

# Independent Component Analysis of Electroencephalographic Data

인공지능학과 이지양

# 논문의 중요성

- EEG 데이터에 ICA 기법을 적용한 최초의 연구 중 하나
- 잡음(artifact) 제거뿐만 아니라 의미 있는 뇌 활동 성분 추출까지 가능해짐
- ERP 분석에도 ICA를 접목해 시간/공간적으로 중첩된 성분을 구분
- EEG 분석 패러다임의 전환
- ICA의 표준화

---

## Independent Component Analysis of Electroencephalographic Data

---

**Scott Makeig**  
Naval Health Research Center  
P.O. Box 85122  
San Diego CA 92186-5122  
scott@cpl.mmhrc.navy.mil

**Anthony J. Bell**  
Computational Neurobiology Lab  
The Salk Institute, P.O. Box 85800  
San Diego, CA 92186-5800  
tony@salk.edu

**Tzyy-Ping Jung**  
Naval Health Research Center and  
Computational Neurobiology Lab  
The Salk Institute, P.O. Box 85800  
San Diego, CA 92186-5800  
jung@salk.edu

**Terrence J. Sejnowski**  
Howard Hughes Medical Institute and  
Computational Neurobiology Lab  
The Salk Institute, P.O. Box 85800  
San Diego, CA 92186-5800  
terry@salk.edu

### Abstract

Because of the distance between the skull and brain and their different resistivities, *electroencephalographic* (EEG) data collected from any point on the human scalp includes activity generated within a large brain area. This spatial smearing of EEG data by volume conduction does not involve significant time delays, however, suggesting that the *Independent Component Analysis* (ICA) algorithm of Bell and Sejnowski [1] is suitable for performing blind source separation on EEG data. The ICA algorithm separates the problem of source identification from that of source localization. First results of applying the ICA algorithm to EEG and *event-related potential* (ERP) data collected during a sustained auditory detection task show: (1) ICA training is insensitive to different random seeds. (2) ICA may be used to segregate obvious artifactual EEG components (line and muscle noise, eye movements) from other sources. (3) ICA is capable of isolating overlapping EEG phenomena, including alpha and theta bursts and spatially-separable ERP components, to separate ICA channels. (4) Nonstationarities in EEG and behavioral state can be tracked using ICA via changes in the amount of residual correlation between ICA-filtered output channels.

# Abstract

- EEG 데이터의 공간적 확산(spatial smearing)
- Volume conduction & no significant time delay
- 독립성분분석(ICA) 알고리즘이 EEG 데이터 블라인드 소스 분리에 적합
- '신호의 식별' 문제와 '신호의 위치 추정 문제'를 분리
- EEG & ERP 데이터에 ICA 알고리즘 적용한 결과
  - 서로 다른 난수 시드에 민감하지 않음
  - 잡음성 EEG 성분을 다른 신호로부터 분리
  - 알파파와 세타파의 중첩된 활동, ERP 성분을 서로 다른 ICA 채널로 분리
  - EEG의 비정상성(nonstationarity)이나 행동 상태의 변화는 ICA 필터링 출력 채널 간의 잔여 상관(residual correlation)의 변화로 추적 가능

# Introduction

- EEG 신호원을 식별하려는 시도는 대부분 신호원의 공간적 분리와 위치 추정에 집중해 왔음
- ICA 알고리즘을 적용하여 신호 식별(어떤 신호인가?)과 신호 위치 추정(어디에서 발생한 신호인가?) 문제를 분리
- ICA 알고리즘은 물리적 위치나 구성을 고려하지 않고  
높이 상관된 EEG 신호들로부터 통계적으로 독립적인 신호원들을 도출
- EEG를 독립적이고 공간적으로 고정된 전위 생성 시스템들의 출력으로 간주

