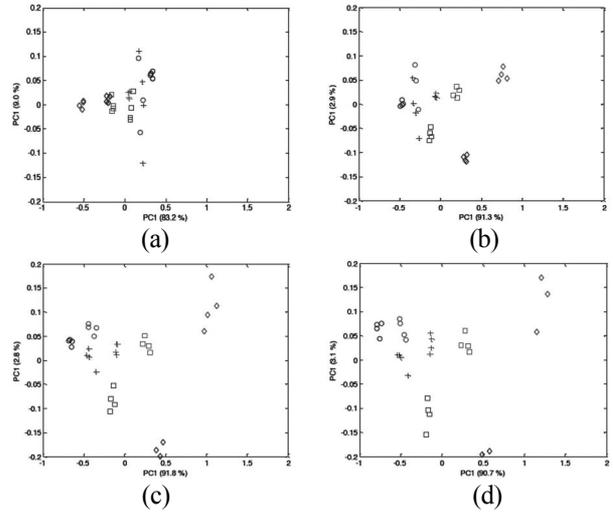


Fig. 11. The result of PCA when SRN is 50; (a) NCRS = 50, (b) NCRS = 100, (c) NCRS = 200, and (d) NCRS = 300.

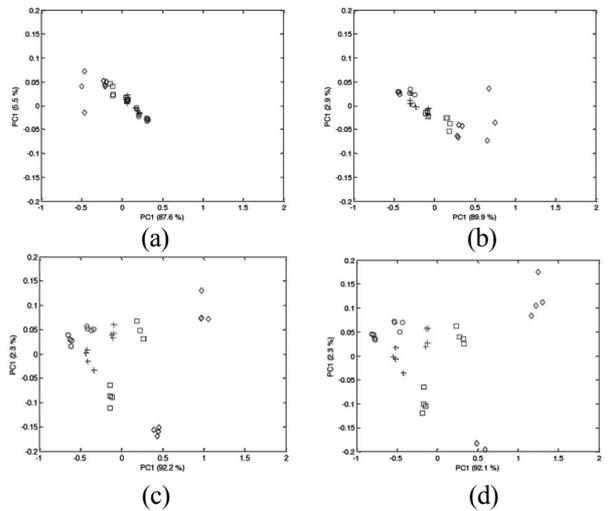


Fig. 12. PCA results in case of SRN=150; (a) NCRS = 50, (b) NCRS = 100, (c) NCRS = 200, and (d) NCRS = 300.

## 3.5 최종 대표 센서 선정 방법

### 3.5.1 선정된 대표센서 후보의 문제점

한 개의 데이터셋을 이용하여 대표센서 후보를 선정하였을 경우, 선정된 대표센서 후보는 다른 가스나 다른 농도에서 측정된 데이터셋에서도 그 대표성이 유지되어야 한다. 선정된 대표센서 후보를 이용하여, 다른 가스와 다른 농도에서 실험한 결과, 대체적으로 좋은 결과를 얻기는 했으나, 분류에 문제가 있는 경우가 발견되었다.

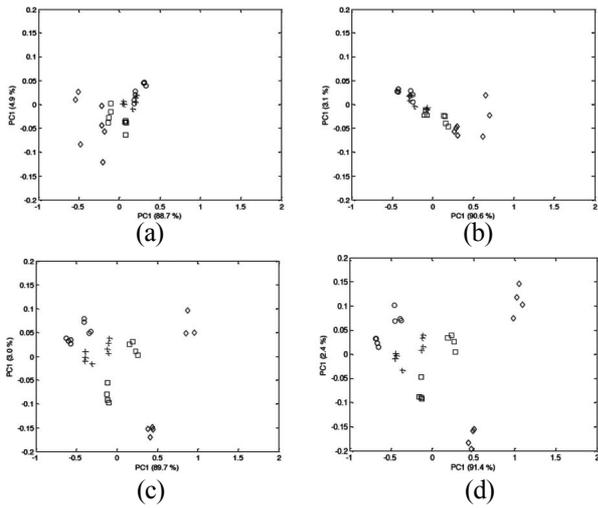


Fig. 13. PCA results in case of SRN=200; (a) NCRS = 50 (b) NCRS = 100, (c) NCRS = 200, and (d) NCRS = 300.

Fig. 11(a), Fig. 12(a), (b)와 Fig. 13(a), (b)의 경우가 분류가 힘든 경우에 해당한다.

