

PCA와 결합된 Fuzzy C-Means 알고리즘을 이용한 전자 혀 시스템 개발

Development of Electronic Tongue System Using Fuzzy C-Means Algorithm Combined to PCA Method

김 정 도*, 정 우 석, 홍 철 호
(Jeong Do Kim, Woo Suk Jung, and Chul Ho Hong)

Abstract : In this paper, we investigate the visual and quantitative analysis at the same time with an electronic tongue(e-tongue) system using an array of ISE(ion-selective electrode). We apply the FCM(fuzzy c-means) algorithm combined with PCA(principal component analysis), which can be reduced multi-dimensional data to third-dimensional data, to classify data patterns detected by E-Tongue system. The proposed technique can be designed to solve the cluster centers and membership grade of patterns combined with the output results obtained by PCA method. According to the proposed technique, the membership grade of unknown pattern, which does not shown previously can be determined and analyzed visually. Conclusionally, the relationship between the standard patterns and unknown pattern can be easily analyzed. Throughout the experimental trials, the proposed technique has been confirmed using developed E-Tongue system.

Keywords : electronic tongue, ISE, FCM algorithm, PCA

I. 서론

인간이 느끼는 맛이란, 감지대상 물질의 화학적 성분들의 양이 얼마나 되는가 하는 정량적 데이터가 아니라 여러 물질이 혀의 미뢰에 어느 정도 영향을 미치느냐에 따라 인간의 뇌에서 종합적으로 분석한 정성적 데이터라 할 수 있다. 물론 맛의 강도에 따라 인간의 뇌는 정량적 분석도 동시에 이루어지기는 하지만 맛 그 자체는 여러 성분의 정성적 분석에 의한 결과이다. 인간의 미각을 모방하기 위하여 전자 혀 (electronic tongue) 시스템의 개발이 활발하게 이루어지고 있는데, 이러한 전자 혀 시스템 시스템의 기본은 서로 다른 이온 선택성을 가진 센서들을 어레이(array)화 하여 사용하는 것이다. 이 센서 어레이에 의해 측정된 데이터는 맛에 따라 구분이 가능한 특징을 보이게 되며, 이는 각각의 센서가 맛의 화학적인 성분에 따라 다른 민감도를 가지는 물질에 의해 제조되기 때문이다. 센서들 사이의 상대응답에 의해 만들어진 패턴은 신호처리 기법에 의해 각 맛을 묘사할 수 있는 일종의 독특한 지문이 될 수 있다[1]. 이는 인간의 미각과 후각 기관인 혀와 코를 대신하여 음식의 판별, 음식의 변질 그리고 음식의 유해성 유무 등을 인간이 직접 맛을 보거나 향을 맡지 않아도 분석이 가능하도록 하여 식품, 음료 생산 공정 및 식음료 신개발 연구용뿐만 아니라 질병의 개인적 진단 등 우리 실생활에서 유용하게 활용할 수 있다. 일반적으로 식음료의 맛이나 향은 관능검사를 통해 비교, 판별되지만 그 정확성은 수치화되지 않는다. 이에 각국에서는 그 맛과 향을 정량화하기 위해 많은 노력을 기울여 왔

다. 종래의 전기적 맛 측정 방법은 이른바 다수의 이온선택 전극을 이용하여 많은 측정치의 재해석을 실행하거나 실리 콘 칩을 이용, 색의 변화로 구분하는 등의 기술을 이용해 왔다. 전자 혀 시스템에서 다중 선택성 화학센서 어레이로 사용되어지는 이온선택성 전극 (ion-selective electrode; ISE)은 1960년대 중반 처음 개발되기 시작한 이래로 최근까지 20년간 그 개발과 응용분야가 급속히 발전하여왔다. 이러한 이온선택성 막전극은 주어진 이온이나 이온들에 선택적인 막(membrane)전위를 가지는 막전극으로 단지 한 종류의 이온에만 선택적으로 감응하지는 않지만 주어진 이온이나 다른 이온들에 대해 어느 정도의 명백한 선택성의 정도는 나타난다. 그러나 이온 선택성 막전극들이 우수한 선택성을 가진다 하더라도 다른 이온에 대해서도 반응을 하게 된다. 따라서 미세한 맛의 변화를 파악하기 위해서는 어레이화된 막전극들을 사용하여야 한다. 어레이화된 여러 개의 막전극의 사용으로 각각의 막전극에 반응하는 패턴을 인식하여 막전극 개수보다 훨씬 다양한 종류의 맛을 분별할 수 있다. 따라서 많은 종류의 패턴들을 적절한 신호처리 알고리즘에 의해 분석할 필요가 있다. 보통 이 신호를 분석하기 위한 알고리즘으로 PCA(principal component analysis) 방법이 사용되고 있다[2]. 다중 센서가 사용되어진 데이터는 다차원의 데이터로 인간의 시각에서는 분석이 이루어 질 수 없다. PCA방법은 다 차원의 데이터를 인간이 시각적으로 분석 가능한 2차원 또는 3차원으로 축소 매핑하는 방법으로 후각과 미각 표현에서 특징을 추출할 수 있는 표준 방법으로 사용되고 있다. 하지만 PCA방법은 시각적 분석 방법이기 때문에 실제 데이터를 정량적으로 표현하거나 성질을 표현하기 위해서는 별도의 방법이 사용되어야 한다.

정량적 표현이 가능하면서 클러스터링이 가능한 대표적인 방법으로는 K-means 알고리즘이 있다. 또한 K-means 알

* 책임저자(Corresponding Author)

논문접수 : 2004. 8. 17., 채택확정 : 2005. 1. 6.