# ICA

- 임의의 벡터  $\mathbf{x} = [x_1 \dots x_N]^T$ 에 대해
- 선형 변환  $\mathbf{u} = \mathbf{W}\mathbf{x} + \mathbf{w}$ 의 원소  $\mathbf{u} = [u_1 \dots u_N]^T$ 가 독립이 되는
- 행렬  $\mathbf{W}$ 와 벡터  $\mathbf{w}$ 를 찾는 기법
- $f_{\mathbf{u}}(\mathbf{u}) = \prod_{i=1}^N f_u(u_i)$ : multivariate probability density function
- $I(u_i, u_j) = 0, \forall ij$ : mutual information go to zero
- Mutual information: 성분들의 모든 고차 통계량을 고려하는 측도
- In EEG: 관측된 혼합 신호로부터 통계적으로 독립적인 신호원을 효과적으로 분리할 수 있게 하는 강력한 분석 기법

# ICA 알고리즘의 가정

- $u_i$ 는 scaling/shifting 이후 모두 동일한 누적분포함수 가진다.
- 우리는 그 분포 함수의 형태  $F_u(u)$ 를 알고 있다.

# ICA 알고리즘의 수행

- 비선형적으로 변환된 벡터  $y = F_u(u)$ 의 엔트로피  $H(y)$  최대화  
-> 행렬  $W$ 와 벡터  $w$ 를 조정하기 위한 확률적 경사 상승법 규칙 도출
- Stochastic gradient ascent rules

$$\Delta \mathbf{W} \propto [\mathbf{W}^T]^{-1} + \hat{\mathbf{y}} \mathbf{x}^T, \Delta w \propto \hat{\mathbf{y}} \quad (1)$$

where  $\hat{\mathbf{y}} = [\hat{y}_1 \dots \hat{y}_N]^T$ , the elements of which are:

$$\hat{y}_i = \frac{\partial}{\partial y_i} \frac{\partial y_i}{\partial u_i} \quad [\text{which if } \mathbf{y} = F_u(\mathbf{u})] = \frac{\partial f_u(u_i)}{\partial F_u(u_i)} \quad (2)$$

# EEG 데이터에 ICA 적용하기

- ICA 기법은 다음과 같은 조건 만족하는 영역에서 적합
  1. 신호원들 간에 독립성 존재
  2. 혼합 매질의 전파 지연이 무시할 수 있을 정도로 작음
  3. 신호원들이 아날로그 형태이며, pdf가 logistic sigmoid gradient와 유사
  4. 신호원 수가 센서 수와 같음
    - > EEG가 정확히 N개의 신호원의 선형 혼합이라고 할 수 있는가?
- ICA 결과를 해석할 때  
입력 채널 차원의 결정과 도출된 ICA 채널의 생리학적 의미가 중요



