

# 신호 패턴 분류를 위한 ADSTM 기법\*

김 아 람\*, 이 승 재\*\*, 김 창 화\*\*

\*강릉대학교 교육대학원 컴퓨터교육전공, \*\*강릉대학교컴퓨터공학과  
e-mail:ciela@cs.kangnung.ac.kr\*, {silveree, kch}@kangnung.ac.kr\*\*

## ADSTM Methodology for Signal Pattern Classification

Aram Kim\*, Seungjae Lee\*\*, Changhwa Kim\*\*

\*Graduate school of Computer Education

\*\*Department of Computer Science and Engineering,  
Kangnung National University

### 요 약

일반적으로 센서 어레이는 많은 채널의 센서를 가지고 있으므로 분석해야 할 데이터의 양이 많다. 따라서 다변량(多變量) 분석 방법을 이용하는데, 크게 통계적 방법과 신경망 방법을 분석하고자 하는 데이터의 특성이나 분석에 필요한 환경 조건에 맞는 분석 방법을 선택하여 이용한다. 센서 어레이의 신호 패턴을 분석하기 위해 본 연구에서는 상태 천이 모델을 이용하여 측정된 가스의 특성을 반영할 수 있는 통계적 방법에 대해 연구하였다. 센서 어레이 신호 데이터를 패턴 모양의 특성을 나타낼 수 있는 상태 천이 모델로 변환하여 가스 종류 식별이 보다 정확하게 이루어 질 수 있도록 모델을 설계하는데 중점을 두고, 모델링 요소인 '상태'는 일정한 시간 간격으로 샘플링 하였을 때의 신호값으로, '천이' 관계는 각 천이 벡터의 각으로 각각 정의하여 각도변이 기반 상태천이 모델링을 고안하였다.

### 1. 서론

센서는 일반적으로 “측정 대상물로부터 정보를 검지 또는 측정하여 그 측정량을 인식 가능한 유용한 신호로 변환하는 장치”, “외계의 정보를 감지하여 신호처리하기 쉬운 전기나 빛의 신호로 변환하는 기능을 지닌 장치” 등으로 정의된다[1]. 보통 그 자체가 독립적으로 기능을 수행하기보다는 다수의 신호처리장치나 작동장치의 일부로 포함되어 기능을 수행하는 경우가 대부분이다. 장치의 일부로써 장치가 목적으로 하는 정보를 얻어내기 위한 입력 장치로서의 역할을 하며, 센서에서 측정된 값은 장치의 분석 부분에서 유용한 정보로 변환되어 장치의 이용자에게 전달된다. 신호를 정보로 변환하는 과정은 단순한 데이터를 임의의 패턴(pattern)으로 변환하고 의미를 부여하는 과정이라고 할 수 있다.

본 연구에서는 가스 신호 패턴의 인식 방법으로 상태 천이 모델(State Transition Model)을 적용하고자 한다. 기존의 가스 신호 패턴 인식 방법들은 가스 신호 패턴의

특성을 잘 반영하지 못하였고, 일부 신경망 모델의 경우에 가스 모델에 대한 학습시간이 길어지는 경우가 많았다.

이에 가스 신호 패턴의 특성을 잘 반영할 수 있고, 학습뿐만 아니라 인식에서도 상대적으로 짧은 시간에 수행 가능한 모델로 상태 천이 모델을 가스 신호 패턴 인식에 적용해 보았다. 상태 천이 모델은 이미 다양한 분야에서 모델링 방법으로 사용되고 있다. 본 연구에서는 다 채널의 센서 어레이 신호들을 분석 대상으로 시간이 경과함에 따라 각 채널별로 독립적으로 변화되는 다 차원의 신호 값들의 특징을 반영할 수 있는 상태 천이 모델링 방법을 설계하고, 방법의 효율성과 특징에 대해 실험해 본다.