김정도, 정우석, 홍철호 : 호서대학교 전기정보통신공학부

(jdkim@office.hoseo.ac.kr/jws7216@hanmail.net/chong@office.hoseo.ac.kr)

고리즘의 단점을 보완한 지능적 방법인 FCM(fuzzy c-means) 알고리즘[3]이 개발되어 더 우수한 클러스터링 능력을 보여 주고 있다. K-means 알고리즘이나 FCM 알고리즘의 경우, unsupervised learning 방법에 의해 클러스터링 작업이 이루어 지기는 하지만 PCA와는 달리 다 차원의 데이터를 그대로 클러스터링하여 클러스터의 센터를 구하는 방법이다. 보통 인간의 경우 사물을 판단함에 있어 시각적 방법에 의존하려는 특성을 가지고 있기 때문에 실제 산업 현장에서는 PCA 방법이나 Spider Map 표현방법을 가장 선호하고 있다. 이러한 이유 때문에 PCA 방법은 반드시 사용되어야 할 필요가 있으나 클러스터링 작업이나 분석작업의 자동화를 위해서는 PCA 방법 이외에 별도의 K-means 알고리즘이나 FCM 알고리즘을 이용하여 정량화된 표현을 하는 것이 필요하다. 만약 PCA 방법과 FCM 방법을 별도로 이용하게 되면 PCA 방법에서의 차원과 FCM 알고리즘의 차원이 다르기 때문에 시각적으로 분석된 데이터와 FCM 알고리즘에 의해 구해진 데이터와는 별도의 상관관계를 가지게 된다.

이러한 문제를 해결하기 위하여 본 논문에서는 PCA방법과 FCM 알고리즘이 상관관계를 가지고 시각적 분석과 정량적 분석과의 일치화를 이루게 하는 것을 목적으로 한다. 이를 위해 먼저 PCA 방법을 통해 3차원 고유벡터를 구한 후, 이 고유벡터에 의해 차원이 축소된 데이터를 이용해 3차원으로 FCM 알고리즘을 구현하는 방법을 이용한다. 이를 통해 새로운 데이터가 들어온다 하더라도 시각적으로 표현 가능하며 이 시각적 데이터와 상관관계를 가지는 클러스터의 센터를 통해 새로 입력된 데이터의 각 클러스터 사이의 소속도를 구할 수 있게 된다.

본 논문에서는 전자 혀 시스템을 개발하고, 후각과 미각의 분류 방법으로 PCA 방법과 FCM 알고리즘을 결합한 알고리즘을 제안한다. 제안된 알고리즘을 이온 선택성 전극 센서어레이에 의해 구현된 전자 혀 시스템에 도입하여 실험하였으며 실험을 통해 그 우수성을 입증하였다.

II. 이온 선택성 전극의 원리에 의한 E-tongue 설계

1. 이온 선택성 전극

전자 혀 시스템에서 다중 선택성 화학센서 어레이로 사용되어지는 이온선택성 전극(ion-selective electrode)은 주어진 이온들에 선택적인 막(membrane)전위를 가지는 막전극으로 단지 한 종류의 이온에만 선택적으로 감응하지는 않지만 주어진 이온이나 다른 이온들에 대해 명백한 선택성의 정도로 나타난다. 이들은 고분자 물질을 전극 막의 지지체(matrix)로 사용하며, 특정이온에 선택성을 주는 전극물질(ionophore)과 함께 비 휘발성 유기용매인 가소제(plasticizer), 경우에 따라 친유기성 첨가제(lipophilic additive)를 포함한다. 고분자막형 이온선택성 전극이 감응하는 원리는 보통 상경 전위차 이론으로 설명한다. 그림 1은 이온선택성 전극을 나타낸 모형도이다. 여기서 이온선택성 전극의 신호는 정상적으로는 시료 경계면에 형성되는 이중막(double layer)의 커패시터의 변화에 의하여 나타난다고 설명할 수 있고 정량적으로는 시료의 화학포텐셜 μ_{sam} 과 유기상(막)의 화학포텐셜 μ_{mem} 이 평형을 이루기 위하여 형성된다고 설명할

수 있다. 상경계면에 이중층의 형성에 의한 커패시터 전위차는 유기상의 막과 시료용액의 화학포텐셜이 평형을 이루는 조건에 의하여 계산할 수 있다.

유기상 이온 선택성 막의 화학 포텐셜에 대한 식은 (1)과 같고

$$\mu_{mem} = \mu_{mem}^{\circ} + RT \ln a_{I,mem} + zF\phi_{mem} \tag{1}$$

시료용액의 화학 포텐셜에 대한 식은 (2)와 같다.

$$\mu_{sam} = \mu_{sam}^{\circ} + RT \ln a_{I,sam} + zF\phi_{sam} \tag{2}$$

where : Z (이온가), F (패러데이 상수),
 R (기체상수), T (절대온도),
 $a_{I,mem}$ (유기상막의 이온 I^{+} 의 활동도),
 $a_{I,mem}$ (시료의 이온 I^{+} 의 활동도)

이 식을 이용하여 평형상태($\mu_{sam} = \mu_{mem}$)에서의 상경계 전위차 $\Delta\phi$ 를 구하면 네른스트 방정식형태의 (3)으로 나타난다.

$$\Delta\phi = -\frac{\mu_{mem}^{\circ} - \mu_{sam}^{\circ}}{zF} + \frac{RT}{zF} \ln \frac{a_{I,sam}}{a_{I,mem}} \tag{3}$$

