

## EEG기반 언어 인식 시스템을 위한 국제 음성기호를 이용한 모음 특징 추출 연구

### EEG based Vowel Feature Extraction for Speech Recognition System using International Phonetic Alphabet

이태주 · 심귀보<sup>†</sup>

Tae-Ju Lee and Kwee-Bo Sim<sup>†</sup>

중앙대학교 전자전기공학부

School of Electrical and Electronics Engineering, Chung-Ang University

#### 요 약

인간과 기계를 연결하는 새로운 인터페이스인 Brain-computer interface (BCI)를 이용해 휠체어를 제어하거나 단어를 입력하는 등, 사용자를 위한 다양한 장치를 개발하는 연구들이 진행되어 왔다. 특히 최근에는 뇌파를 이용한 음성인식을 구현하고 이를 통해 무음 통신 등에 적용하려는 시도가 있었다. 본 논문에서는 이러한 연구의 일환으로 electroencephalogram (EEG) 기반의 언어 인식 시스템을 개발하기 위한 기초 단계로서, 국제음성기호에 기반을 둔 모음들의 특징을 추출하는 방법에 대한 연구를 진행하였다. 실험은 건강한 세 명의 남성 피험자를 대상으로 진행되었으며, 한 개의 모음을 제시하는 첫 번째 실험 과정과 두 개의 연속된 모음을 제시하는 두 번째 실험 과정으로 두 단계에 나누어서 실험이 진행되었다. 습득된 64개의 채널중 선택적으로 32개의 채널만을 사용해 특징을 추출하였으며, 사고 활동과 관련된 전두엽과 언어활동에 관련된 측두엽을 기준으로 영역을 선택하였다. 알고리즘 적용을 위해서 특징으로는 신호의 고주 값을 사용하였고, support vector machine (SVM)을 이용하여 분류를 수행하였다. 실험 결과, 첫 번째 단계의 실험을 통해서, 언어의 뇌파를 분석하기 위해서는 10차원 이상의 특징 벡터를 사용해야 됨을 알게 되었고, 11차원의 특징 벡터를 사용할 경우, 평균분류율은 최고 95.63 %로 /a/와 /o/를 분류할 때 나타났고, 가장 낮은 분류율을 보이는 모음은 /a/와 /u/로 86.85 %였다. 두 번째 단계의 실험에서는 두 개 이상의 모음을 발음하는 것이 단일 모음 발음과 어떤 차이가 있는지 확인해 보았다.

**키워드** : 뇌-컴퓨터 인터페이스, 음성인식, 뇌파, 상상 음성.

#### Abstract

The researchs using brain-computer interface, the new interface system which connect human to machine, have been made to implement the user-assistance devices for control of wheelchairs or input the characters. In recent researches, there are several trials to implement the speech recognitions system based on the brain wave and attempt to silent communication. In this paper, we studied how to extract features of vowel based on international phonetic alphabet (IPA), as a foundation step for implementing of speech recognition system based on electroencephalogram (EEG). We conducted the 2 step experiments with three healthy male subjects, and first step was speaking imagery with single vowel and second step was imagery with successive two vowels. We selected 32 channels, which include frontal lobe related to thinking and temporal lobe related to speech function, among acquired 64 channels. Eigen value of the signal was used for feature vector and support vector machine (SVM) was used for classification. As a result of first step, we should use over than 10th order of feature vector to analyze the EEG signal of speech and if we used 11th order feature vector, the highest average classification rate was 95.63 % in classification between /a/ and /o/, the lowest average classification rate was 86.85 % with /a/ and /u/. In the second step of the experiments, we studied the difference of speech imaginary signals between single and successive two vowels.

**Key Words** : Brain-computer interface, Speech recognition, Electroencephalogram, Imagined speech.

접수일자: 2013년 9월 1일

심사(수정)일자: 2013년 9월 7일

게재확정일자 : 2014년 2월 13일

<sup>†</sup> Corresponding author

본 논문은 본 학회 2013 추계 학술대회에서 선정된 우수논문입니다.

본 논문은 한국연구재단 중견연구지원사업(No.2012-0008726)에서 지원하여 연구하였습니다. 연구비 지원에 감사드립니다.

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.

## 1. 서 론

Brain-computer interface (BCI)는 인간과 기계 사이를 뇌의 활동에서 발생하는 전기적 신호를 이용하여 연결 해 주는 새로운 인터페이스로, 세계 각국에서 활발한 연구가 진행되고 있다[1]. 그 중에서도 특히 비침습적이고 사용자 친화적인 Electroencephalogram (EEG)를 이용해 휠체어를 제어하거나 단어를 입력하는 장치가 개발되어 BCI 구현의 대표적인 사례로 잘 알려져 있다. 특히 BCI는 사용자의 움직임 심상을 획득할 수 있어 사용자가 외부 장치의 움직임을 생각만을 이용하여 작동하는데 유용하기 때문에 신체 활동에 어려움이 있는 사람을 위한 보조기구나, 사람이 직접 활동하기 힘든 작업 환경에 대응할 수 있는 장치 개발을 위한 연구들에 적용되고 있다.

