Speech artifact removal from EEG recordings of spoken word production with tensor decomposition

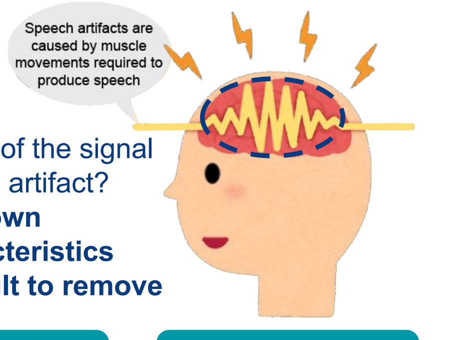
Holy Lovenia[1,3], Hiroki Tanaka[1,2], Sakriani Sakti[1,2], Ayu Purwarianti[3], and Satoshi Nakamura[1,2]

1) Nara Institute of Science and Technology, Japan
2) RIKEN, Center for Advanced Intelligence Project AIP, Japan
3) Department of Informatics, Bandung Institute of Technology, Indonesia

Email: holy.lovenia@gmail.com

1. Introduction

A. Background

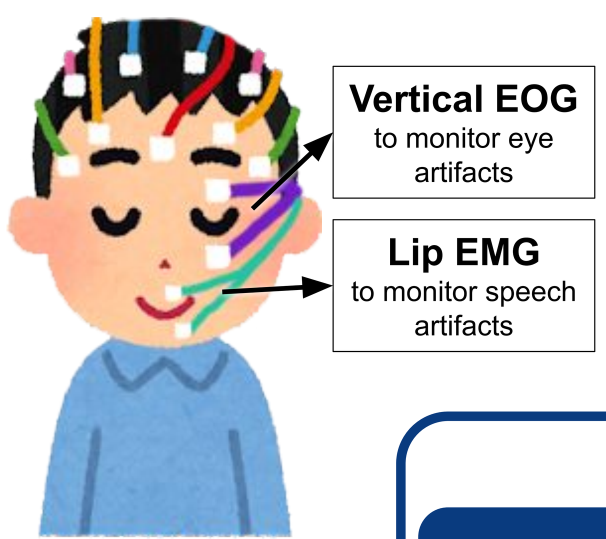


Which part of the signal is a speech artifact?
- **Unknown characteristics**
- **Difficult to remove**

Speech artifact contamination → Unable to process the EEG data

B. Overview

Electrode placement



Research pipeline

Data collection (picture-naming experiment) → **Preprocessing** → **Eye artifact removal** → **Speech artifact removal**: **Proposed method** (tensor-based); **Previous work methods**: **SAR-ICA**, **BSS-CCA** → **Evaluation** (grand-average correlation with lip EMG during 0-1350 ms)

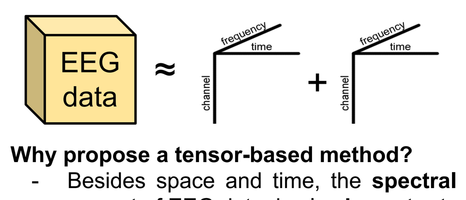
Related works

- **SAR-ICA** by Porcaro et al, 2015
- **BSS-CCA** by Vos et al, 2010

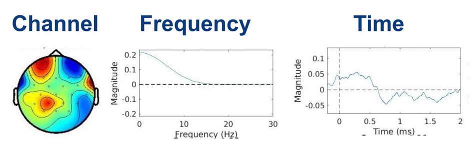
Both proposed a speech artifact removal method with matrix decomposition.

SAR-ICA > BSS-CCA, according to Porcaro et al, 2015

Tensor decomposition



Why propose a tensor-based method?
- Besides space and time, the **spectral aspect** of EEG data is also **important**
- The **multi-way nature** of EEG data



C. Research Objectives

1. **Propose a new method** for removing speech artifacts with **tensor decomposition** for sources reconstruction
2. **Evaluate the performance of the proposed method** against existing methods (**SAR-ICA** and **BSS-CCA**)

2. Methods

SAR-ICA (Porcaro et al, 2015)

Independent component analysis (ICA) → **Analysis**:
- **Visual inspection** (topographical distribution, single-trial, averaged trials)
- **Statistical calculation** (entropy, kurtosis, global kurtosis)
- **Spectral calculation** (PSD correlation with lip EMG)