여기서

$$\Delta\phi = \phi_{mem} - \phi_{sam} \tag{4}$$

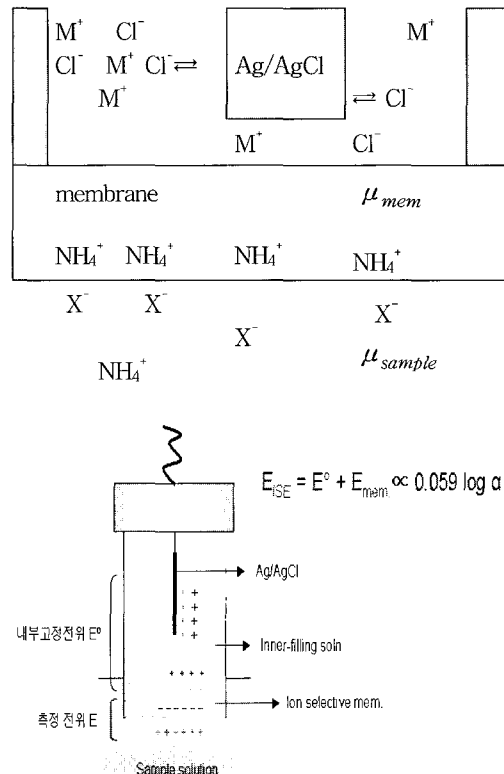


그림 1. 이온 선택성 전극.
 Fig. 1. Ion-selective electrode(ISE).

이때 $a_{I, mem}$ 는 거의 상수이므로 상정계전위차는 시료 중 이온농도의 로그에 비례하는 값으로 나타난다. 전자 혀 시스템은 여러 종류의 이온선택성 전극의 이온 농도에 비례하는 전위차를 측정하여 이들의 감응 패턴으로 시료의 맛을 구분하게 된다. 이와 같이 유기용매/고분자 유형 이온 선택성 전극막(solvent/polymeric ion-selective membrane)은 전극에 장착된 이온선택성 막이 특정 이온이나 분자와 결합했을 때 막 표면에 형성되는 전하 분리(charge separation)층에 의하여 발생하는 전위차를 측정하여 시료내 함유된 이온을 정량할 수 있다. (그림 1) 특히 이온선택성 막전극은 우수한 이온선택성(selectivity)을 나타내고, 감응시간이 수 초 이내로 매우 짧으므로 대량분석이 용이하며, 분광분석과는 달리 분석방법이 시료의 색깔이나 탁도 등에 영향을 받지 않으므로 시료의 전처리 단계가 필요 없고, 전극막의 제조가 용이하며 다른 분석방법에 비해 가격이 저렴하다는 장점을 가지므로 다양한 응용분야에서 여러 이온 종류에 대해 연구가 활발히 진행되어 왔다.

2. 이온 선택성 전극을 이용한 다중 배열 센서구조

이온선택성 전극막을 이용한 전극은 크게 두 가지 형태로 나눌 수 있다.

이온선택성 전극막과 내부금속전극(internal reference metal electrode) 사이에 내부기준용액(internal reference filling solution)을 필요로 하는 재래식 이온선택성 전극과 이를 필요로 하지 않는 고체상(solid-state)전극이다. (Fig. 2) 이들 두 전극형태의 차이점은 재래식 이온선택성 막전극은 전극체에서 막을 고정시켜 전극막의 이탈을 막아 줄 수 있으나, 고체상(solid-state)전극의 경우 이온선택성 전극막이 전극표면에 아무 고정체도 없이 노출되어 있으므로 전극막의 전극표면에 대한 접착력이 전극의 전기화학적 특성을 결정짓는 중요한 요인으로 작용한다는 것이다. 그러나 고체상(solid-state)전극은 다음과 같은 장점을 가지므로 앞으로 자동화 분석 장비의 개발과 연계되어 상업적인 센서 개발의 주요 전극형태가 될 것으로 판단된다. 고체상(solid-state)전극이 갖는 장점은 (1) 내부기준용액을 필요로 하지 않으므로 소형화에 유리하며, (2) 한 개의 소자(chip)에 여러 이온을 동시에 검출할 수 있는 다중센서의 제작이 가능하고, (3) 대량생산이 용이하므로 가격의 저렴함까지 가져 올 수 있다.

전자 혀 시스템을 개발하기 위해 다양한 종류의 이온선택성 전극 물질들을 비교 조사하였고, 특히 주요 전해질 이온(K⁺, NH₄⁺, Na⁺, Cl⁻, NO₃⁻, NO₂⁻, CO₃²⁻, Li⁺, H⁺, Ca²⁺, Mg²⁺ 등)에 대해서는 특성 분석을 하였다. 각각의 이온선택성 전극은 그 선택성을 중요시 하여 낮은 선택성을 갖는 전극은 관심을 끌지 못한다. 이해 반해 전자 혀 시스템은 각각의 이온선택성 전극이 사용되기는 하나 그 전극의 선택성보다는 여러 시료들을 측정했을 때의 각 센서들 상호간의 영향과 이온 또는 분자들 간의 영향, 이온 또는 분자들과 센서간의 반응에 의한 일차적인 데이터를 처리하여 정확한 맛에 대한 정보를 얻기 위해 주성분 분석, 다 변량 회귀분석, 퍼지 논리분석, 인공 신경망분석 등 통계적인 데이터 분석 알고리즘을 통해 미지시료의 맛을 판별하기위한 기준을 정하게 된다.

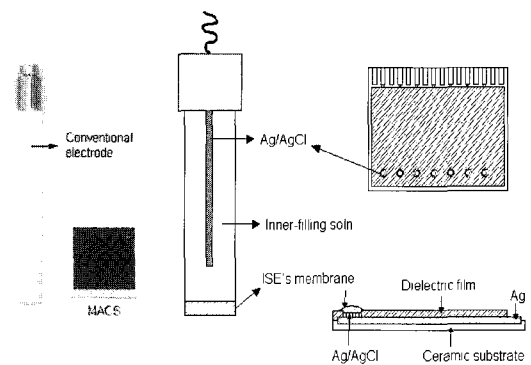


그림 2. 기존의 전극과 다중배열 화학 센서 어레이.

Fig. 2. Conventional electrode and MACS.

따라서 특정 이온선택성 물질이 첨가되지 않은 blank막도 전극의 구성요소로 이용되어진다.