최근에는 움직임 상상을 이용한 장치뿐만 아니라, 음성인식분야에서도 BCI의 활용 영역이 점차 확대되고 있다. BCI 시스템을 이용하여 특정 음성을 말하고자 할 때 발생하는 뇌파를 구분할 수 있다면 기존에 많은 연구가 진행된 음성인식 기술과 결합해 기계 제어 및 더 진보된 인간-기계간 상호작용을 구현할 수 있을뿐더러, 발생에 장애가 있는 사람들을 위한 의료용 재활기구로도 사용할 수 있는 길이 열릴 것이다. 특히 미국 국방고등연구계획국(DARPA)에서는 Irvine의 University of California (UCI) 연구팀과 함께 뇌파를 이용하여 소리 없이 병사들 사이에 통신을 할 수 있도록 하는 연구를 진행하였다[2]. 이 같은 사례를 통해, 언어 인식 분야로도 BCI를 활용 할 수 있음을 알 수 있다.

BCI를 이용한 언어인식 시스템을 구현하기 위해서 진행된 연구들을 살펴보면, 사용자가 발생에 사용하는 근육의 움직임을 상상할 때 발생하는 신호를 분석하는 방법과 사용자가 직접 언어를 상상할 때 발생하는 신호를 습득하는 두 가지 접근법으로 크게 나뉘어 연구가 진행되고 있음을 볼 수 있다[3-4]. 움직임 상상을 통한 접근법의 경우에는, 기존에 진행되어 왔던 많은 움직임 상상 연구와 결합하여 연구를 진행하기가 수월하고 상대적으로 쉽게 결과를 얻을 수 있다는 장점이 있다. 그러나 다른 신체 부위와 달리, 얼굴이나 목의 경우 말 뿐만 아니라, 표정을 짓는 등의 행동을 하므로 노이즈가 발생하기 쉽다. 또 말을 하는 행동에 대한 연구가 아닌 근육 움직임에 의해 발생하는 신호를 연구하게 됨으로써, 본질적으로 언어와 관련된 신호를 습득한다고 보기에 힘들다. 따라서 언어 인식 시스템을 구현하기 위해서는 후자와 같이, 사용자가 언어를 상상할 때 발생하는 신호를 습득하여 분석하는 과정이 필요하다. 이러한 접근은 움직임 상상에 기반을 둔 접근에서 나타나는 단점을 피할 수 있지만, 음성언어와 상상한 언어 사이의 관계를 정확히 찾아 낼 수 있는지, 다양한 뇌파 발생 패턴을 구분할 수 있는가와 같은 난관을 극복해야 된다[5].

본 논문에서는 이러한 단점들을 극복한 EEG 기반의 언어인식 시스템을 개발하기 위한 그 기초 단계로서, 국제 음성기호에 기반을 둔 모음 구분 방법에 대한 연구를 진행 하였다. 기존의 연구들은 영문 혹은 자국의 언어에 기반을 두어 연구가 진행돼, 전 세계 언어에 대해 일반화하기 힘들다는 단점이 존재한다. 따라서 본 논문에서는 연구 결과를 한글을 비롯한 다양한 언어에서 사용이 가능하도록 국제음성기호(International Phonetic Alphabet)를 기준으로 발음을 제시하고, 이에 의해 발생하는 뇌파를 분류하였다. 4 종류의 단일 모음을 분류하는 실험과 함께, 더 긴 단어 및 어휘의

분류 가능성을 확인하고자 두 글자로 된 모음을 피험자에게 제시하여 한 글자로 된 모음을 제시했을 경우와 어떤 점이 달라지는지 확인하는 실험을 수행하였다.

이어지는 2장에서는 본 논문에서 사용한 알고리즘들에 대한 이론적 배경을 소개하고, 3장에서는 실험을 통한 데이터의 습득 과정에 대해 기술하였다. 4장에서 실험의 결과를 분석하여 마지막 5장에서 본 논문의 결론을 제시하였다.