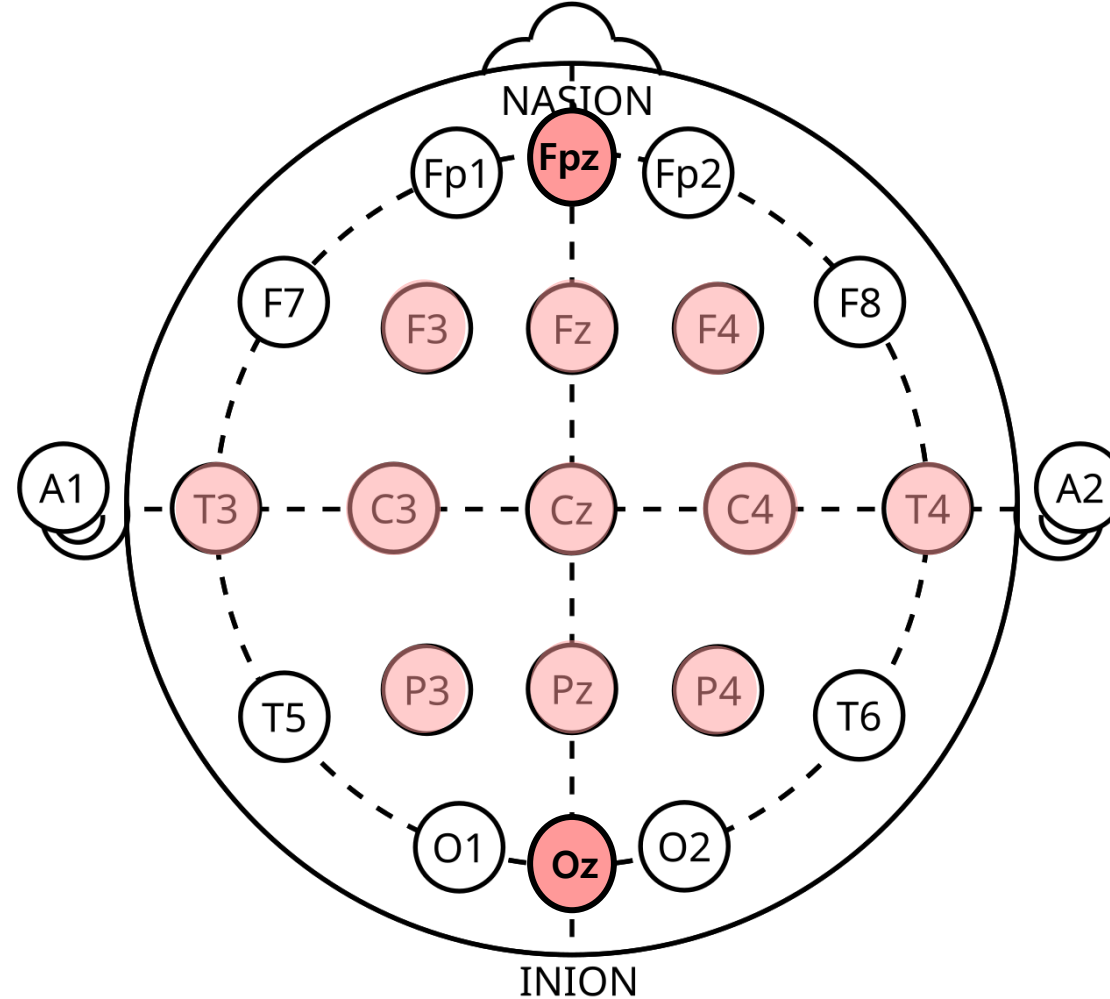
# EEG 및 행동 데이터 수집

- S.Makeig & M.Inlow (1993) Lapses in alertness: Coherence of fluctuations in performance and EEG spectrum
- 10명의 성인 자원자는 30분 분량의 세션에 3회 이상 참여
- 지속적으로 제시되는 배경 소음의 강도가 잠시 증가하는 청각 자극 감지할 때마다 버튼을 한 번 누르도록 지시받음
- 각성 상태 감소 관찰 가능성을 높이기 위해 아늑한 분위기에서 눈을 감은 상태로 진행
- 청각 자극:
  - 62dB 백색소음 배경의 350ms 간 강도 증가
  - 평균 10회/분의 빈도로 제시
  - 39Hz click train for 39Hz SSR
  - 과제와 무관한 짧은 톤도 타겟 자극 사이에 2~4초 간격으로 삽입

# EEG 및 행동 데이터 수집

- EEG는 International 10-20 System 상의 13개 전극에서 수집
- Right mastoid(우측 유양돌기)를 기준 전극으로 사용
- 샘플링 속도 312.5Hz
- EOG 채널도 함께 기록 – 눈 움직임으로 인한 잡음 보정 및 제거
- Hits: 탐지 성공. 타겟 자극 제시 후 100~3000ms 이내에 반응한 경우
- Lapses: 탐지 실패. 반응하지 못한 경우
- Local error rate:  
convolve the irregularly-sampled performance index time series(Hit=0/Lapse=1) with a 95s smoothing window advance for 1.64s steps