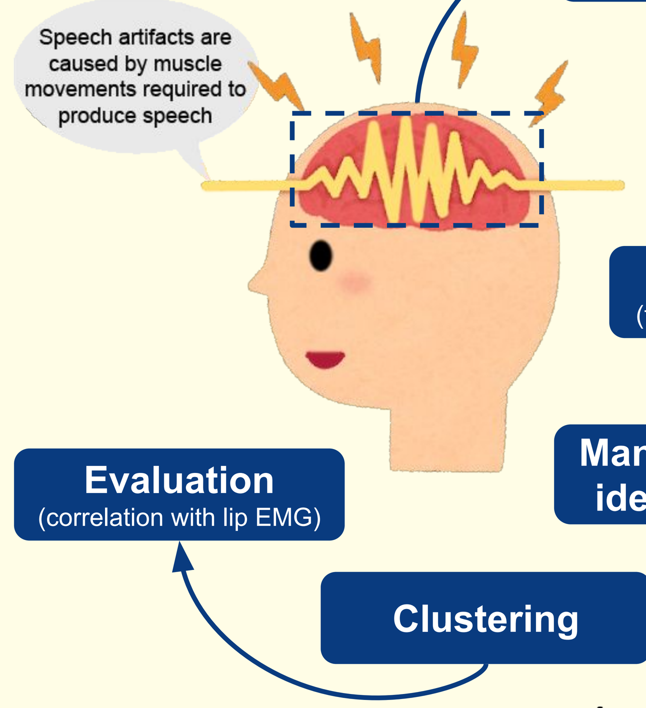
→ **Manual artifact detection** → **Control cycle** (to ensure only artifact is removed) → **Clustering** →

1. **Speech artifact cluster (SAR-ICA)**
2. **Cleaned data cluster (SAR-ICA)**
3. **Raw data without EOG cluster**

Eye artifact removal
Performed with SAR-ICA, using PSD correlation with EOG channel as the spectral calculation

A. Pipelines

Proposed method (with tensor decomposition)



EEG data → **DIFFIT** (by Timmerman et al, 2000) —Number of components→ **CPD tensor decomposition** → **Analysis** (time, spectral, & spatial) —Features→ **Manual artifact identification** —Predictions→ **Clustering** → **Evaluation** (correlation with lip EMG)

1. **Raw data without EOG cluster**
2. **Speech artifact cluster (CPD)**
3. **Cleaned data cluster (CPD)**

Post-removal validation
Manual observation of the difference between before and after removing the artifactual component

The difference of fit (DIFFIT)
- Number of components range: 5-26
- Estimated number of components (across subjects)
  - Lowest : 8
  - Highest : 24
  - Average : 15

Component features used in analysis
1. Visualization of channel mode in space
2. Visualization of frequency mode
3. Frequency mode correlation with lip EMG
4. Visualization of time mode
5. Time mode correlation with lip EMG

Time mode correlation calculated between

$$[SO_{avg}(subject), \min(T, SO_{avg}(subject) + 1)]$$

where
$SO_{avg}(subject)$ is averaged speech onset of the subject, and T is end of trial

Proves to be the **significant differentiator** in **method performance**

BSS-CCA (Vos et al, 2010)

PSD analysis (using quantitative criterion)

It is an **EMG activity if** $M \geq \frac{E}{n}$

where
- M is average power in EMG band (approx. by 15-30 Hz)
- E is average power in EEG band (approx. by 0.1-15 Hz)
- n is constant (default: 7)

Assumption: **EEG → lower power at high freq, EMG → higher power at high freq**

BSS-CCA decomposes EEG signals into sources in decreasing order of autocorrelation
- **EMG** activity is **weakly autocorrelated** over time
- **Brain** activity is **more autocorrelated**

Blind source separation (BSS) using canonical correlation analysis (CCA) as ground truth → **PSD analysis** → **Artifact marking** → **Automatic muscle-related artifact removal** → **Data reconstruction** → **Cleaned data cluster (BSS-CCA)**

B. Key Differences

Overview comparison of the existing methods and the proposed method (CPD tensor decomposition)

Aspect	Speech artifact removal method		
	SAR-ICA	BSS-CCA	Proposed method (CPD)
Decomposition method	ICA	BSS	**CPD**
EMG channels usage	Yes	No	**Yes**
Need of visual inspection	Yes	No	**Yes**