## 2. 분류 알고리즘

EEG 신호에서 특징 추출을 하기 위해서 본 논문에서는 covariance matrix의 eigen value를 사용하였다.  $t$  시간동안  $n$ 개의 전극에서 데이터를 얻을 경우, 데이터 행렬  $X$ 는  $t \times n$  행렬로 표현된다. 이때의, covariance matrix  $\Sigma$ 는 다음 식 (1)과 같이 구할 수 있다.

$$\Sigma = E[(X - E[X])(X - E[X])^T] \quad (1)$$

여기서  $E[\cdot]$ 은 행렬의 expectation을 의미하고  $T$ 는 transpose를 의미한다. 이 계산식을 통해  $\Sigma$ 은  $n \times n$  차원의 행렬임을 알 수 있다. 따라서  $\Sigma$ 의 eigen value는  $n \times 1$  차원의 벡터로 표시될 수 있다.

특징 추출 과정을 통해서 얻어낸 벡터들의 클래스를 분류하기 위한 알고리즘으로 support vector machine (SVM)을 사용하였다. SVM은 교차 학습을 기반으로 하는 분류기법이다. 두 클래스로 이루어진 데이터가 있을 때, 각 클래스 최외각 데이터를 나타내는 벡터인 support vector 사이의 여백이 최대가 되는 hyperplane을 설정하는 것을 통해 분류를 수행하는 알고리즘이다. SVM의 기본이 되는 수식은 다음 식 (2)와 같다[6].

$$y_i = w \cdot x_i + b \quad (2)$$

여기서  $x_i$ 는 학습을 위한 support machine들 중에서  $i$  번째 벡터를 의미하고  $y_i$ 의 값이  $\pm 1$ 일 때는 어느 한 쪽 클래스에 속한다는 것을 의미한다.  $y = 0$ 인 경우는 특징 공간에서 두 클래스를 나누는 hyperplane을 나타내게 된다.  $w$  벡터는 hyperplane에 수직한 벡터가 되어, 두 support vector 사이에 존재하는 여백의 크기는  $2/\|w\|$ 로 나타내어진다. 그러므로 SVM을 푸는 것은 다음 식(3)과 같이  $\|w\|$ 의 값을 최소화하는  $w$ 와 hyperplane의 offset을 결정하는  $b$  값을 찾는 것이다. 이는 최적화 문제로 바꾸어 풀 수 있으며, 일반적으로 Karush-Kuhn-Tucker 조건을 이용하여 풀게 된다.

$$\begin{aligned} \min P(w, b) &= \frac{1}{2} w^2 \\ \text{subject to } \forall i, & y_i(w \cdot x + b) \geq 1 \end{aligned} \quad (3)$$

현대적인 SVM 알고리즘에서는 커널 함수를 활용하여 선형적인 데이터뿐만 아니라 비선형적인 데이터도 원활히 분류가 가능하다. 또 cross validation을 이용하면 작은 크

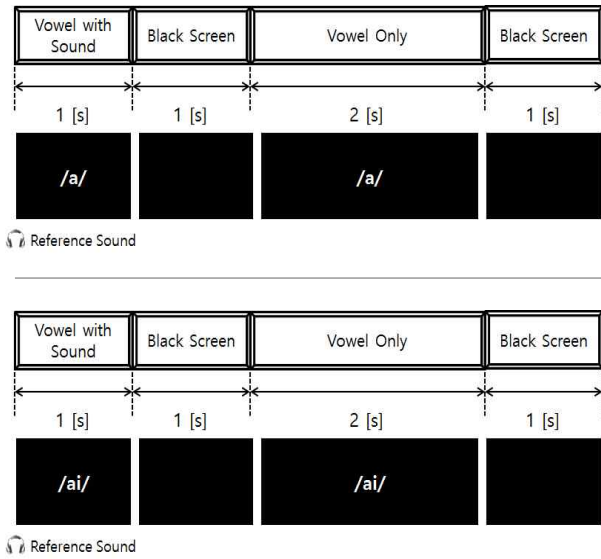


그림 1. 각 실험에 사용된 timeflow와 그 때의 모니터 화면에서 (위-첫 번째 실험 과정/아래-두 번째 실험 과정)  
 Fig. 1. Timeflow of each experiment and example of monitor screen at the time. (up-first step of the experiments /below-second step of the experiments)

기의 데이터에서도 높은 분류 결과를 얻을 수 있다. 본 논문에서는 주어진 데이터를 5-fold하여 SVM 알고리즘을 수행하였다.