이러한 문제에 대한 이유를 알아내기 위해, 본 연구에서는 대표 센서 후보 선정에 사용되지 않은 다른 농도 및 가스의 데이터셋을 대상으로 대표센서 후보들을 비교, 분석하였다. 분석결과, 두 가지의 문제가 발견되었다. 첫 번째, 대표센서 후보에 사용된 가스에서는 안정한 센서였으나, 다른 가스에서는 불안정한 동작을 하는 센서가 존재하며, 이 센서가 대표센서 후보로 선정되었다. 두 번째, 대표센서 후보에 사용된 농도에서는 동작범위에 해당하는 센서였으나, 더 높은 농도에서는 포화된 센서가 존재하였다. Fig. 14는 대표센서 후보 중 다른 가스 혹은 다른 농도에서 불안정한 동작을 보이는 센서를 보여주고 있다.

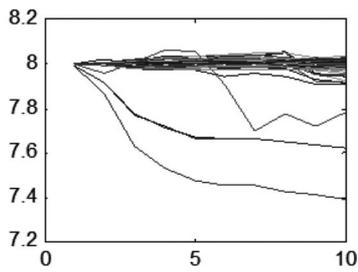


Fig. 14. Unstable sensors of when CRS chosen by ethanol dataset is applied to butanone dataset.

Ethanol 9,670 ppm의 4번째 실험 데이터셋에 의해 추출된 대표센서 후보를 부탄원의 데이터셋에 적용시켜본 결과, 에탄올에서는 안정한 센서였으나 부탄원에서는 불안정하거나 포화된 센서가 존재하는 것을 볼 수 있다.

### 3.5.2 각각의 데이터셋에 존재하는 불안정 및 포화 센서를 제거한 최

### 종대표센서 선정

하나의 데이터셋(dataset)을 통해 선정된 대표센서 후보는 다른 가스 혹은 다른 농도에서도 대부분 안정한 상태를 보이나, 농도가 높아질 경우 포화되는 경우가 존재하며, 다른 가스가 인가되었을 경우 불안정한 반응을 보이는 센서가 존재한다. 이 문제를 해결하기 위해 제안된 알고리즘을 다음과 같이 수정하였다.

- 1) 하나의 데이터셋을 선정 후, 대표센서 후보를 추출한다. 대표센서 후보의 추출 방법은 3.4절에서 설명한 바와 같다.
- 2) 대표센서 후보를 선정된 데이터셋을 제외하고, 다른 가스 및 다른 농도에서 측정된 훈련 데이터셋들을 대상으로 각각의 데이터셋에서 존재하는 불안정 센서 및 포화센서를 추출한다. 지나치게 불안정한 센서 및 포화센서는 다른 센서들과 중첩성이 매우 작기 때문에 중첩도가 5~10 사이에 존재하는 센서를 추출하면 된다.
- 3) 기존에 선정된 대표센서 후보 중에서, 각각의 데이터셋에서 추출된 불안정 센서 및 포화센서들을 제거한다. 불안정 및 포화센서가 제거된 나머지 대표센서 후보를 최종 대표센서 (Representative Sensors : RS)로 결정한다.

Fig. 15은 각각의 데이터셋에 존재하는 불안정 및 포화센서들을 제거하지 않은 대표센서 후보만을 이용한 PCA 결과 중에서 분류가 어려운 것만을 표시하였다.

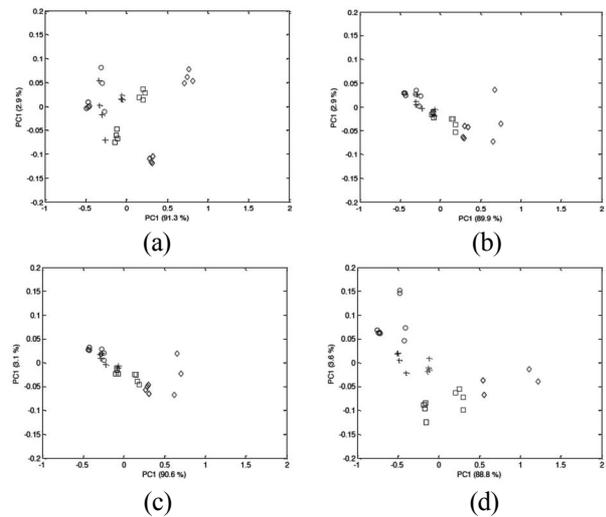


Fig. 15. The case of having some difficulty in classification (When CRS extracted from Ethanol 3,433ppm dataset are applied to other datasets); (a) NCRS=100 (SRN=50), (b) NCRS=100 (SRN=150), (c) NCRS=100 (SRN=200), and (d) NCRS=350 (SRN=200).