이온선택성 전극막의 제작을 위해 막구성 요소-지지체(matrix) 33wt%, 이온선택성 물질(ionophore) 1wt%, 가소제(plasticizer) 66wt% 및 친유성 첨가제(lipophilic additive)-의 종류 및 조성변화를 통해 다양한 이온선택성 막을 제작하였다. 대부분의 막 조성은 기본 조성비와 크게 다르지 않다. 제작 방법은 지지체, 가소제, 친유성 첨가제 및 이온선택성 물질을 유기용매에 녹여 유리판 위에 고정시킨 직경 22mm의 유리관에 부어 성형시키고, 용매는 실온에서 대기 중에 하루 이상 방치하여 증발시켰다.

위와 같은 방법으로 제작된 각종 이온선택성 전극막은 직경 5.5mm의 원형으로 잘라 Philips 전극체에 장착하여 작동 전극으로 사용하였으며 내부기준용액으로는 적정농도의 타겟(target)이온을 포함하는 용액을 사용하였다. 고체상 염화 이온 선택성 막에 구비되는 이온선택성 막을 형성하기 위해서는 막 조성물을 용매에 희석시킨 후 상기 용액을 마이크로실린지(microsyringe)를 이용하여 고체상 전극 표면에 air fluid dispensor(1000XL: EFD Inc., East Providence, RI, USA)를 사용하여 막 조성물을 얇게 입혀 실온에서 하루 이상 건조시켰다. 이렇게 제조된 전극은 적절한 원충용액에서 기초 성능테스트를 거친 뒤 여러 가지 시료 분석에 사용되어진다.

III. E-Tongue 하드웨어 설계

전자 혀 시스템은 액체 시료를 이온 선택성 전극으로 구성된 센서 어레이가 가지는 특정한 전기적 신호값으로 측정된다. 다시 말하면 기준전극과 측정전극사이의 전위차를 측정하여 이를 분석하게 된다. 그림 3은 전자 혀 시스템 데이터 취득회로의 구성을 보여주고 있다. 센서로부터 출력되는 신호는 수mV~수배mV로 낮기 때문에 센서 입력단에 버퍼를 구성하여 측정회로가 신호원에 영향을 주지 않도록 하였다. 버퍼를 거친 신호는 저역통과필터를 통과하여 신호원에 포함된 고주파잡음을 제거하게 된다. 멀티플렉서는 센서 각각의 필터를 거친 신호를 선택적으로 입력받아 출력하게 되며, 다시 버퍼를 거쳐 A/D변환기로 입력된다. A/D변환기는 측정 데이터 값을 정밀하게 디지털 값으로 변환시키기 위해 16bit의 분해능을 가지도록 설계하였다.

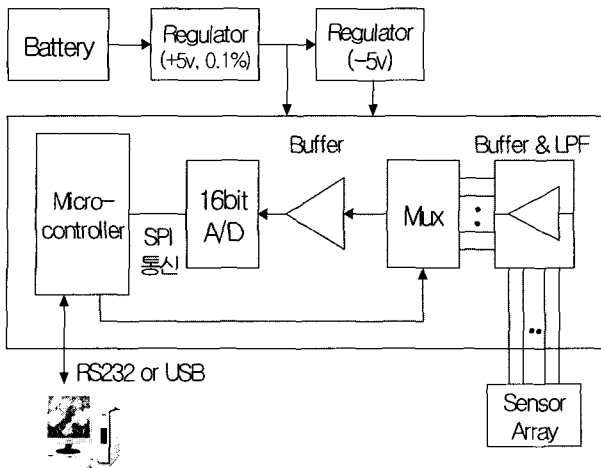


그림 3. 데이터 취득 회로의 구성.
Fig. 3. Block diagram of data acquisition circuit.

A/D변환기가 16비트의 분해능을 가진다 하더라도 기준 전압이 변동한다면 측정 정밀도는 떨어질 수밖에 없다. 그래서 A/D 변환기에 인가되는 기준전압을 만들어 주는 전압 레귤레이터를 0.1%의 변동률을 가지는 것을 사용하였다. 하지만 0.1%의 변동률로는 16bit의 분해능에 해당하는 정밀도를 이룰 수가 없기 때문에 매 1초당 200번 이상의 데이터를 받아 평균을 내는 방법을 사용하였으며, 이에 따라 측정된 전위값의 오차범위는 0.15mV 이하가 된다. A/D 변환에 의해 측정된 데이터들은 SPI통신을 통해 원칩 마이크로 컨트롤러로 전달되며, RS232 혹은 USB와의 연결을 통해 PC로 전달되어 PC에서 데이터를 알고리즘을 통해 분석하게 된다.

IV. PCA-FCM 알고리즘

PCA 방법은 다차원 데이터를 인간이 가시적으로 확인이 가능한 2차원이나 3차원의 데이터로 패턴을 매핑 하는 알고리즘이다.

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1L} \\ X_{21} & X_{22} & \dots & X_{2L} \\ \dots & \dots & \dots & \dots \\ X_{N1} & X_{N2} & \dots & X_{NL} \end{bmatrix} \quad (5)$$

여기서 X는 센서가 받아들인 입력 데이터이다. 이때 N은 패턴의 수이고 L은 센서의 수이다. 이러한 N×L의 데이터를 이제 (6)와 같은 N×D의 데이터 형태로 바꾸게 된다.

$$Y = \begin{bmatrix} Y_{11} & Y_{12} & \dots & Y_{1D} \\ Y_{21} & Y_{22} & \dots & Y_{2D} \\ \dots & \dots & \dots & \dots \\ Y_{N1} & Y_{N2} & \dots & Y_{ND} \end{bmatrix} \quad (6)$$

여기서 Y는 매핑 데이터이고 D는 원하는 차원의 수이다. 만약 k = 1, 2, ..., n 이면 (7)을 이용해서 각각의 센서에 대한 평균을 구하게 된다.