### 3. 실험방법

EEG를 이용한 모음 구분을 하기위하여 실험에서는 국제 음성기호에 따라 분류된 모음 중, 가장 두드러지게 발음이 구분되는 /a/, /i/, /u/, /o/ 네 개의 모음을 사용하였다. 만 25세에서 27세 연령대의 세 명의 건강한 남성 피험자를 대상으로 뇌파를 수집하는 실험이 진행되었다. 모든 피험자들은 오른손잡이이며, 이전에 운동심상과 관련된 BCI 실험에 참여한 적이 있다. 실험 중에 피험자들은 의자에 편안한 자세로 앉아, LCD 모니터 화면을 응시하도록 하였으며, 발음을 할 때 얼굴의 움직임을 제외한 기타 움직임들은 최소화하도록 하였다. 실험 진행 중에 피험자는 피드백을 받지 않도록 하였고, 신호의 분석은 모든 습득과정이 끝난 후, 오프라인에서 이루어 졌다. 뇌파신호를 습득하기 위해 사용한 장비는 NeuroScan사의 Stim2와 Synamps2이고, 신호 처리 및 분석에는 같은 회사의 Curry7과 MathWorks의 Matlab을 사용하였다. 피험자의 두피에 접촉하는 전극은 international 10-20 system을 따르는 64채널 cap을 사용하여 측정되었으며, 추가로 5개의 전극을 배치해 눈의 깜빡임 등에서 발생하는 잡음을 제거하였다.

실험은 크게 두 과정으로 이루어졌다. 첫 번째 실험 과정은 피험자에게 하나의 모음만을 제시하는 것이고, 두 번째 실험 과정은 연속된 두 개의 모음을 제시하는 것이다. 첫 번째 실험 과정은 그림 1의 위쪽에 있는 timeflow에 따라, 실험이 진행되었다. 한 epoch에서 한 모음이 화면에 두 번 표시 되는데, 첫 번째 표시 될 때에는 표준 음성과 함께 1초간 제시되고, 두 번째 표시 될 때에는 음성 없이 화면에만

글자가 2초 동안 나타나게 된다. 화면에 표시된 모음은 검은색 배경에 노란색 글씨로, 영문자 /a/, /i/, /o/, /u/ 형태로 나타나게 되며, 모음 등장 사이의 시간 간격은 1초이다. 모음이 화면에 표시되는 동안 피험자는 화면에 나타난 문자를 말하는 상상을 수행한다. 첫 번째 제시된 모음은 사용자의 학습과 집중력 향상을 위해 제시되었으며, 피험자들이 통일된 발음을 상상하는 것을 유도하기 위하여 표준음성과 같이 제시되도록 하였다. 한 번의 실험동안 모음은 4종류의 모음 중 임의의 순서대로 각 9번씩 등장하게 되며, 총 5번의 실험을 수행해 뇌파를 습득하였다. 결과적으로 각 모음별로 145개의 데이터를 사용하게 되고, 전체 실험 데이터는 580개가 된다. 두 번째 과정에서는 그림 1의 아래와 같이, 첫 번째 실험 과정의 패러다임을 그대로 사용하되, 제시하는 두 개의 모음을 피험자에게 제시하였다. 이때 사용한 모음은 /ao/, /ui/, /oa/, /iu/의 네 종류로, 피험자가 제시된 모음에서 특별한 심상을 떠올리지 않도록 한국어에서 의미를 갖지 않는 모음 조합을 사용하여 실험을 진행하였다. 이번 과정 역시, 임의의 순서대로 자극이 제시되었으며, 각 자극은 6번 등장하였다. 총 실험을 4회 반복하여 데이터를 습득하였다. 따라서 모음 자극별 72개의 데이터를 얻었으며, 총 실험 데이터의 개수는 288개이다. 실험에 사용된 전극은 64채널 cap을 사용하여 피험자의 모든 영역에서 뇌파를 습득하였고, Curry7을 이용해 언어 및 사고와 관련된 브로카 영역 및 전두엽을 포함한 32개의 전극만을 선택, 사용하였다. 분석에는 자극이 주어지고 난 후 첫 1초 동안의 데이터를 이용하였다. 실험에 사용된 sampling rate는 250 Hz이었다.

데이터 특징 추출 과정에서 특징 벡터가 너무 커지거나 작아지지 않도록 특징 벡터를 normalize하는 과정을 거치도록 하였다. 각각의 차원에서 가장 큰 값으로 나눈 후, 가장 작은 값과 가장 큰 값이 -30부터 30까지의 값 사이에 매핑될 수 있도록 하였다. 또, 실험의 첫 번째 단계에서 분류를 수행할 때 SVM 알고리즘이 사용되었다. 현재 multi class 데이터에 적용하기 위한 SVM 알고리즘 연구가 진행되고 있으나, 선도적인 방법이 없기 때문에 두 개의 클래스에 대해서 분류가 가능한 일반적인 SVM을 선택하여 분류를 수행하였다. 따라서 분류는 /a/, /i/, /u/, /o/ 중 두 개씩 선택하여 수행되었으며, 총 6개의 분류 결과를 얻을 수 있었다. 또한 SVM 알고리즘이 학습에 의해 진행되고 최적화 문제를 푸는 과정이 포함되어 있어 항상 같은 결과를 얻을 수 없고 지역 최적해에 수렴하는 경우가 발생하므로, SVM 알고리즘을 10회 반복하여 평균 분류 결과를 이용하였다.