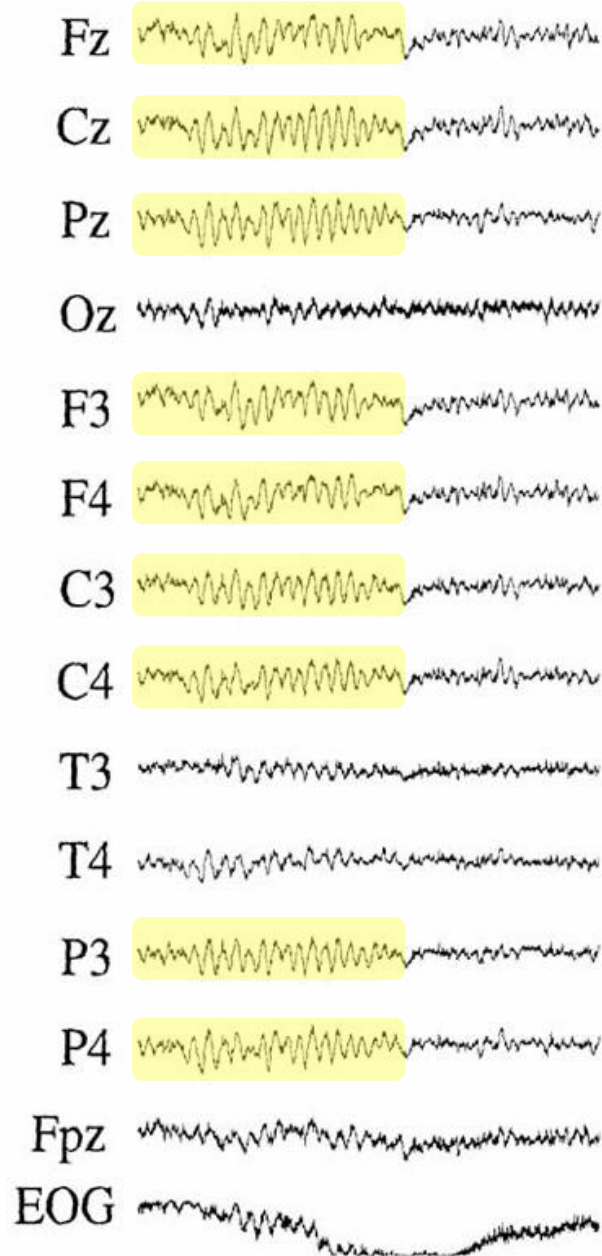
# International 10-20 System



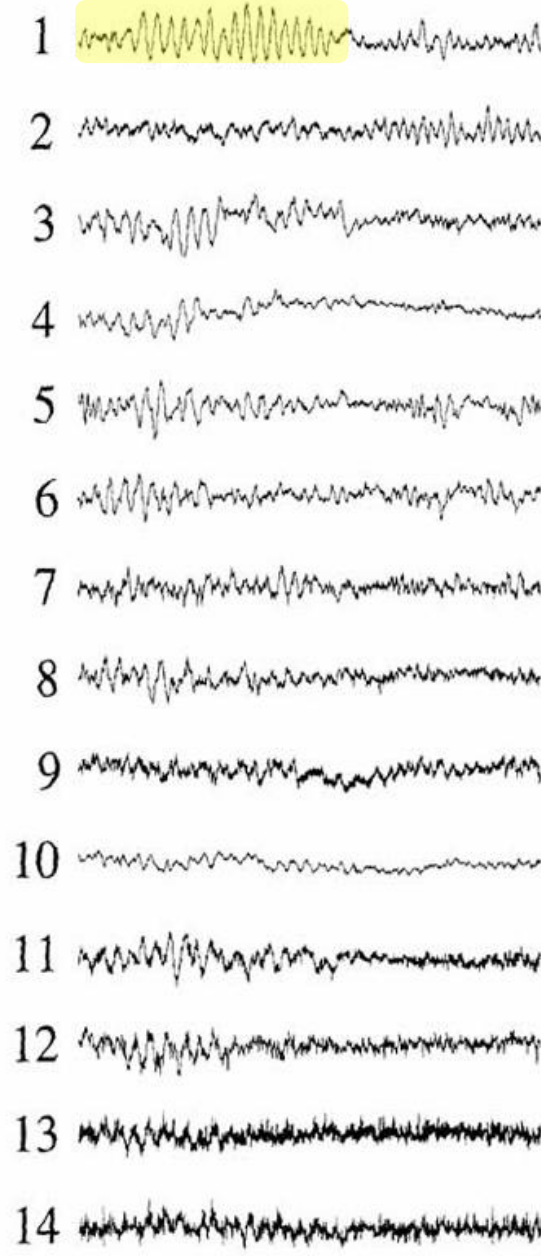
# ICA 알고리즘의 적용

- ICA 알고리즘을 14채널 EEG 기록에 적용
- 신호의 정상성(stationarity) 확보를 위해 시간 인덱스 permute
- 14차원 시점 벡터를 하나씩 14x14 ICA Network에 입력
- 데이터 pre-whitening 처리하여 1,2차 통계량 제거