Fig. 16은 Fig. 17의 (a), (b), (c)와 (d)에서 각 데이터셋에서 존재하는 불안정 및 포화센서를 제거한 후의 PCA결과를 보여주고 있

다. Fig. 15과 16에서 알 수 있듯이, 불안정 센서 및 포화 센서가 제거된 결과는 모든 데이터셋에서 분류가 가능함을 보여주고 있다. Fig. 16의 (a)에서 나타난 것처럼 76개의 대표센서만을 사용한다 하더라도 좋은 분류 성능을 보여주고 있음을 알 수 있다.

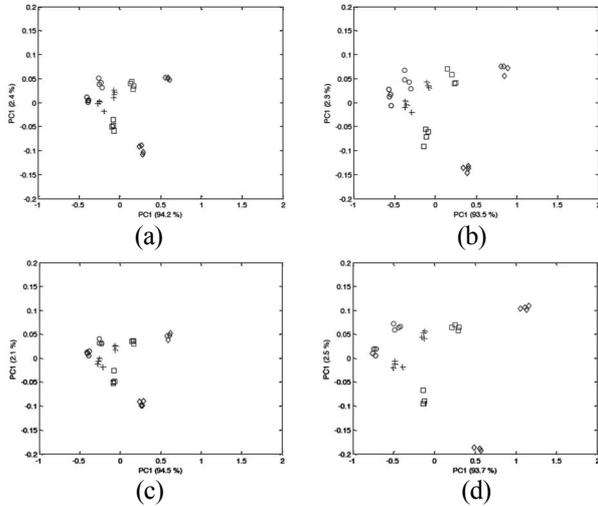


Fig. 16. PCA result (in case that RS eliminating unstable and saturated sensors are used for classification); (a) 76 RS chosen from 100 CRS (SRN=50), (b) 149 RS chosen from 200 CRS (SRN=50), (c) 79 RS chosen from 100 CRS (SRN=200), and (d) 286 RS chosen from 350 CRS (SRN=200).

Figs. 17과 18은 3.4절에서 제안된 방법과 3.5절에서 제안된 방법을 시각적으로 비교하기 위한 그림이다.

Fig. 17은 대표센서후보를 사용하였지만 부탄원에서 존재할 수 있는 불안정 센서를 제거하지 않았기 때문에 같은 가스와 같은 농도에서도 저항변화가 다르게 나타나는 것을 볼 수 있다.

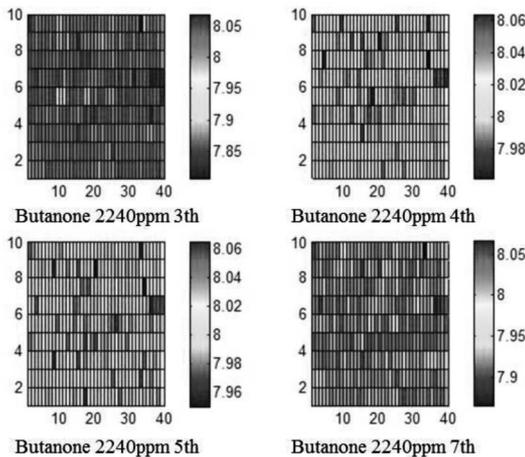


Fig. 17. Resistance distribution of 400 CRS (When 400 CRS extracted from Ethanol 3,433 ppm dataset are applied to butanone datasets).

하지만 Fig. 18은 에탄올 데이터셋에 의해 결정된 대표센서 후보 중에서 부탄원에서 불안정한 대표센서 후보를 제거한 경우로서 같은 가스와 농도에서 거의 비슷한 저항 분포를 가진다는 것을 확인할 수 있다. 이로써 3.5절에서 제안된 방법이 매우 효율적임을 시각적으로 확인할 수 있다.