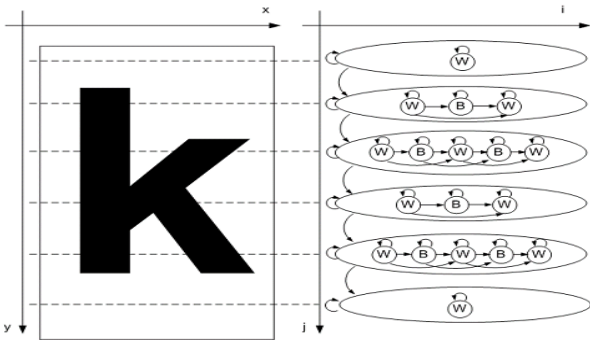
### 2. 관련연구

상태 천이 모델은 시스템의 프로세스나 어떠한 현상의 연역관계 등을 도형화하거나 시각화하는 모델링 도구로 주로 사용되었다. 이 모델을 특정 패턴을 분석하는데 적용하기 위해서 통계적 기법을 적용하여 사용한다. 본 논문에서는 센서 어레이에서 출력된 가스 신호값들을 각 채널별로 2차원 좌표 평면상에 나타내었을 경우, 평면상에 나타

\* 본 연구는 산업자원부 지방기술혁신사업(RTI05-01-02) 지원으로 수행되었음

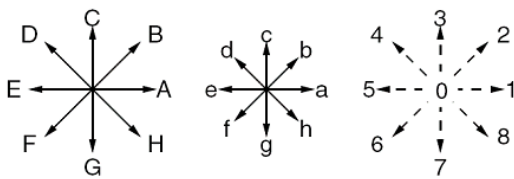
나는 시간에 따른 신호값의 변화 그래프를 분석 대상으로 하였다.

2차원 상에 표현된 데이터를 분석하는데 사용된 상태 천이 모델에는 Pseudo 2D 모델링, Substroke 모델링, Nebulous stroke 모델링과 같이 주로 문자 인식 모델링 방법이 있다.

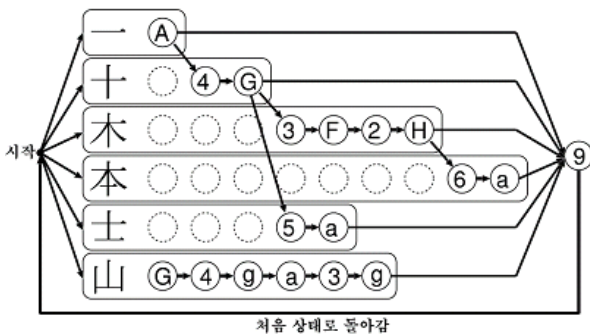


<그림 1> Pseudo 2D model

2차원 픽셀 맵(pixel map)에 표현된 문자를 흑(black)과 백(white)으로 구분하여 상태 천이 모델을 만든다. <그림 1>에서와 같이 y축 방향으로 진행하면서 x축 방향으로 상태가 바뀌는 각 경우를 모델링 한다[4].



<그림 2> substroke 획의 방향

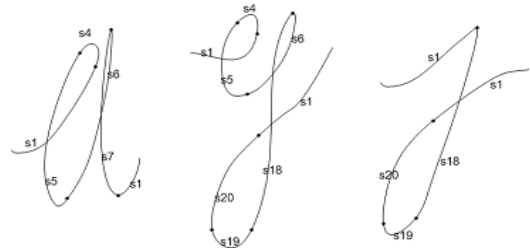


<그림 3> Substroke model

Substroke 모델은 일본어 간지를 모델링 할 때 사용한 방법으로 같은 획이 여러 문자에서 공동으로 사용되는 것을 이용하여 획의 방향과 길이, 펜을 들었을 때와 인식 판에 대었을 때의 펜의 이동 방향 등 <그림 2>에서와 같이 3요소로 모델링을 하였다. <그림 3>은 문자를 모델링 하였을 경우의 시퀀스 탐색 예시를 보여준다[2].

필기체 문자의 획을 여러 단위로 잘게 나누어, 각 부분 획을 기준으로 모델링 한 것이 <그림 4>의 Nebulous

stroke 모델이다[3].

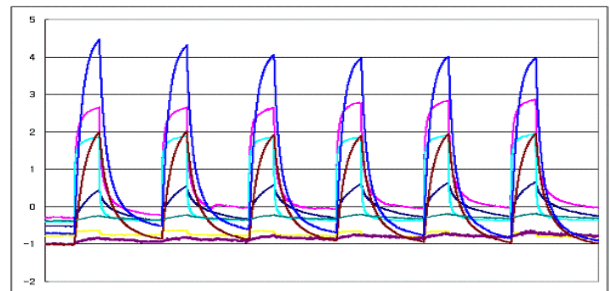


<그림 4> Nebulous stroke model

2차원으로 표현된 데이터를 상태 천이 모델로 표현하기 위해 이와 같은 상태 천이 모델링 방법들이 사용된다.