두 번째 실험단계에서는 두 개의 연속된 모음을 하나의 단일 모음에서 습득한 특징 벡터를 이용해 분류 가능한가에 대한 실험이었다. 가능성을 확인하기 위해서 두 개의 단일 모음 벡터들을 이용하여 SVM을 학습시키고, 연속된 두 개의 모음 데이터를 0.1초 단위로 나누어 학습된 SVM을 이용해 분류를 한다. 이러한 과정을 통해서, 특정 시간과 특징 벡터가 유의미하게 관계가 있음을 확인할 수 있으면, 연속된 언어상상이라도 시간별로 나누어 분석하는 것을 통해 단일 모음의 집합으로 바꾸어 볼 수 있을 것이다.

### 4. 실험결과 및 고찰

특징 추출된 데이터의 차원이 32이므로, 최적화된 EEG 신호 분류를 하기위하여, 몇 차원의 고유벡터를 사용해야

표 1. 11차 특징벡터에서 평균 피험자 분류 성공률  
Table 1. Average classification rate with 11th order feature vector

	/i/	/u/	/o/
/a/	94.2222 %	86.8519 %	95.6296 %
/i/	-	94.0741 %	92.8889 %
/u/	-	-	87.7778 %

되는지 찾기 위해 1부터 32까지 특징 벡터의 차원을 변화시켜가며, 그 분류 정확도를 확인하였다. 결과는 그림 3에서 보이는 바와 같다. 예상 할 수 있는 대로, 특징 벡터의 차원이 낮을수록 낮은 분류 성공률을 보이고 있다. 이는 낮은 차원을 선택할수록 데이터가 갖고 있는 정보량이 손실되기 때문으로 보인다. 점차로 특징 벡터의 차원이 커지면서 분류 성공률도 같이 증가하는 경향을 보이고 있으나, 특정 구간에서 극적으로 분류율이 증가하는 모습을 보이고 있어 이 구간 이후에는 특징 벡터의 차원을 증가시켜도 분류율의 향상이 미미한 것을 확인 할 수 있다. 특징 벡터의 차원이 높을수록, 계산량이 증가하여 시스템 구현에 어려움이 있기에, 적당한 범위 내에서 특징 공간의 차원을 결정해야할 것이다. 본 논문에서는 80 % 이상의 분류 성공률을 보인 11차원의 데이터를 사용했다. 이때의 분류 성공률을 표 1에서 확인 할 수 있다. 입모양이 비슷한 /o/와 /u/의 경우 87.7778 %의 분류율을 보여주고 있으며, 그림 3 및, 표 1의 결과처럼 /a/와 /u/가 다른 모음 분류 결과에 비해 상대적으로 좋은 결과를 얻지 못하는 것을 볼 수 있다. 이는 입모양이나 발음이 비슷한 모음에 대해서 뿐만 아니라, 전혀 다른 입모양 혹은 발음을 갖고 있는 모음 사이에도 분류 성공률을 낮추는 요인이 존재한다는 것을 보여준다. Pei의 연구 결과에 따르면[5], 기존 연구에서는 60 % 미만의 분류 성공

표 2. 11차 특징벡터에서 피험자별 분류 성공률  
Table 2. Classification rates with 11th order feature vector of each subject

	/i/	/u/	/o/
/a/	84.4444 %	71.8889 %	79.8889 %
/i/	-	84.7778 %	77.1111 %
/u/	-	-	75.0000 %

subject A

	/i/	/u/	/o/
/a/	75.0000 %	71.4444 %	91.3333 %
/i/	-	79.1111 %	94.3333 %
/u/	-	-	91.4444 %

subject B

	/i/	/u/	/o/
/a/	85.8889 %	71.1111 %	82.4444 %
/i/	-	87.7778 %	88.0000 %
/u/	-	-	82.3333 %