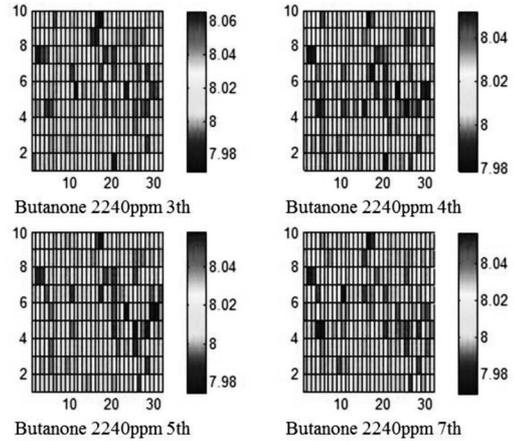


Fig. 18. Resistance distribution of 328 RS (when 328 RS eliminating unstable and saturated sensors from ethanol 3,433 ppm dataset are applied to butanone datasets).

#### 4. 확률신경망을 이용한 분류

PNN (Probabilistic Neural Network)은 1990년 Specht에 의해 제시되었으며, 은닉층의 역할을 하는 Radial basis function (RBF) network layer와 출력층의 역할을 하는 경쟁층(competitive network layer)로 구성된다.

PNN은 각각의 훈련 패턴이 각각의 가우시안 함수에 저장되기 때문에 연상 메모리와 비슷하다. PNN은 패턴의 수와 동일한 뉴런 수를 가져야 하기 때문에 패턴의 수가 많을 경우, 상당한 메모리가 필요하게 된다. 만약 한 개의 입력 패턴이 50개의 데이터로 구성되고, 총 24개의 훈련 패턴이 존재한다면 RBF 네트워크층에서 24 × 50 = 1200개의 메모리 공간이 필요하다. 하지만 별도의 학습 시간이 필요없기 때문에 off-line 학습 없이 실시간으로 전방향 계산에 의해 미지의 데이터의 분류가 가능하다. 또한, 다른 신경회로망과는 달리 국부적 해(local minima)에 빠질 염려가 없는 최적해를 갖는 분류기이다. 만약 국부적 해를 가지는 신경망을 사용할 경우, 대규모 센서 어레이는 입력패턴이 상당히 많기 때문에 분류를 불가능하게 할 수 있다. PNN이 많은 메모리가 필요하다는 점이 문제일 수 있으나 현재의 반도체 기술은 초소형 고속 대규모 메모리 칩이 가능하기 때문에 이 문제는 본 시스템에서는 무시해도 되는 문제라 가정한다.

냄새를 분류하는 경우, 측정환경이나 상황을 측정 시마다 항상 일치화 시킬 수 없다. 센서의 반응은 온도, 습도, 공기 흐름뿐만 아

나라 아주 작은 화학적 변동에 의해서도 달라질 수 있기 때문에 범  
 새의 측정은 반드시 오차가 존재하게 된다. 이러한 문제는 대규모  
 센서를 사용하고, 센서의 반응을 확률적으로 처리하는 것이 가장 좋  
 은 방법일 수 있다. Fig. 19는 PNN의 구조를 보여주고 있다.

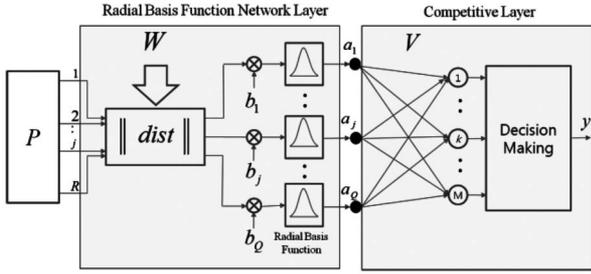


Fig.19. The network structure of PNN.

먼저 한 개의 훈련 입력 패턴은 R개의 데이터로 구성되며,  
 학습을 위한 Q개의 입력패턴이 존재하고 있다고 가정하면, 훈련  
 입력행렬 P는 Q × R 행렬이 된다.

$$P = \begin{bmatrix} p_1^1 & p_2^1 & \dots & p_i^1 & \dots & p_r^1 \\ p_1^2 & p_2^2 & \dots & p_i^2 & \dots & p_r^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ p_1^r & p_2^r & \dots & p_i^r & \dots & p_r^r \\ \vdots & \vdots & & \vdots & & \vdots \\ p_1^Q & p_2^Q & \dots & p_i^Q & \dots & p_r^Q \end{bmatrix} = \begin{bmatrix} P^1 \\ P^2 \\ \vdots \\ P^r \\ \vdots \\ P^Q \end{bmatrix}, \quad i=1,2,\dots,R \quad (4)$$

RBF의 가중치를 W로 가정하면, 식(5)와 같이 정의된다.

