# Behaviour of Empirical Mode Decomposition Algorithms for Classification of Single-Channel EEG Manifesting McGurk Effect

Arup Kumar Pal[1], Dipanjan Roy[2], G. Vinodh Kumar[2], Bipra Chatterjee[3], L. N. Sharma[1], Arpan Banerjee[2]*, and Cota Navin Gupta[3]*

[1] Indian Institute of Technology, Guwahati, India
EMST Lab, Dept. of Electronics and Electrical Engineering
[2] National Brain Research Centre, Gurgaon, India
Cognitive Brain Dynamics Lab
[3] Indian Institute of Technology, Guwahati, India
Neural Engineering Lab, Dept. of Biosciences and Bioengineering
* Both the senior authors spent equal time

**Abstract.** Brain state classification using electroencephalography (EEG) finds applications in both clinical and non-clinical contexts, such as detecting sleep states or perceiving illusory effects during multisensory McGurk paradigm, respectively. Existing literature considers recordings of EEG electrodes that cover the entire head. However, for real world applications, wearable devices that encompass just one (or a few) channels are desirable, which make the classification of EEG states even more challenging. With this as background, we applied variants of data driven Empirical Mode Decomposition (EMD) on McGurk EEG, which is an illusory perception of speech when the movement of lips does not match with the audio signal, for classifying whether the perception is affected by the visual cue or not. After applying a common pre-processing pipeline, we explored four EMD based frameworks to extract EEG features, which were classified using Random Forest. Among the four alternatives, the most effective framework decomposes the ensemble average of two classes of EEG into their respective intrinsic mode functions forming the basis on which the trials were projected to obtain features, which on classification resulted in accuracies of 63.66% using single electrode and 75.85% using three electrodes. The frequency band which plays vital role during audio-visual integration was also studied using traditional band pass filters. Of all, the Gamma band was found to be most prominent followed by the alpha and beta bands as in previous studies.

**Keywords:** EMD, EEG, McGurk effect, Random Forest

## 1 Introduction

Electroencephalography (EEG) is an electrophysiological technique to record the electrical signals within the brain. Although EEG signals exhibit very high temporal resolution in the order of milliseconds, they lack appreciably in spatial resolution because the neural activity detected from the surface of the skull is the summation of the excitatory and inhibitory postsynaptic potentials of billions of neurons firing simultaneously at different depths within the brain. So the signal remains contaminated with artefacts from various sources which include, but not constrained to, various physiological factors (like heart beat, respiratory pulses, eye movement, muscle movement, etc) resulting in poor Signal-to-Noise (SNR) ratio eventually making subsequent extraction of relevant information for analysis purposes troublesome. To reduce artefacts, digital filters may be applied, but with requisite precautions, as they tend to filter out the EEG activity of interest alongside and as a result contort EEG waveforms severely.

Humans get an idea of the external world by assimilating information reaching to the brain from multiple sensory organs. In case of perception of speech, along with the auditory signal, the movement of the speaker's lips plays a crucial role in influencing perception. For instance, when an auditory signal is presented alone or the visual stimulus presented along with it is congruent, then it is most likely to be correctly interpreted by the listener. But, when the same audio is presented along with a semantically incongruent visual signal, then the listener tends to hear a phoneme which is completely different from both the audio and visual cue - commonly known as the McGurk Effect [1], the rate of which is greatly affected by the relative timing of the audio and the video. In this work, we analysed the EEG recordings of 15 healthy subjects when they were presented with an audiovisual (AV) clip created by superimposing /**pa**/ as audio and lip movements generated by the utterance of /**ka**/ as visual stimuli and in most cases they confirmed of hearing /**ta**/.

Empirical Mode Decomposition (EMD) helps in decomposing a signal into different frequency bands, commonly known as Intrinsic Mode Functions (IMFs) [2]. Since the introduction of EMD by Huang et al., it has proven itself to be a very effective tool for processing and analysis of various non-stationary and non-linear signals like image, climate data, seismic waves, ocean waves, EEG, etc. to name a few [3], [4], [5], [6], [7]. It has already been reported in literature that the differences during unimodal and cross-modal perception of AV speech stimuli gets reflected primarily in the gamma band, with some information also in the alpha and beta band [8] while posterior superior temporal sulcus (pSTS) being the main region of human brain responsible for audio-visual integration [9]. But a suitable EEG-based classification algorithm which can decipher whether a subject perceives the McGurk effect or not was missing till date. As EMD has already been quite effective in extracting features from EEG signals in different other paradigms, we tried to implement various EMD based feature extraction techniques in this work to classify the EEG data from various subjects, recorded in McGurk paradigm. The feature vectors thus obtained are then classified using a non-linear classifier, Random Forest. A comparative study of the different feature extraction methods proposed has been presented here.