$$\bar{X}_k = \frac{1}{m} \sum_{i=1}^m X_{ik} \quad (7)$$

여기서 n은 센서의 수이고 m은 패턴의 수이다. 그 평균 데이터를 가지고 (8)을 이용하여 입력에 대한 공분산 행렬 (covariance matrix)를 구하게 된다. 공분산 행렬 C는 각각의 원소인 C_{ij}가 모든 classes 위에 중심점이 되어가도록 하기 위해서 입력 패턴 X에서 평균 벡터 \bar{X}_k 를 빼어감에 따라 계산되어 진다.

$$C_{ij} = \frac{\sum_{k=1}^m (X_{i1} - \bar{X}_i)(X_{j1} - \bar{X}_j)}{m} \quad (8)$$

다음에는 k = 1, 2, ..., n 에 대한 고유치 λ_k와 고유 벡터 u_k를 구한 다음 (9)와 같이 정의한다.

$$Cu_k = \lambda_k u_k \quad (9)$$

계산된 λ_k 중에서 (만약 3차원 매핑의 경우) 가장 큰 고유치 λ₁, λ₂와 λ₃를 선택하고 이와 연관되어 있는 고유 벡터 u₁, u₂와 u₃를 주성분으로 해서 다차원 데이터를 원하는 3차원으로 매핑하게 되는 것이다.

보통 3차원으로 매핑하게 되면 어느 정도의 특징을 가지고 있는 경우, 대부분 데이터의 큰 손실 없이 매핑이 가능하다. 이 PCA 방법을 이용하면 시각적으로 클러스터링 분석이 가능하다. 하지만 새로운 데이터가 취득되었을 경우, 이 새로운 데이터가 어느 클러스터에 가까운가 혹은 어느 클러스터에 소속되어 있는 가를 정량적으로 판단하기 위해서는 별도의 알고리즘이 필요하다.

이를 위해 본 연구에서는 PCA 알고리즘을 이용해 구해진 3차원 고유 벡터를 기준으로 하여 구해진 3차원 데이터들을 FCM(fuzzy c-means) 알고리즘을 이용하여 클러스터의 센터를 구하고, 구해진 클러스터의 센터를 이용하여 입력된 데이터의 소속도를 구하는 방식을 제안한다. FCM 알고리즘은 비지도 학습(unsupervised learning) 이기는 하지만, 맛에 대한 표준 데이터를 미리 정의하여 놓았다면 같은 데이터에 대해서는 같은 클러스터로 결정될 수 있기 때문에 지도학습(supervised learning)의 형태로 사용하는 것이 가능하다.

별도의 k-means 알고리즘이나 FCM 알고리즘을 이용하여 정량화된 표현을 하는 것은 PCA 방법에서의 차원과 FCM 알고리즘의 차원이 다르기 때문에 시각적으로 분석된 데이터와 FCM 알고리즘에 의해 구해진 데이터와는 별도의 상관관계를 가지게 된다. 이렇게 PCA의 결과를 이용한 FCM의 사용은 PCA에 의해 표현가능한 시각적 데이터와 상관관계를 가지는 클러스터의 센터를 통해 새로 입력된 데이터의 각 클러스터 사이의 소속도를 구할 수 있게 된다.

정성적 분석을 하기위한 FCM 알고리즘은 PCA 과정으로 매핑된 데이터를 입력벡터로 하여 Y₁의 요소인 y를 벡터로 표시하면

$$Y_{n3} = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ \vdots & \vdots & \vdots \\ y_{n1} & y_{n2} & y_{n3} \end{bmatrix} \quad (10)$$

여기서 n은 입력패턴수이고 PCA를 통해 3차원으로 줄인

행렬을 나타내고 있다.

다시 입력패턴으로만 표시하면 다음과 같고

$$Y_i = \{y_{1i}, y_{2i}, \dots, y_{ni}\} \quad (11)$$

입력패턴 집합에 대한 클러스터의 중심은 다음과 같이 나타낸다.

$$Z_j = \{z_{1j}, z_{2j}, \dots, z_{cj}\}, \quad 1 \leq j \leq c \quad (12)$$

여기서 c는 클러스터 중심의 갯수이고 Y_i의 i는 입력 패턴수 그리고 Z_j의 j는 클러스터의 수이다. 따라서 i번째 패턴과 j번째 클러스터의 중심의 소속도 W_{ij}는 다음과 같이 나타낸다.

$$W_{ij} \in [0, 1]; \quad 1 \leq i \leq n, \quad 1 \leq j \leq c \quad (13)$$

$$\sum_{j=1}^c W_{ij} = 1; \quad 1 \leq i \leq n \quad (14)$$

$$0 < \sum_{i=1}^n W_{ij} < n; \quad 1 \leq j \leq c \quad (15)$$

위에서 패턴들은 서로 다른 소속도와 클러스터를 가질 수 있다. 그리고 (14)와 (15)에서와 같이 하나의 패턴에 대한 소속도의 총합은 1로 정규화 되고 각각의 패턴 소속도는 목적함수 J_m이 최소가 되는 클러스터의 패턴을 찾는다.

$$d_{ij} = \| Y_i - Z_j \| = \left[\sum_{k=1}^c (Y_{ki} - Z_{kj})^2 \right]^{1/2}$$

$$J_m(W, Z) = \sum_{i=1}^n \sum_{j=1}^c W_{ij}^m d_{ij}^2 \quad (16)$$

목적함수 J_m에서 d_{ij}는 입력패턴과 클러스터 중심 사이의 유클리드 거리(euclidean distance)이다. 이것은 각각의 패턴에 대한 유사성을 측정한다. 여기서 m은 소속 가중치 지수(membership weighting exponent)며 이 값에 대한 이론적이고 정당화된 법칙이 존재하지 않기 때문에 일반적으로 m=2로 선택한다. Bezdek가 제안한 FCM 알고리즘의 목적함수의 해를 구하는 방법은 (17), (18)과 같이 소속도와 클러스터의 중심을 구한다.