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1i} & \dots & w_{1R} \\ w_{21} & w_{22} & \dots & w_{2i} & \dots & w_{2R} \\ \vdots & \vdots & & \vdots & & \vdots \\ w_{j1} & w_{j2} & \dots & w_{ji} & \dots & w_{jR} \\ \vdots & \vdots & & \vdots & & \vdots \\ w_{Q1} & w_{Q2} & \dots & w_{Qi} & \dots & w_{QR} \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_j \\ \vdots \\ W_Q \end{bmatrix} \quad (5)$$

보통 W는 훈련 입력 패턴 P가 주어지면 별도의 학습 없이 W =  
 P로 결정할 수 있다. r번째 훈련 입력 벡터가 인가되었을 때 j번째  
 RBF에서 r번째 훈련벡터와 가중치벡터 사이의 거리는 식(6)과 같  
 이 정의된다.

$$d_j^r = \|dist_j^r\| = \|W_j - P^r\| = \sqrt{\sum_{i=1}^R (w_{ji} - p_i^r)^2} \quad (6)$$

만약 미지의 입력벡터 P<sup>u</sup> = [P<sub>1</sub><sup>u</sup>, P<sub>2</sub><sup>u</sup>, ..., P<sub>R</sub><sup>u</sup>]이 인가되면 j번째  
 RBF에서 미지의 입력벡터와 가중치 행렬 사이의 거리는 식(7)와 같  
 이 정의된다.

$$d_j^u = \|dist_j^u\| = \|W_j - P^u\| = \sqrt{\sum_{i=1}^R (w_{ji} - p_i^u)^2} \quad (7)$$

계산된 입력벡터와 가중치 벡터 사이의 거리 \|dist<sub>j</sub>\|와 바이어  
 스 b<sub>j</sub>를 곱한 값을 RBF의 입력으로 인가한다. 즉, RBF의 입력은  
 식(8)와 같이 정의된다.

$$n_i = \|dist_i\| \times b_i \quad (8)$$

RBF는 보통 가우시안 함수를 사용하며, W와 P사이의 거리가  
 작아질 수록 RBF의 출력은 커지게 된다. W와 P사이의 거리가 0  
 에 도달하면, 출력은 1과 같아지고, 이 값이 최대출력이 된다. 식(9)  
 는 RBF의 출력을 나타낸다.

$$a_i = e^{-n_i^2} = e^{-\|W_i - P\|^2 b_i^2} \quad (9)$$

RBF의 출력이 결정되면, 이 값은 경쟁층으로 입력된다. 경쟁층  
 은 M개의 뉴런으로 구성되는데, M은 클래스의 수로 정의된다. 만  
 약 Q개의 훈련입력패턴이 총 M개의 클래스로 분류되어 있다면,  
 경쟁층의 뉴런은 M개가 된다. 경쟁층의 accumulation neuron에  
 서는 입력패턴이 속한 클래스의 확률을 근사화시키는 역할을 하며,  
 k번째 accumulation neuron의 출력 Q<sub>k</sub>는 식(10)으로 정의된다.  
 여기서 c는 클래스의 수를 나타낸다.

$$Q_k = \sum_{j=1}^Q V_{kj} a_j \quad k=1,2,\dots,c \quad (10)$$

만약, 5개의 훈련입력패턴 [P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, P<sub>4</sub>, P<sub>5</sub>]이 존재하고, P<sub>1</sub>  
 과 P<sub>2</sub>는 1번 클래스, P<sub>3</sub>는 2번 클래스, P<sub>4</sub>와 P<sub>5</sub>는 3번 클래스로  
 구분되어 있다고 가정하면,

$$V = \begin{bmatrix} v_{11} & v_{12} & v_{13} & v_{14} & v_{15} \\ v_{21} & v_{22} & v_{23} & v_{24} & v_{25} \\ v_{31} & v_{32} & v_{33} & v_{34} & v_{35} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad (11)$$

클래스를 결정하기 위한 최종 결정은 Bayes optimal decision  
 rule에 의해 결정되는데, 일단 Accumulation neuron에서 O<sub>1</sub>, O<sub>2</sub>,  
 ..., O<sub>M</sub>을 구한 다음, 이중 가장 높은 확률값을 갖는 뉴런의 번호를  
 출력하면 된다. 즉, 만약 [O<sub>1</sub>, O<sub>2</sub>, ..., O<sub>M</sub>]중에서 3번째 뉴런이  
 가장 큰 확률값을 가지면 3의 값이 출력된다.