Matrix decomposition vs tensor decomposition as sources reconstruction method

Matrix decomposition (ICA & BSS) → 2 dimensional
Tensor decomposition (CPD) → multiple modes

The proposed method takes spectral features of speech artifact into account **during decomposition**, while the previous methods do it after decomposition

3. Experiment

EEG Data Collection

Experiment design

Task: picture-naming

Participants: 9 native Japanese speakers
- Mean age 23.3, SD 2.6
- 7 males, 2 females
- 8 right-handed, 1 corrected-to-right
- 1 excluded because of many errors

Recordings: 27 (+ 1 reference) EEG channels, 2 EMG channels, & 2 EOG channels

Picture stimuli: 45 line-drawings of common objects (Nishimoto et al, 2005)

Pre-experiment → **Picture-naming experiment**: **Practice block** → **Rest** → **Experiment block**

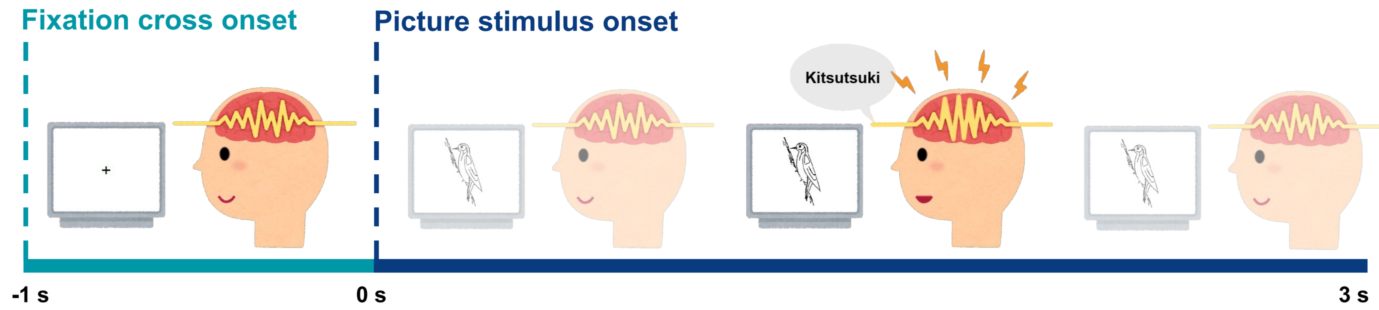
Preparation
- Read a booklet of picture-name pairs

Instructions
- Blink only after naming
- Reduce movements during experiment

Block design
- Consists of **45 trials**
- Displays each picture exactly once randomly
- Only EEG data from experiment block will be processed

Trial design

At the beginning of every trial, a fixation cross appears for one second. This period is called pre-stimulus. Afterwards, it is replaced with a picture stimulus. The participant is to name the displayed stimulus as quickly and as accurately as possible. The stimulus remains on the display for three seconds. A trial lasts for four seconds in total, the end of which leading to the beginning of the next trial.



4. Results

A. Assessments

Evaluation

Aim
To **compare performances** of all methods

Technique
Calculate the **Pearson correlation (R)** between **grand-average clusters and the lip EMG** during 0-1350 ms in the time domain

Interpretation
The higher the correlation, the stronger the association with lip EMG. Lip EMG is assumed to represent the original speech artifacts.
- **Speech artifact**: the **higher**, the better
- **Cleaned data**: the **lower**, the better

Advanced validation

Aim
To ensure the preservation of brain signals

Technique
Calculate **R** between **the cleaned data cluster and raw data without EOG cluster** during:
- 0-700 ms: before the earliest speech onset
- 0-900 ms: before the grand-average speech onsets