### 3. ADSTM 기법

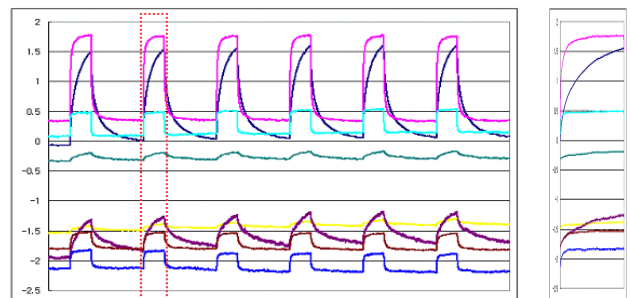
센서 어레이에서 출력된 신호 그래프는 <그림 5>와 같이 나타낼 수 있다. 그림은 8개의 채널(센서)로 구성된 센서 어레이로 대상 가스를 6회 반복 측정된 결과를 보여준다. 데이터 특성을 살펴보면, 측정된 경우마다 8개의 채널은 각기 일정한 패턴을 보이고 있다.



<그림 5> 센서 어레이에서 출력된 신호

각도 변이 기반 상태천이 모델링(ADSTM : Angle Difference based State Transition Model)은 <그림 5>의 패턴에서 각도 변이 관계를 상태 천이 관계로 변환하여 모델링 하는 방법으로 의미 구간 추출, 추출 구간의 선형 배치, 양자화, ADSTM 변환의 과정으로 모델링 방법을 고안하였다.

#### 3.1 의미 구간 추출



<그림 6> 의미 구간 추출

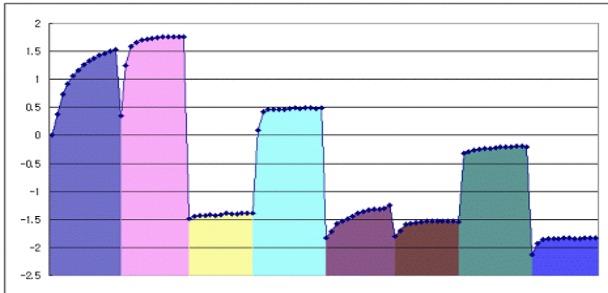
센서의 패턴에서 측정 결과의 의미를 지니는 부분으로 측정 시작시점부터 값이 최고점에 이르는 사이 구간이다. 이 구간은 센서가 가스에 반응하여 그 값의 변화를 보여주는 구간이다. 의미 구간 추출 과정은 <그림 6>에서와 같이 나타낼 수 있다.

### 3.2 추출 구간의 선형 배치

추출된 의미 구간을 일정한 채널 순서로 선형 배치한다. 선형으로 배치한 결과는 <그림7>과 같이 나타낼 수 있다.

### 3.3 양자화

양자화 과정으로 각 채널별로  $k$ 번 샘플링하여 값을 추출한다. <그림 7>과 같이 패턴의 양자화로 각 요소 사이의 프로세스 관계를 추출할 수 있다.



<그림 7> 추출 구간의 선형 배치 및 양자화

센서 어레이의 채널의 수를  $l$ 개라 할 경우, 총 양자화 요소의 개수  $m$ 은 식 (1)과 같이 나타낼 수 있다.

$$m = k \times l \quad (1)$$

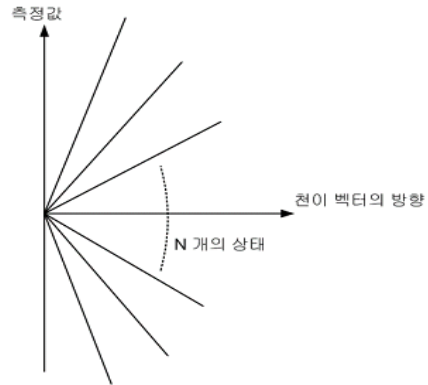
양자화 결과 모두  $m$ 개의 요소가 생성된다. 양자화 된 각 요소는 식 (2)와 같이 나타낼 수 있다.

$$q_1, q_2, q_3, \dots, q_m \quad (2)$$

### 3.4 ADSTM 변환

ADSTM을 구성하는 최소 요소는 상태(state)와 상태들 사이의 천이(transition)관계이다. 각 측정값 사이의 관계를 천이 요소로 하여 모델을 만든다.  $q_i$ 에서  $q_{i+1}$ 로 향하는 모든 벡터를 구하면  $\vec{q_1q_2}, \vec{q_2q_3}, \dots, \vec{q_{m-1}q_m}$ 와 같이  $(m-1)$ 개의 벡터를 얻을 수 있고, 이 벡터를 천이 벡터로 정의한다. 또한 본 논문에서 천이 벡터의 각은  $x$ 축과 천이 벡터 사이의 각으로 정의한다. 본 논문에서 센

서 어레이 신호를 모델링함에 있어 천이 벡터의 방향을 상태 천이 모델에서의 '상태'로 이용한다.