### 5. PNN을 이용한 인식 실험 결과

본 연구에서 제안된 대표센서와 PNN을 사용한 분류 성능을 평  
 가하기 위해 중첩도를 50으로 하여 100개의 대표센서 후보를 추출

한 후, 각각의 훈련 데이터셋에 존재하는 불안정 및 포화 센서들을 제거한 76개의 대표센서만을 이용하여 실험하였다.

PNN의 훈련 입력 데이터로서 24개의 데이터셋이 이용되었기 때문에 총 입력 데이터  $P$ 는  $76 \times 24$  행렬이며 목표 데이터  $T$ 는  $24 \times 1$  벡터가 된다.

$$P = [S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{11}, S_{12}, S_{13}, S_{14}, S_{15}, S_{17}, S_{19}, S_{20}, S_{21}, S_{22}, S_{23}, S_{25}, S_{26}, S_{28}, S_{30}, S_{31}, S_{32}]$$

$$T = [1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 7, 7, 7, 8, 8, 8]$$

여기서  $S_1$  은 L2\_EtOH5scm\_4에서 대표센서 76개만을 추출한 후, 버터워스 필터링한 데이터 중에서 특정 시간 프레임의 데이터 76개를 의미하는 훈련입력 벡터이다. Fig. 20은 훈련입력벡터를 보여주고 있는데, 빨간 부분이 데이터 추출을 위한 특정 프레임이다.

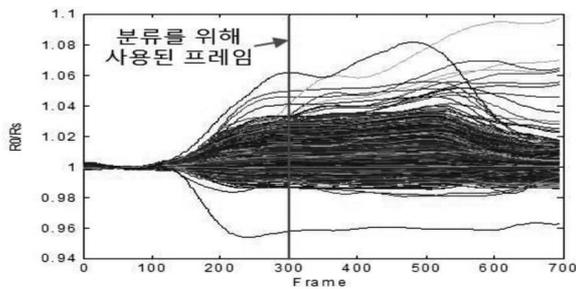


Fig. 20. The frame for extraction of input training vector.

또한, Table 1의 훈련데이터가 인가되어 훈련데이터의 분류가 이루어진 후, 미지의 데이터에 의한 분류 성능을 파악하기 위해 Table 2의 미지의 데이터를 인가하였다.

Table 1. Training data

훈련 입력벡터	훈련 출력벡터	내용
$S_1, S_3, S_4$	1	Ethanol 1716 ppm의 4,6,8차데이터셋에서 추출된 훈련벡터
$S_6, S_7, S_8$	2	Ethanol 3433 ppm의 5,6,8차데이터셋에서 추출된 훈련벡터
$S_9, S_{11}, S_{12}$	3	Ethanol 6867 ppm의 4,6,8차데이터셋에서 추출된 훈련벡터
$S_{13}, S_{14}, S_{15}$	4	Ethanol 13734 ppm의 4,5,6차데이터셋에서 추출된 훈련벡터
$S_{17}, S_{19}, S_{20}$	5	Butanone 2240 ppm의 3,5,7차데이터셋에서 추출된 훈련벡터
$S_{21}, S_{22}, S_{23}$	6	Butanone 4880 ppm의 3,4,5차데이터셋에서 추출된 훈련벡터
$S_{25}, S_{26}, S_{28}$	7	Butanone 9760 ppm의 3,4,7차데이터셋에서 추출된 훈련벡터
$S_{30}, S_{31}, S_{32}$	8	Butanone 19520 ppm의 4,5,7차데이터셋에서 추출된 훈련벡터

Table 2. Unknown data

미지 입력벡터	실제 출력값	내용
$U_1$	1	Ethanol 1716 ppm의 5차 입력벡터
$U_2$	2	Ethanol 3433 ppm의 4차 입력벡터
$U_3$	3	Ethanol 6867 ppm의 5차 입력벡터
$U_4$	4	Ethanol 13734 ppm의 8차 입력벡터
$U_5$	5	Butanone 2240 ppm의 4차 입력벡터
$U_6$	6	Butanone 4880 ppm의 7차 입력벡터
$U_7$	7	Butanone 9760 ppm의 5차 입력벡터
$U_8$	8	Butanone 19520 ppm의 3차 입력벡터