$$W_{ij} = 1 / \sum_{k=1}^c \left[\left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}} \right] \quad (17)$$

d_{ij} > 0, ∀ i, j이고, 만약 d_{ij} = 0 이면 W_{ij} = 1 이고 j ≠ k 일 때 W_{ij} = 0 이 된다.

$$Z_j = \sum_{k=1}^n W_{ik}^m Y_{kj} / \sum_{k=1}^n W_{ik}^m \quad (18)$$

FCM 알고리즘은 (17)와 (18)에 의해 분석적으로 해결되지 않고 (16)의 목적함수의 최소값을 근사화시키는 반복적인 알고리즘이다. 따라서 (17)와 (18)의 학습으로 최적의 값으로 수렴된다.

이 알고리즘은 비지도 패턴인식 기술(unsupervised pattern

recognition technique)이기 때문에 그에 맞는 학습 절차가 필요하다. 이 학습절차는 입력패턴의 클러스터 센터에 대한 소속도를 가져오고 클러스터 센터에 대한 각각의 입력패턴의 거리를 소속도로 나타낸다.

학습을 위한 FCM 알고리즘 절차를 단계별로 정리하면 다음과 같다.

1) 초기 소속함수 Z(0) 결정 및 클러스터 j와 패턴 i의 소속도 W_{ij}의 계산 (17)

2) 클러스터 중심 Z(k+1)의 계산 (18)

3) || Z(k) - Z(k+1) || 의 계산

만약 || Z(k) - Z(k+1) || > ε면 k = k+1로 하고 처음으로 피드백하고 그렇지 않으면 알고리즘을 종료한다.

이렇게 PCA과정을 통해 차원을 줄인 결과를 다시 FCM 알고리즘을 통한 클러스터링을 함으로써 기존 데이터에 대한 새로운 데이터의 소속감, 시각적인 효과 그리고 정량적인 결과값을 동시에 얻을 수 있었다.

이러한 과정을 통해 미리 얻어진 표준 패턴들에 대한 클러스터의 중심과 각 표준 패턴들의 클러스터 소속도를 구할 수 있다. 여기서 구해진 클러스터 중심들은 미리 컴퓨터에 저장되어 미지의 샘플이 입력되었을 경우에 미지의 샘플에 대한 소속도를 구하기 위해 사용된다.

$$d_j = \| S - Z_j \| \quad (19)$$

여기서 S는 미지의 샘플 벡터이며 j=1,2, ..., c 표준 패턴의 중심과 샘플 패턴과의 소속도는 다음 식(20)에 의해서 구할 수 있다.

$$V_j = \frac{1}{\sum_{i=1}^n \left[\frac{d_{ij}}{d_i} \right]^2} \quad (20)$$

(20)에서 V_j는 j번째 클러스터에 대한 샘플 벡터의 소속도를 나타낸다.

표 1. 맥주에 대한 입력 데이터.

Table 1. Input data of beers.

종류 \ 센서	센서1	센서2	센서3	센서4	센서5	센서6	센서7
라거1	46.5	-61.4	-52.3	-89.0	-16.5	228.5	178.9
라거2	39.8	-59.7	-53.4	-85.3	-18.4	217.0	179.7
라거3	41.5	-63.7	-51.1	-93.3	-19.4	221.7	180.6
코로나1	41.4	-49.4	-51.1	-91.8	-12.5	191.5	170.6
코로나2	40.4	-47.1	-48.3	-87.5	-14.6	189.6	178.3
코로나3	41.2	-46.7	-49.3	-89.8	-13.5	189.4	177.8
기네스1	26.6	-58.5	-38.5	-79.0	-11.6	208.4	174.3
기네스2	24.2	-54.1	-32.8	-76.8	-12.4	224.9	182.3
기네스3	20.9	-53.4	-35.7	-77.2	-13.1	222.7	179.1
레드라1	32.9	-64.2	-46.5	-95.4	-16.2	196.2	171.4
레드라2	34.6	-64.6	-45.2	-92.5	-16.1	195.3	172.2
레드라3	33.2	-65.7	-47.0	-94.4	-15.9	198.6	173.8

V. 실험 결과

1. 맥주에 대한 PCA 분석

표 1은 센서를 통해 들어온 맥주에 대한 입력데이터를 나타냈고 이 데이터를 PCA를 통해 시물레이션 하였다.

우선 입력데이터에 대한 고유값과 고유벡터를 구한다. 맥주에 대한 입력데이터에 따른 고유값은 총 7개로 7차원이지만 3차원으로 축소하여 특징을 얻기 위해 7개의 고유값 중 큰 순서로 3개만을 취했으면 그에 상응하는 고유벡터 3개만을 얻어 3차원으로 매핑했다. 이렇게 7차원의 입력데이터를 PCA를 통하여 특징만을 추출하여 3차원의 패턴 클러스터링한 데이터를 표 2에서 보여주고 있다. 그림 4는 PCA의 결과이고 같은 특징의 데이터간에 클러스터링되어 있는 것을 볼 수 있다. 이렇게 고유값과 고유벡터를 이용하여 PCA를 사용한 결과 시각적으로 뚜렷이 데이터의 클러스터링 된 모습을 확인하였다.

2. PCA 결과에 의한 FCM 분석

위와 같이 PCA 과정을 거쳐 차원 축소된 매핑데이터는 다시 FCM의 입력데이터로 사용하여 각 데이터의 센터값과 그에 따른 입력데이터의 소속도를 수치적으로 확인 할 수 있다. 위에서 보았던 표 2의 PCA를 통해 3차원으로 출력된 결과값은 다시 맥주에 대한 FCM의 입력데이터로 다시 사용된다. 표 3은 FCM 입력데이터에 따른 각 클러스터의 센터를 나타냈고 표 4는 입력값과 센터값의 사이의 거리를 나타내는 소속도를 보여주고 있다. 표 4에서 보는 바와 같이 각각의 맥주의 센터에 대한 입력데이터의 소속도를 수치적으로 확실하게 확인할 수 있고 얼마나 소속해 있는가를 쉽게 판단할 수 있다.