<그림 8> 상태(state)의 구분

벡터의 방향은 <그림 8>에서와 같이 좌표축의 원점을 기준으로 했을 때, 1, 4 사분면 방향으로 구간  $(-\frac{\pi}{2}, \frac{\pi}{2})$ 에 존재한다. 따라서 구간  $(-\frac{\pi}{2}, \frac{\pi}{2})$ 을  $n$ 개의 일정한 크기의 구간으로 나누어 각 구간을 하나의 상태로 정의한다. 상태의 각 범위를  $s_i$ 라 하면  $i$ 번째 상태  $s_i$ 는 식 (3)과 같다.

$$\begin{aligned} \frac{\pi}{n} \times (i-1) \leq s_i < \frac{\pi}{n} \times i & \quad (0 < i \leq \frac{n}{2}) \\ -\frac{\pi}{n} \times (i - \frac{n}{2}) \leq s_i < -\frac{\pi}{n} \times (i - \frac{n}{2} + 1) & \quad (\frac{n}{2} < i \leq n) \end{aligned} \quad (3)$$

$0 < i \leq \frac{n}{2}$ 의 범위는 천이 벡터의 기울기가 양수임을,

$\frac{n}{2} < i \leq n$ 의 범위는 천이 벡터의 기울기가 음수임을 나타낸다. <표 1>은 천이 벡터의 각을 상태 순서대로 표현한 것이다.

<표 1> 상태 기울기 범위

$i$	상태의 기울기 범위
1	$0 \leq S_1 < \frac{\pi}{n}$
2	$\frac{\pi}{n} \leq S_2 < \frac{2\pi}{n}$
$\vdots$	$\vdots$
$\frac{n}{2}$	$\frac{\pi}{2} - \frac{\pi}{n} \leq S_{\frac{n}{2}} < \frac{\pi}{2}$
$\frac{n}{2} + 1$	$-\frac{\pi}{n} \leq S_{\frac{n}{2}+1} < 0$
$\frac{n}{2} + 2$	$-\frac{2\pi}{n} \leq S_{\frac{n}{2}+2} < -\frac{\pi}{n}$
$\vdots$	$\vdots$
$n$	$-\frac{\pi}{2} \leq S_n < -\frac{\pi}{2} + \frac{\pi}{n}$

천이 벡터의 각을 천이 관계로 나타냈을 때 이 각의 크기는 센서의 측정값인 좌표 평면 내에서 천이 벡터의  $x$  값을 1로 정규화 하면  $y$ 값만으로 아크탄젠트 함수를 이용하여 구할 수 있다. 기울기 값을 차례로 나열했을 때, 각의 천이 벡터 각  $d_i$ 는 식 (4)와 같이 구할 수 있다.

$$d_i = \arctan(q_i - q_{i-1}) \quad (1 \leq i \leq m-1) \quad (4)$$

이렇게 계산된 천이 벡터의 각의 순차열을 이용하여 가스 신호 패턴 인식을 위한 상태 천이 모델이 만들어 진다. 구간  $(-\frac{\pi}{2}, \frac{\pi}{2})$ 를  $n$ 개의 상태(state)로 나누었을 때, 각 상태마다 각의 범위는 <표 1>과 같이 나누어지고 천이 벡터의 각이 속하는 범위가 순차열에 표시된다. 생성된 순차열은  $n \times n$ 의 상태 천이 행렬을 만들어 상태로 변환될 수 있다. 상태  $i, j$ 가  $1 \leq i \leq n, 1 \leq j \leq n$  일 경우, 상태  $i$ 에서 상태  $j$ 로 천이가 발생하여  $a_{ij}$  상태 천이 행렬이 생성되며 <표 2>와 같이 나타낼 수 있다. 각 상태 간에 천이 관계가 존재할 경우, 관계가 있다는 표시로 '1'의 값을 삽입하여 최종 ADSTM 모델을 완성한다.