미지의 데이터  $P'$ 는 다음과 같은 순서로 인가하였다.

$$P' = [U_3, U_5, U_8, U_4, U_7, U_6, U_2, U_1]$$

유용성 실험은 크게 두 부분으로 이루어졌다. 한 부분은 정상 상태 즉, 가스 인가 후 충분한 시간이 흐른 후의 frame에서의 실험이며, 다른 부분은 과도상태 즉 가스가 인가 한 후 정상상태에 도달하기 전 프레임에서 데이터를 추출한 실험이다.

정상상태와 과도상태의 기준은 약 200 프레임으로 결정하였다. 100 프레임 부근에서 가스가 인가되었으며, 500 프레임 부근에서 클리닝이 시작되었다. 한 개의 프레임이 약 1.24초이기 때문에 정상상태와 과도상태의 기준인 200 프레임은 가스 주입 후 약 12.4초 후라 할 수 있다. Table 3은 미지의 데이터에 대한 분류 성능결과를 보여주고 있다.

Table 3. The classification result with unknown data

프레임	분류결과
Steady	200 3,5,8,4,7,6,2,1 100% 인식
State	250 3,5,8,4,7,6,2,1 100% 인식
	300 3,5,8,4,7,6,2,1 100% 인식
	350 3,5,8,4,7,6,2,1 100% 인식
	400 3,5,8,4,7,6,2,1 100% 인식
	450 3,5,8,4,7,6,2,1 100% 인식
Transient	190 3,5,8,4,7,6,2,1 100% 인식
State	180 3,5,8,4,7,6,2,1 100% 인식
	170 3,5,8,4,7,6,2,1 100% 인식
	160 3,5,8,4,7,6,7,1 1개 오인식
	150 3,5,3,4,7,1,7,1 3개 오인식

## 6. 결론

본 연구의 목적은 생물학적 후각 시스템을 모사한 방법으로 가스를 분류하는 것이다. 이 목적을 위해, NEUROCHEM 프로젝트에

의해 만들어진 4,096개의 센서 어레이를 대상으로 새로운 신호처리 방법을 제안하였다.

기본적으로 대규모 센서 어레이를 개발하기 위해서는 Conducting Polymer형 센서가 사용되어야 하는데, Conducting Polymer형 센서는 제조과정에서 불안정한 센서를 포함할 수 있다는 단점과 수명이 일정하지 않을 수 있다는 단점을 가진다. 또, 생물학적 후각 시스템은 수백~수천만 개의 수용체 중 비슷한 반응을 보이는 수용체를 모아 사구체에 연결하여 분석에 이용하기 때문에 실제 사용되는 신호는 수백 개 정도라는 생물학적 후각 시스템의 특성에 기인하여, 본 연구에서도 비슷한 반응을 보이는 센서의 대표를 선정하여 사구체의 역할을 대신하게 하는 방법을 사용하였다. 만약 대표센서가 일정 시간이 지난 후 불안정해지거나 죽는 일이 발생한다면 대표센서와의 소속도가 가장 큰 인접 센서를 다시 대표센서로 지정하면 된다. 이러한 기능을 위해 수많은 클러스터링 알고리즘 중에서 FCM의 사용을 제안하였다.

대표센서만을 사용한다 하더라도 기존의 센서 어레이보다 훨씬 더 많은 센서가 사용되고 있으며, 대표센서 간의 특징이 충분히 서로 다르기 때문에, 기존에 수많은 연구에서 시도하였으나 좋은 결과를 얻지 못한 과도상태에서의 분류가 가능할 것으로 예측하였으며, 그 예측을 입증하기 위한 실험에서 그 가능성을 충분히 입증하였다.

PNN이 분류 성능의 분석을 위해 사용되었다. PNN은 훈련데이터만을 필요로 하고, 별도의 학습이 필요하지 않으면서 국부적 해와 같은 마이너스 효과에 빠지지 않기 때문에 수많은 센서를 사용해야 하는 생물학적 후각 모사 시스템에서는 가장 적절한 알고리즘일 수 있다.