표 2. 맥주에 대한 PCA 분석 결과.

Table 2. PCA analysis results of beers.

종류	차원1	차원2	차원3
라거1	317.55	-1.3548	-3.1097
라거2	308.06	0.34679	-0.60142
라거3	314.89	2.3532	-5.5616
코로나1	284.35	14.734	5.8177
코로나2	285.43	12.794	12.058
코로나3	285.79	14.743	11.95
기네스1	299.57	-17.706	-1.9161
기네스2	306.23	-25.147	4.8463
기네스3	302.87	-24.426	3.4417
레드락1	290.48	9.6711	-8.9671
레드락2	289.55	8.9531	-7.9272
레드락3	293.67	8.6099	-8.9884

표 3. 클러스터의 센터.

Table 3. Cluster center.

종류	주성분1	주성분2	주성분3
라거	313.98	0.41914	-3.3531
코로나	285.26	14.035	10.132
기네스	303.22	-22.905	2.5512
레드락	291.25	9.0387	-8.5773

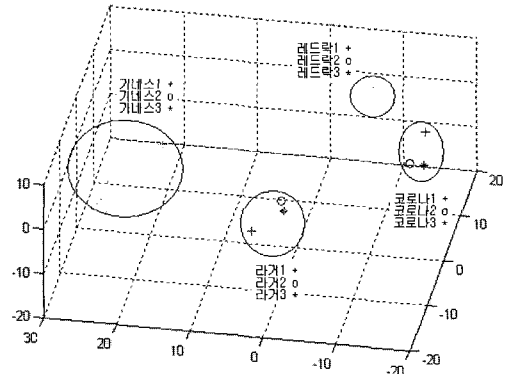


그림 4. 맥주에 대한 PCA 분석 그래프.

Fig. 4. PCA analysis graph of beers.

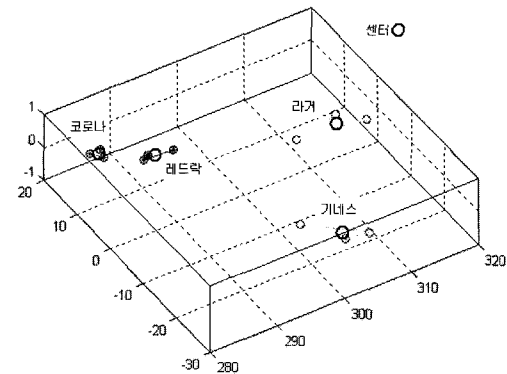


그림 5. PCA 결과에 의한 FCM 결과 그래프.

Fig. 5. FCM graph of PCA result.

표 4. 입력데이터 소속도.

Table 4. Membership grades of input data.

센터	라 거 1	라 거 2	라 거 3	코 로 나 1	코 로 나 2	코 로 나 3	기 네 스 1	기 네 스 2	기 네 스 3	레 드 락 1	레 드 락 2	레 드 락 3
라 거	0.8134	0.9553	0.9918	0.0101	0.0030	0.0022	0.0215	0.0605	0.0108	0.0008	0.0024	0.0049
코 로 나	0.0361	0.0092	0.0017	0.9110	0.9816	0.9861	0.0104	0.0422	0.0072	0.0027	0.0088	0.0129
기 네 스	0.0911	0.0236	0.0040	0.0155	0.0042	0.0032	0.9497	0.8148	0.9769	0.0017	0.0047	0.0117
레 드 락	0.0592	0.0117	0.0022	0.0631	0.0109	0.0083	0.0182	0.0823	0.0149	0.9446	0.9839	0.9703

표 5. 미지의 데이터.

Table 5. Unknown data.

	센서1	센서2	센서3	센서4	센서5	센서6	센서7
데이터	16.2	95.5	65.04	71.3	9.45	14.7	27.68

위의 그림 5는 PCA 결과에 의한 FCM 결과를 3차원으로 나타내고 있고 센터와 입력데이터들과의 소속정도를 그림으로 보여주고 있다. 지금까지 PCA를 통하여 7차원의 데이터를 3차원으로 축소한 후에 고유벡터를 이용하여 시각적으로 표시하고 수치적으로 볼 수 없는 단점을 FCM을 통하여 소속도를 정량적으로 표현하여 표준 데이터를 얻었다.

3. 미지의 데이터에 대한 PCA-FCM 분석

지금까지 PCA-FCM을 통해 미리 알고 있던 맥주의 데이터로 표준데이터 값을 만들었다. 이 표준데이터를 이용하여 새로 들어온 미지의 데이터가 어떠한 데이터에 속해 있는지를 다시 PCA를 통해 시각적으로 확인한 후에 FCM을 통해 수치적으로 얼마나 소속해 있는지를 판단할 수가 있다.

표 5에 나타난 데이터는 알 수 없는 미지의 데이터를 보여주고 있다. 표 5에 나타난 미지의 데이터가 어떠한 데이터 인지를 PCA로 시뮬레이션 한 모습은 그림 6과 같다.

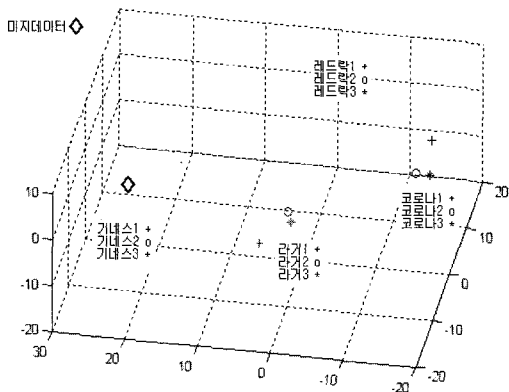


그림 6. 미지 데이터의 PCA 결과 그래프.