subject C

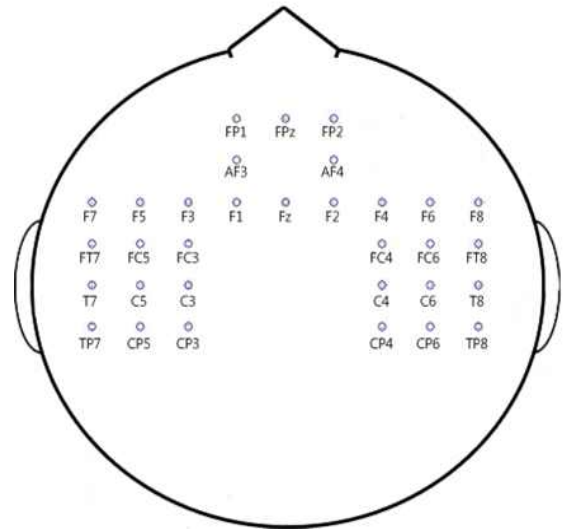


그림 2. 실험에 사용된 전극의 배치도  
Fig. 2. The location of electrode used for experiment

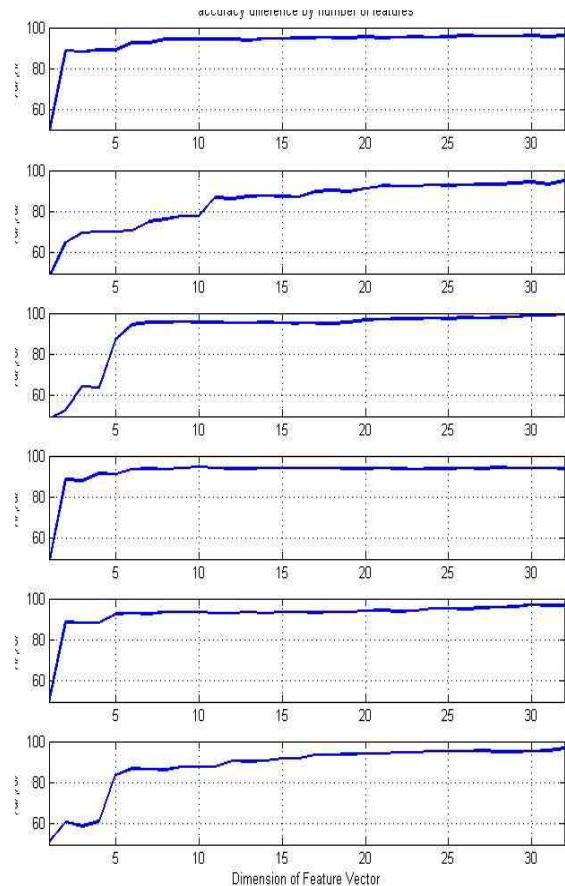


그림 3. 특징 벡터의 차원에 따른 분류 정확도  
Fig. 3. Classification accuracy rate by increasing of feature vector dimension

를을 보이기 때문에 제안한 방법이 더 효과적인 것을 알 수 있다.

또, 피험자 별로 결과를 살펴보면 각 피험자마다 상이한 결과를 보여주는 것을 확인 할 수 있다. 비교하는 모음에 따라 17 % 이상의 분류율 차이를 보여주는 경우도 있음을 볼 수 있는데, 실험이 통제된 환경에서 이루어졌고 여러 번 반복해서 수행되었기 때문에 이는 피험자 사이에서 발생할 수 있는 집중력 차이 혹은 실험에 대한 이해도 차이라고 보기는 어렵다. 특히 피험자 A의 경우에는 상대적으로 낮은 수치를 보였다. 운동 심상을 이용한 다른 BCI 패러다임의 경우, 80-100% 수준의 정확도를 얻을 수 있는 사람이 전체 피험자의 19%에 이르는 선형 연구에 비추어 보았을 때 [7], 피험자 A는 심상과 관련된 BCI 시스템 학습이 어려운 유형의 피험자인 것으로 생각된다. 그러나 작은 크기의 데이터를 사용했음에도 불구하고, 세 명의 피험자 모두 70 % 이상의 분류 성공률을 보이고 있고, 전체 평균 분류 성공률 역시 최소 80 %가 넘는 성공률이므로 언어인식 시스템 구현에 있어 국제음성기호에 기반을 둔 발음 체계를 사용하는 것이 학습이 어려운 유형의 피험자에게도 어느 정도 효과를 보인다고 할 수 있겠다.

두 번째 단계의 실험에서는 연속된 두 개의 모음이 단일 모음에서 얻어진 특징을 이용해 표현이 가능한지에 대한 여부를 확인하는 것이었다. 실험을 수행한 네 종류의 모음 /ao/, /ui/, /oa/, /iu/를 각각 /a/와 /o/, /u/와 /i/, /o/와 /a/, 그리고 /i/와 /u/에서 얻어진 특징 벡터로 학습한 SVM을 이용해 0.1초 단위로 분류를 수행하였다. 그림 4는 두 번째 단계의 실험 결과를 보여준다. 실선으로 된 것은 연속된 모음 중에서 먼저 발음된 모음의 SVM 분류 성공률을 나타내고, 점선으로 된 것은 두 번째 발음된 모음의 SVM 분류 성공률을 의미한다. 분류 성공률이 높다는 것은 해당 시간대