## REFERENCES

- [1] Persaud K and Dodd G, "Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose", *Nature*, Vol. 299, p. 5881, 1982
- [2] Arthur L.P.S. Bailey, Anna Maria Pisanelli, and Krishna C. Persaud, "Development of conducting polymer sensor arrays for wound monitoring", *Sensor Actuat. B-Chem.*, Vol. 131, pp. 5-9, 2008.
- [3] R. E. Schaak, D. B. Dahlberg, and K. B. Miller, "The electronic nose: studies on the fundamental response and discriminative power of metal-oxide sensors, electronic noses and sensor array based systems design and application", *Proceedings of the 5th International Symposium on Olfaction and the Electronic Nose*, pp. 27-30, Baltimore, MD, USA, 1998.
- [4] T. Hamacher., J. Niess, P. S. Lammers, B. Diekmann, and P. Boeker, "Online measurement of odorous gases close to the odour threshold with a QMB sensor system with an integrated preconcentration unit", *Sensor Actuat. B-Chem.*, Vol. 95, pp. 39-45, 2003.
- [5] H. G. Byun, K. C. Persaud, S. M. Khaffaf, P. J. Hobbs, and T. H. Misselbrook, "Application of unsupervised clustering methods to the assessment of malodour in agriculture using an array of conducting polymer odour sensors", *Computers and Electronics in Agriculture*, Vol. 17, pp. 233-247, 1997.
- [6] J. W. Gardner, E. L. Hines and M. Wilkinson, "Application of artificial neural network to an electronic olfactory system", *Meas. Sci Technol.*, Vol. 1, pp. 446-451, 1990.
- [7] Beccherelli R, Zampetti E, Pantalei S, Bernabei M, and Persaud K. "Design of a very large chemical sensor system for mimicking biological olfaction", *Sensor Actuat. B-Chem.*, Vol. 146, No. 2, pp. 446-452, 2010.
- [8] M. Bernabei and K. C. Persaud. "Development and test of a large array of conductive polymers sensors resembling the redundancy of the biological olfactory system", *Chemical Senses*, 2011.
- [9] J. G. Hildebrand and G. M. Shepherd, "Molecular mechanisms of olfactory discrimination: converging evidence for common principles across phyla", *Annu. Rev. Neurosci.*, Vol. 20, pp. 595-631, 1997.
- [10] M. Ma and G.M. Shepherd, "Functional mosaic organization of mouse olfactory receptor neurons", *Proc. Natl. Acad. Sci.*, Vol. 97, No. 23, pp. 12869-12874, U.S.A., 2000.
- [11] Buck L and Axel R. "A novel multigene family may encode odorant receptors: a molecular basis for odor recognition", *Cell*, Vol. 65, pp.175-187, 1991.
- [12] Reed, R, R., "Signaling pathways in odorant detection", *Neuron*, Vol. 8, pp. 205-209, 1992.
- [13] C. Distanto, M. Leo, and K. C. Persaud, "Wavelet transform for electronic nose signal analysis", In: Hannu Olkkonen, ed. *Discrete Wavelet Transforms - Biomedical Applications*. Intechweb.org. 2011
- [14] K. Brudzewski, S. Osowski, K. Wolinska, and J. Ulaczyk, "Smell similarity on the basis of gas sensor array measurements", *Sensor Actuat. B-Chem.*, Vol. 129, pp. 643-651, 2008.

- [15] W. P. Jones and G. W. Furnas, "Pictures of relevance: a geometric analysis of similarity measures", *J. Am. Soc. Inform Sci.*, Vol. 36, pp. 420-442, 1987.
- [16] J. W. Gardner and P. N. Bartlett, "Electronic noses principles and applications", *Oxford Science Publications*, 1999.
- [17] J. D. Kim, H. G. Byun, D. J. Kim, Y. K. Ham, W. S. Jung, and C. O. Yoon, "A simple taste analyzing system design for visual and quantitative analysis of different tastes using multi-array chemical sensors and pattern recognition techniques", *Talanta*, Vol. 70, pp. 546-555, 2006.
- [18] J. D. Kim, J. J. Kim, S. D. Park, H. G. Byun, K. C. Persaud, and S. J. Lim, "The redundancy reduction using fuzzy C-means clustering and cosine similarity on a very large gas sensor array for mimicking biological olfaction", *J. Sensor Sci. & Tech.*, Vol. 21, No. 1, pp. 59-67, 2012.