<표 2> 상태천이 행렬

$j \backslash i$	$s_1$	$s_2$	$s_3$	...	$s_n$
$s_1$	$a_{11}$	$a_{12}$	$a_{13}$	...	$a_{1n}$
$s_2$	$a_{21}$	$a_{22}$	$a_{23}$	...	$a_{2n}$
$s_3$	$a_{31}$	$a_{32}$	$a_{33}$	...	$a_{3n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$s_n$	$a_{n1}$	$a_{n2}$	$a_{n3}$	...	$a_{nm}$

센서 어레이에서 새로운 가스가 측정되어 생성된 상태 천이 행렬을  $O$ , 모델 저장소에 보관된 비교 대상 모델의 상태 천이 행렬을  $M$ 이라 하자. 각각 1의 값을 가지는 요소의 개수를  $n(O)$ ,  $n(M)$  이라 할 때, 두 행렬의 유사도(類似度)인 **similarity**는 식(5)와 같은 방법으로 계산할 수 있다.

$$similarity = \frac{n(O \cap M)}{n(O \cup M)} \quad (5)$$

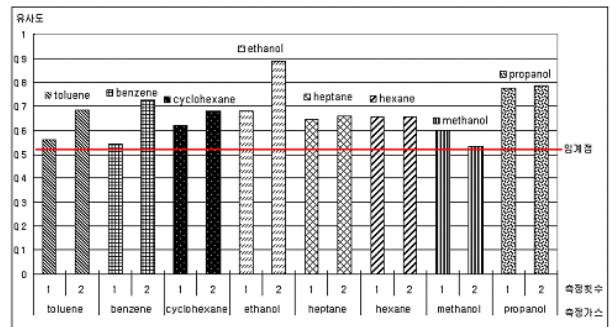
여기에서 교집합( $\cap$ )은 두 상태 천이 행렬에서 같은 위치의 행과 열에 동일하게 1의 값을 갖는 요소의 개수를 의미한다. 합집합( $\cup$ )은 두 행렬에서 같은 위치에 어느 한 쪽이라도 1의 값을 가지고 있는 것이 개수를 의미한다. 이 유사도는 상대적 일치도 측정에 매우 유용한 방법이다.

4. 결론

본 논문에서 사용한 ADSTM 기법에 대해 여덟 가지 가스를 6회 반복 측정된 결과를 데이터를 검증에 사용하였다. 6회의 측정 결과 중에서 무작위로 4회의 측정 결과를 선택하여 ADSTM 모델링 한 후에 임계값을 결정하고, 남은 2차례 측정 결과로 검증하는 과정을 무작위 반복하였다.

모델링 변인 중에서 인위적으로 조절 가능한 상태의 개수( $n$ )과 샘플링 횟수( $k$ )의 값을 조절하여 실험한 결과  $n$ 은 20, 60, 100의 변인 중 20의 경우에,  $k$ 는 7, 13, 25의 변인 중에서 7의 경우에 최적화된 결과를 나타내었다. <그림 9>는  $n$ 과  $k$ 의 값을 결정한 후, 임계값을 이용한 가스 패턴 인식 실험 결과이다.

실험을 통해 설계한 모델링 방법이 가스 패턴을 구분하는데 있어 적용 가능한 방법이라는 것을 알 수 있었고, 인식 결과 가스 구분에 효과적으로 작동하였다. 가스 센서 인식 분야에서 기존에 거의 시도되지 않았던 상태 천이 모델을 적용하여 인식 가능한 모델을 보일 수 있었던 것이 하나의 성과라 할 수 있다.



<그림 9> 임계값을 이용한 실험 결과

참고문헌

- [1] 김상진, “자동화를 위한 센서”, 연학사, pp. 15-20, 1998.
- [2] Hiroshi SHIMODAIRA, Takashi SUDO, Mitsuru NAKAI and Shigeki SAGAYAMA, On-Line Overlaid-Handwriting Recognition Based on Substroke HMMs, Proceeding of the Seventh International Conference on Document Analysis and Recognition(ICDAR 2003), pp. 1-5, 2003.
- [3] Jianying Hu and Michael K. Brown and William Turin, HMM Based On-Line Handwriting Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol 18 No 10, pp. 1039-1041, 1996.
- [4] Oscar E. Agazzi and Shyh-Shiaw Kuo, Hidden Markov Model based Optical Character Recognition in the Presence of Deterministic Transformations, Pattern Recognition Vol 26 No 12 Pattern Recognition Society, pp. 1814-1815, 1993.