- 학습률은 0.03에서 0.0001로 점진적 감소
- 학습 절차 중단
  1. ICA 출력 채널 간 평균 상관관계가 0.05 미만
  2. ICA 가중치 행렬에 더 이상 유의미한 변화 없음

## EEG



## ICA



세타파 (주요 성분)

알파파

짧은 진동성 활동들(3-8)

근육 관련 고주파 노이즈

느린 안구 운동 관련(EOG+Fpz)

60Hz 전력선 노이즈

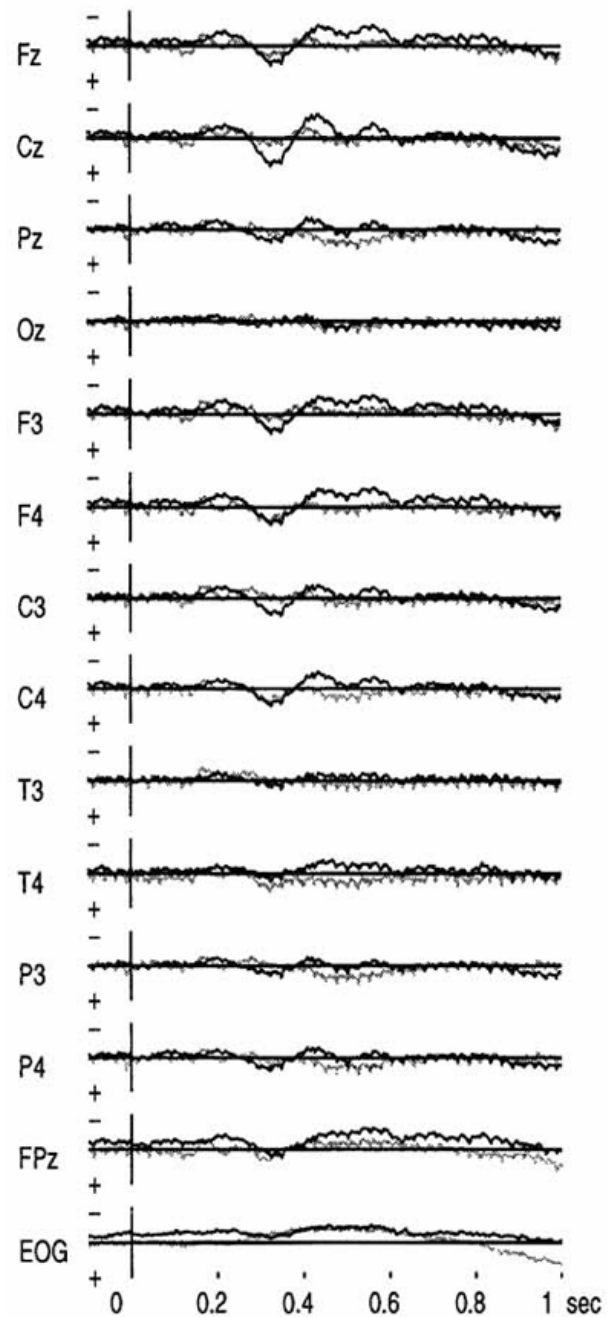
근육 관련 고주파 노이즈

— 1 sec.