Interpretation
Preserved brain signals → **high** correlation

B. Performance Measure

Evaluation results

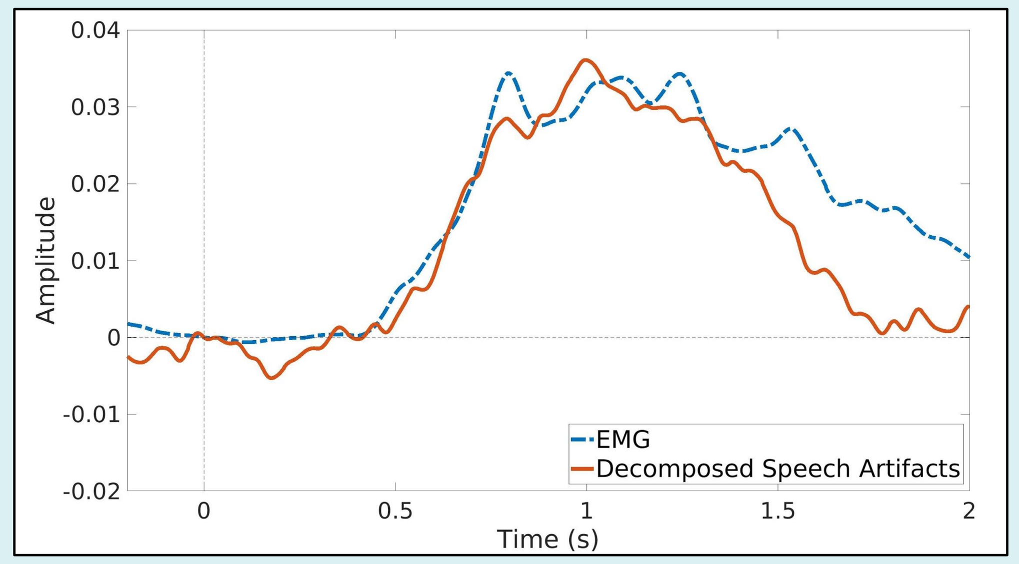
The table below presents the **detailed evaluation results**, which is the absolute Pearson correlation between speech artifacts and cleaned data with lip EMG in the time domain for 0-1350 ms ($p < 0.01$). **The proposed method is written as CPD.**

Grand-average cluster	R (0-1350 ms)
SAR-ICA's Speech Artifact	0.875
CPD's Speech Artifact	**0.985**
SAR-ICA's Cleaned Data	0.351
BSS-CCA's Cleaned Data	0.413
CPD's Cleaned Data	**0.101**

The proposed method (CPD) outperforms SAR-ICA and BSS-CCA, **both in detecting speech artifact** (0.985) **and cleaning data** (0.101)

Speech artifact by the proposed method

Below is the comparison of the normalized **grand-average lip EMG** and **decomposed speech artifacts by the proposed method**.



- **Almost identical** at most time points to lip EMG
- Slight difference → possibly caused by the inability of lip EMG channels to pick up every single speech movement

Advanced validation results

Grand-average cluster	R (0-700 ms)	R (0-900 ms)
CPD's Cleaned Data	0.927	0.942

The high validation results (0.92-0.94) during the pre-speech onset **indicate** the **quality** of the cleaned data is **enough for subsequent EEG processing**.

5. Closing

Sum Up and Future Direction

Conclusion

- We **proposed a speech artifact removal method** using **CPD tensor decomposition**
- **The proposed method (CPD) surpasses former methods** (SAR-ICA and BSS-CCA), both in identifying speech artifacts (0.985) and producing cleaned data (0.101)

Future works

Present study relied on manual visual inspection and EMG channels usage, so **future research should...**
- **Automate** all of the steps needed for speech artifact identification
- **Use other means** to determine the speech artifacts

6. References

C. Porcaro, M. T. Medaglia, and A. Krott, “Removing speech artifacts from electroencephalographic recordings during overt picture naming,” NeuroImage, vol. 105, pp. 171 – 180, 2015.

F. Cong, Q.-H. Lin, L.-D. Kuang, X.-F. Gong, P. Astikainen, and T. Ristaniemi, “Tensor decomposition of eeg signals: A brief review,” Journal of Neuroscience Methods, vol. 248, pp. 59 – 69, 2015. [Online].

M. de Vos, D. Maarten Vos, S. Ries, K. Vanderperren, B. Vanrumste, F.-X. Alario, S. Huffel, S. Huffel, and B. Burle, “Removal of muscle artifacts from EEG recordings of spoken language production,” Neuroinformatics, vol. 8, pp. 135–150, 06 2010.

M. E. Timmerman and H. A. L. Kiers, “Threemode principal components analysis: Choosing the numbers of components and sensitivity to local optima,” British Journal of Mathematical and Statistical Psychology, vol. 53, no. 1, pp. 1–16, 2000.

T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” SIAM Review, vol. 51, no. 3, pp. 455–500, September 2009.

T. Nishimoto, K. Miyawaki, T. Ueda, Y. Une, and M. Takahashi, “Japanese normative set of 359 pictures.” Behavior research methods, vol. 37 3, pp. 398–416, 2005.