에 어떤 모음의 특징이 더 잘 나타난다는 것을 의미한다. 모든 결과에서 두 모음의 분류 성공률의 크기가 역전되는 경우가 없어, 뇌파의 변화 양상이 두드러지게 나타난다고 말하기는 어렵다. 그러나 네 종류의 그래프를 보았을 때, 0.2초에서 0.4초까지 그래프의 변화 양상이 어느 정도 유사한 점을 보이고 있다. 먼저 발음된 실선의 그래프는 첫 얼마의 시간동안 증가하는 것을 보이고, 그 이후 감소하게 된다. 두 번째 발음된 모음은 이와 반대의 경향을 보이고 있다. 이 역시 Pei의 연구에 따르면[5], 0.25초에서 0.5초 사이에 언어를 상상할 때 뇌가 가장 활성화되는 것을 알 수 있다. 따라서 이 시간동안 증가와 감소의 경향성은 언어활동과 관련이 있음을 알 수 있고, 이러한 경향성을 이용한다면 추가적인 연구에서 시간에 따른 모음을 구분해 낼 수 있는 단서로 사용 가능할 것이다. 재미있는 것은 모음이 어떤 순서로 발음되었는가에 따라 모음의 분류율이 높아지는 것이 아니라, 어떤 모음인가에 따라 그 분류율이 결정된다는 것이다. 실험의 첫 번째 결과와 세 번째 결과, 그리고 두 번째 결과와 네 번째 결과를 비교했을 때, 두 연속된 모음들은 등장 순서가 바뀌어 있으나, 주로 높은 분류 결과를 내는 것은 같은 모음이다. 이를 고려한다면, 언어는 시간적인 순서와 관계없이 사용자가 생각한 언어의 종류에 따라 사용자의 뇌파가 활성화 되는 것이라 할 수 있다. 기존에 단어를 대상으로 구분을 하였던 기존 연구에서는 block화 된 단어의 경우, 어휘의 내용보다 temporal correlation patterns에 영향을 받는다는 결과가 있는데[8], 단일 모음에서는 이와 반대되는 흥미로운 결과를 얻을 수 있었다. 이러한 결과를 고려한다면, 언어인식 시스템을 구현하기 위해서는 단어보다 개별 음소를 분석할 수 있는 방향으로 연구가 진행되어야 한다고 볼 수 있다.

### 5. 결론 및 향후연구

본 논문은 EEG를 이용한 언어인식 시스템을 구현하고자, 그 첫 단계로서, 모음을 인식하는 방법에 대한 연구를 수행하였다. 보다 다양한 언어에 적용하기 위하여 특정한 언어에 종속적으로 연구가 진행되었던 기존 방식과 달리, 국제음성기호에 기반을 두어 뇌파 신호를 추출하였으며, 다양한 언어에 적용할 수 있다. 본 논문에서는 국제음성기호의 네 가지 모음인 /a/, /i/, /o/, /u/를 오프라인 방식을 이용해 서로 간에 분류를 수행하였으며, 이 때 해당 분류율이 특징 벡터의 차원에 영향을 받는다는 것을 확인하였다. 모든 분류 결과가 처음으로 80 % 이상의 정확도를 보이는 특징 벡터의 차원은 11 차원이었으며, 이때의 최대 분류율은 약 /a/와 /o/의 약 95.63 %였으며, 최소 분류율은 /a/와 /u/의 약 86.85 %였다.

두 번째 실험을 통하여 시간에 따라 모음의 특징을 분류해 내는 것이 쉽지 않음을 확인할 수 있었다. 그러나 SVM을 이용한 분류율의 변화를 보았을 때, 자극이 주어지고 난 후 뇌가 언어에 대해 활성화되는 시간 동안, 4 종류의 모음 자극에서 발생하는 결과의 증감 양상이 비슷하게 나타나는 것을 볼 수 있었다. 이 부분에 대해서는 추가적인 연구와 구체적인 실험 패러다임을 구성하여 개선된 결과를 얻을 수 있도록 해야겠다.