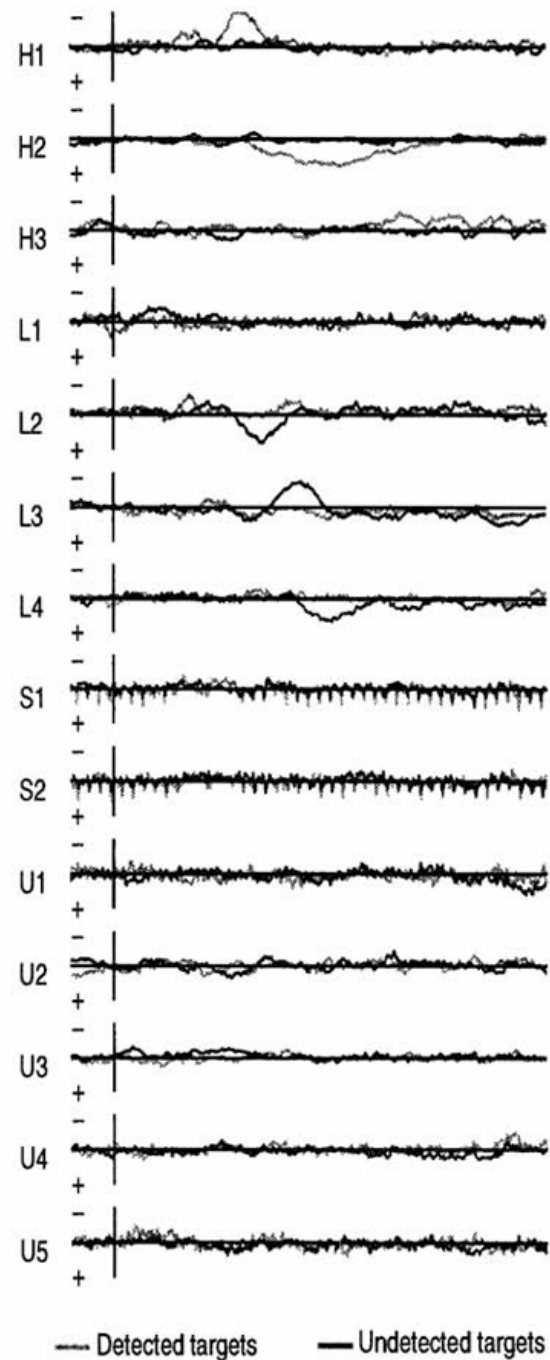
# 결과

- Source 해는 학습률, 초기 조건에 크게 영향을 받지 않음
- 같은 세션에서 다른 EEG 데이터 구간을 초기화 다르게 한 두 ICA 행렬로 필터링
- 출력 채널 14쌍 중 11쌍은 상관계수 0.95 이상, 최소 상관계수 0.894  
-> 다르게 훈련한 ICA 모델도 거의 동일한 분리 결과
- 각성 상태 데이터로 훈련한 모델은  
각성 상태 데이터에서는 출력 채널 간 상관관계 ↓  
졸림 상태 데이터에서는 출력 채널 간 상관관계 ↑
- 졸림 상태 데이터로 훈련한 모델은 반대의 양상
- 각성과 졸림 상태의 뇌 역학 및 공간 구조의 변화를 반영

Scalp ERPs



ICA ERPs



Hit:

시간적으로 제한된 반응 성분(H1-3)

Lapse:

상대적으로 더 큰 반응 성분(L1-4)

39Hz SSR(S1-2)

이 분리된 성분들은 뇌 반응활동 위치의 시간 경과를 반영한다.

이는 유발 반응을 시간적으로 겹쳐 있는 하위 구성 요소들로 나누는 하나의 해법이 될 수 있다.