Fig. 6. PCA result graph of unknown data.

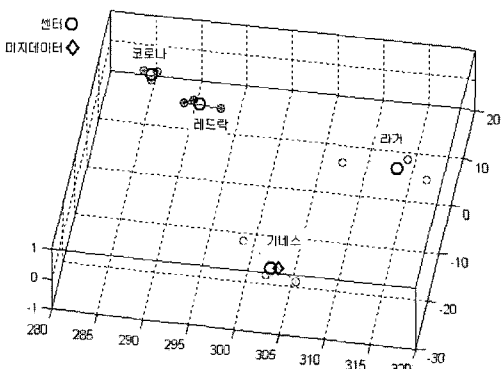


그림 7. PCA 결과에 의한 미지 데이터의 FCM 결과.

Fig. 7. FCM result of unknown data through PCA.

표 6. 미지 데이터의 소속도.

Table 6. Memberships grade of unknown data.

	라거	코로나	기네스	레드락
소속도	0.0032207	0.0011879	0.99392	0.001667

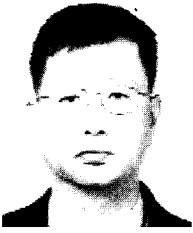
그림 6에서 보는 바와 같이 미지의 데이터가 속한 클러스터를 시각적으로 뚜렷이 확인할 수 있다. 이렇게 PCA를 통해 얻어진 미지 데이터의 결과는 다시 FCM 알고리즘을 통해 수치적으로 해석될 수 있고 그 결과인 소속도를 표 6에 나타내었다.

VI. 결론

본 연구에서 살펴본 전자 혀 시스템은 인간의 혀의 기능인 맛을 구별하는 미각기능을 구현한 시스템이다. 이온 선택성 전극의 원리에 의한 센서 어레이와 16bit A/D를 통한 데이터를 이용하여 각 센서에서 읽어들이는 패턴을 가지고 분석하였다. 그 결과 소수의 센서를 가지고 센서 수보다 훨씬 더 많은 맛을 구별할 있었다. 또한 센서들의 패턴을 분석을 위해 시각적인 분석인 PCA와 정량적인 분석인 FCM을 통해 데이터를 클러스터링하는 방법을 살펴보았다. 선형 분석 방법인 PCA를 통해 입력데이터의 차원을 축소하고 3차원으로 표현하여 시각적으로 특징을 확인할 수 있게 하였으며 데이터의 성질에 대한 수치적인 표현하지 못하는 PCA의 단점을 FCM과정을 거쳐 클러스터의 센터를 통하여 입력데이터의 소속도를 표현할 수 있는 것을 실험을 통하여 확인해보았다. 본 연구에서 보여준 전자 혀 시스템으로 인간이 직접 맛으로 확인 하지 않아도 되는 환경이나 직접 확인하기 힘든 물질들을 처리해야 하는 부분에서 이 시스템이 많이 쓰여 질 수 있을 것이다. 향후발전 과제는 표준 데이터의 밀도를 확인하여 입력데이터의 소속도를 처리하는 부분과 미각센서의 외부환경에 대한 반응을 고려한 어떠한 환경에도 영향을 받지 않고 측정할 수 있는 시스템 구현이라 생각되어진다.

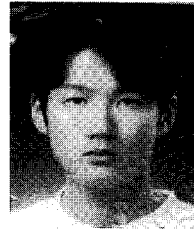
참고문헌

- [1] A. V. Legin, A. M. Rudnitskaya, G. Vlasov Y, C. Di Natale, A. D'Amico, "The features of the electronic-tongue in comparison with characteristics of the discrete ion-selective sensors", *Sensors and Actuator B*, 1999, V. 58, pp. 464-468.
- [2] G. N, Man, "A fuzzy neural network combining wavelet denoising and PCA for sensor signal estimation", *원자력 학회지*, 32권 5호, pp. 485-494, 2000.
- [3] 안강식, "내부 클러스터를 이용한 개선된 FCM 알고리즘에 대한연구", *정보처리학회 논문지*, 9권 2호, pp. 205-214, 2002.
- [4] R. W. Cattrall, "Chemical sensors", Oxford University Press, pp. 4-23, 1997.
- [5] D. T. Sawyer, A. Sobkowiak, J. L. Robert jr. "Electrochemistry for chemists", *John Wiley & Sons inc.*, pp. 24-52, 1995.
- [6] J. W. Gardner, P. N. Bartlett, "Electronic noses principles and applications", *Oxford Science Publications*, 1991.
- [7] L. Lvova, S. S. Kim, A. Legin, Yu. Vlasov, G. S, Chaa, H. Nam, "All-solid-state electronic tongue its application for beverage analysis", *Analytica Chimica Acta*, 2002, vol. 468 Iss. 2, pp. 303-314.



김 정 도

1987년 성균관대 전자공학과 졸업. 동대학원 석사(1990). 동대학원 박사(1995). 1995년~2004년 삼척대 컴퓨터응용 제어 공학과 교수. 2004년~현재 호서대 전기 정보통신공학부 교수. 관심분야는 센서 응용 및 로봇틱스.



정 우 석

2003년 호서대 전기정보통신공학부 졸업. 동대학원 석사과정(2004). 관심분야는 지능 알고리즘, 센서응용.



홍 천 호

1977년 성균관대 전자공학과 졸업. 동대학원 석사(1981). 동대학원 박사(1989). 1983년~1989년 인덕대 전자과 교수. 1989년~현재 호서대 전기정보통신공학부 교수. 관심분야는 로봇틱스 및 시스템 제어.