본 논문에서 수행한 모음 구분을 통해 언어인식의 가능성이 있고, 둘 이상의 언어인식 또한 지속적인 연구를 통해

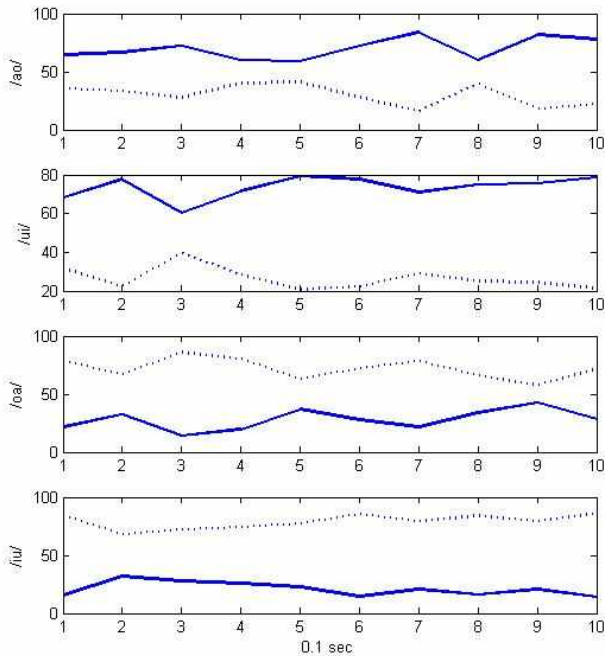


그림 4. 두 개의 단일 모음을 이용한 연속된 모음에 대한 SVM 분류 성공률

Fig. 4. SVM classification rate of successive two vowels using two single vowels

가능성이 있음을 확인하였다. 그러나 선행된 연구[5]에서 언급하고 있는 것처럼, 모든 발음을 뇌파 패턴을 이용해 표현할 경우, 패턴의 종류가 많아져 분류에 어려움이 발생하는 문제점은 여전히 남아있으며, 본 논문에서 얻은 결과에서 볼 수 있듯이, 기존의 특징 추출 방법을 사용하게 되면 10차원 이상의 특징 벡터를 사용해야 하므로, 시스템의 구현에 어려움이 발생할 것은 자명하다. 따라서 추후 연구 과제로 많은 언어를 인식하기 위한 특징 추출 방법과 multi-class 분류 알고리즘을 구현하는 방안에 대해서 추가적인 연구가 필요하다. 뿐만 아니라, 실제 사용에서는 off-line 방식이 아닌 on-line방식의 시스템 구현이 필요하다. 간소화된 장치를 이용하여 무선으로 실시간 처리를 할 수 있도록 연구를 진행할 예정이다.

### References

[1] J. R. Wolpaw, E. W. Wolpaw, "Brain-computer interfaces: something new under the sun," Brain-computer interfaces: principles and practice. Oxford University Press, Oxford, pp. 3-12, 2012.

[2] R. Bogue, "Brain-computer interfaces: control by thought," *Industrial Robot: An International Journal*, vol. 37, issue 2, pp. 126-132, 2010.

[3] K. Brigham, B. V. K. V. Kumar, "Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy," *Conf Bioinf. and Biomed. Eng.* 2010, pp. 1-4, 2010.

[4] N. Yhoshimura, A. Satsuma, C. S. DaSalla, T. Hanakawa, M. Sato, Y. Koike, "Usability of EEG Cortical Currents in Classification of Vowel Speech Imagery," *Int. Conf. Virtual Rehabilitation 2011*, pp. 1-2, 2011.

[5] X. Pei, J. Hill, G. Schalk, "Silent Communication: Toward Using Brain Signals," *IEEE Pulse*, vol. 3, issue 1, pp. 43-46, Jan. 2012.

[6] L. Bottou, C-. J. Lin, "Support Vector Machine Solvers," *Large-Scale Kernel Machines*, MIT Press, pp. 1-27, 2007.

[7] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabona, F. Gramatica, G. Edlinger, "How many people are able to control a P300-based brain-computer interface (BCI)?," *Neuroscience Letters*, vol. 462, issue 1, pp. 94-98, Sep. 2009.

[8] A. Porbadnigk, M. Wester, J. P. Calliess, T. Schultz, "EEG-based Speech Recognition - Impact of Temporal Effects" *2nd International Conference on Bio-inspired Systems and Signal Processing (Biosignals 2009)*, 2009.

### 저 자 소 개



#### 이태주(Tae-Ju Lee)

2013년 : 중앙대학교 전자전기공학부 공학사  
 2013년~현재 : 중앙대학교 대학원  
 전자전기공학부  
 석박사통합과정.

관심분야 : Brain-computer Interface, Intention recognition, Neuro-Robotics, Soft Computing 등.  
 Phone : +82-2-820-5319  
 E-mail : bindaman@cau.ac.kr



#### 심귀보(Kwee-Bo Sim)

1990년 : The University of Tokyo  
 전자공학과 공학박사  
 1991년~현재 : 중앙대학교 전자전기공학부 교수  
 2006년~2007년 : 한국지능시스템학회 회장

관심분야 : 인공생명, 뇌-컴퓨터 인터페이스, 의도인식, 감성인식, 유비쿼터스 지능형로봇, 지능시스템, 컴퓨터이셔널 인텔리전스, 지능형 홈 및 홈 네트워크, 유비쿼터스 컴퓨팅 및 센서 네트워크, 소프트 컴퓨팅(신경망, 퍼지, 진화연산), 다캐체 및 자율분산로봇시스템, 인공면역시스템, 지능형 감시시스템 등.  
 Phone : +82-2-820-5319  
 E-mail : kbsim@cau.ac.kr  
 Homepage URL : http://alife.cau.ac.kr