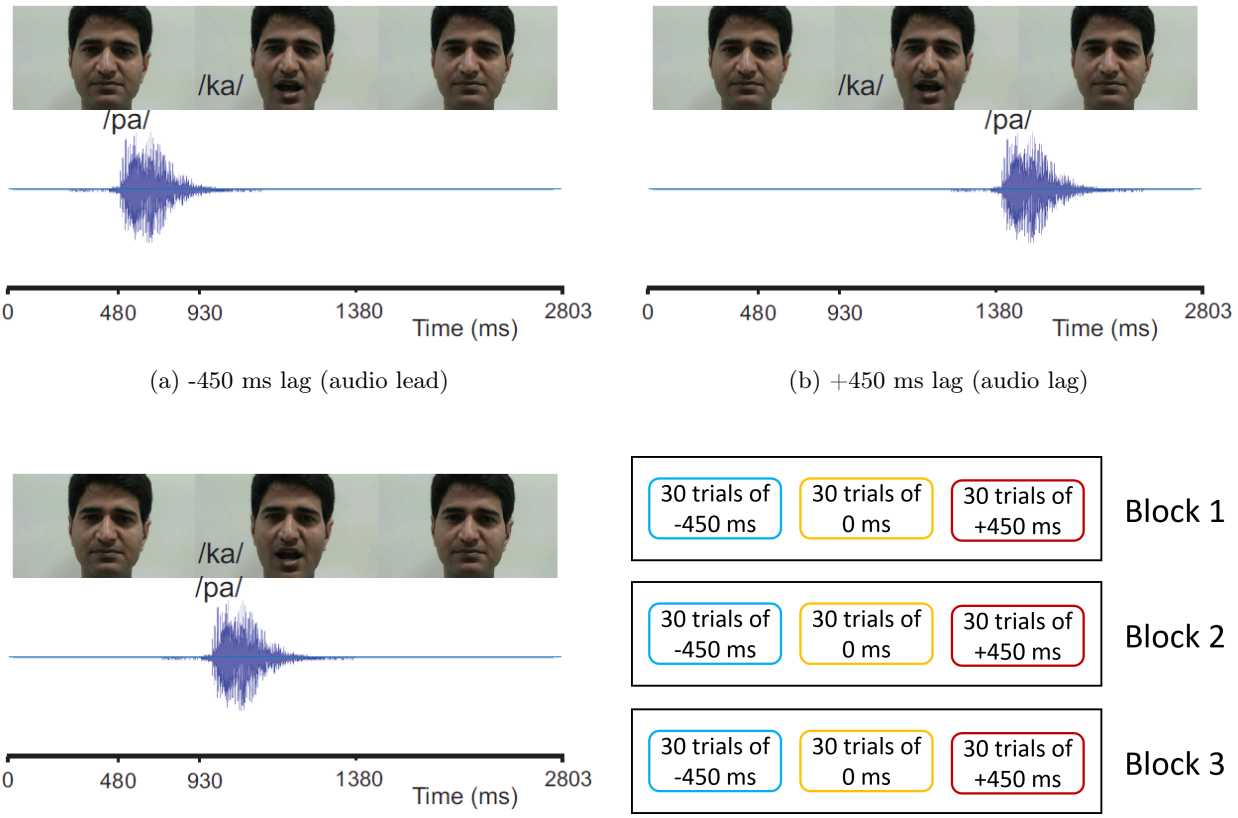
## 2 Data and Pre-processing

The AV stimulus presented to the subjects in this paradigm is created by superimposing an audio signal of /**pa**/ over the visual of lip movements formed while uttering /**ka**/, both of duration 450 ms in an audio-visual clip of 2803 ms. As the relative timing of the audio and movement of lips greatly affects the rate of cross-modal perception, the stimulus was presented in three variations :

- **−450 ms lag** where the audio signal leads the visual by 450 ms as in **Figure 1a**.
- **+450 ms lag** where the audio lags visual by 450 ms as in **Figure 1b**.
- **0 ms lag** where both the audio and visual stimuli are synchronous to each other as in **Figure 1c**.

30 repetitions of each of the variations of AV stimulus mentioned above were taken together and randomized to form a trial block. Each subject was presented with 3 such blocks. Thus, a total of $(30 \times 3) \times 3 = 270$ AV stimuli were presented as elaborated in **Figure 1d**. After each clip was played, the subjects had to confirm what they heard from three options: */ta/*, */pa/* and */others/*. The inter-trial interval was varied somewhat between 1200 and 2800 ms to minimize expectancy effects and during that time no clip was shown on the screen.

When the stimulus was played on the screen, simultaneous recording of EEG was done from the scalp by placing 64 electrodes in 10-20 fashion, mounted on a Neuroscan cap. Reference electrodes were linked




(a) -450 ms lag (audio lead)

(b) +450 ms lag (audio lag)

(c) 0 ms lag (synchronous)

(d) Representation of blocks of trials

Fig. 1: **a-c**: Different AV stimuli which were presented to the subjects, adapted from [10]. **d** - Pictorial representation of the trial blocks.

to mastoids and grounded to AFz. As the data from 64 electrodes (known as 64 channel EEG data) was recorded at a sampling frequency of 1 kHz while the stimulus being of 2803 ms duration, each trial of EEG signal turned out to be of dimension $2803 \times 64$ (for 64 channels).

After referencing, EEG data from 15 subjects was band-passed through a zero phase-shift offline filter of 0.2-200 Hz. After that, a notch filter with a cut-off frequency of 50 Hz was applied on the data. Next, a baseline correction was performed on the filtered data by removing the temporal mean of the EEG signal on an epoch-by-epoch basis. Epochs with a maximum amplitude above 100 µV or minimum below -100 µV were removed to eliminate the trials contaminated with ocular and muscle-related activities [11]. On removal of

| Lag ↓ \ Responses → | /**ta**/ | /**pa**/ |
|---|---|---|
| 0 ms | 622 | 200 |
| − 450 ms | 458 | 296 |
| + 450 ms | 456 | 288 |

Table 1: No. of trials left for analysis after artifact rejection.

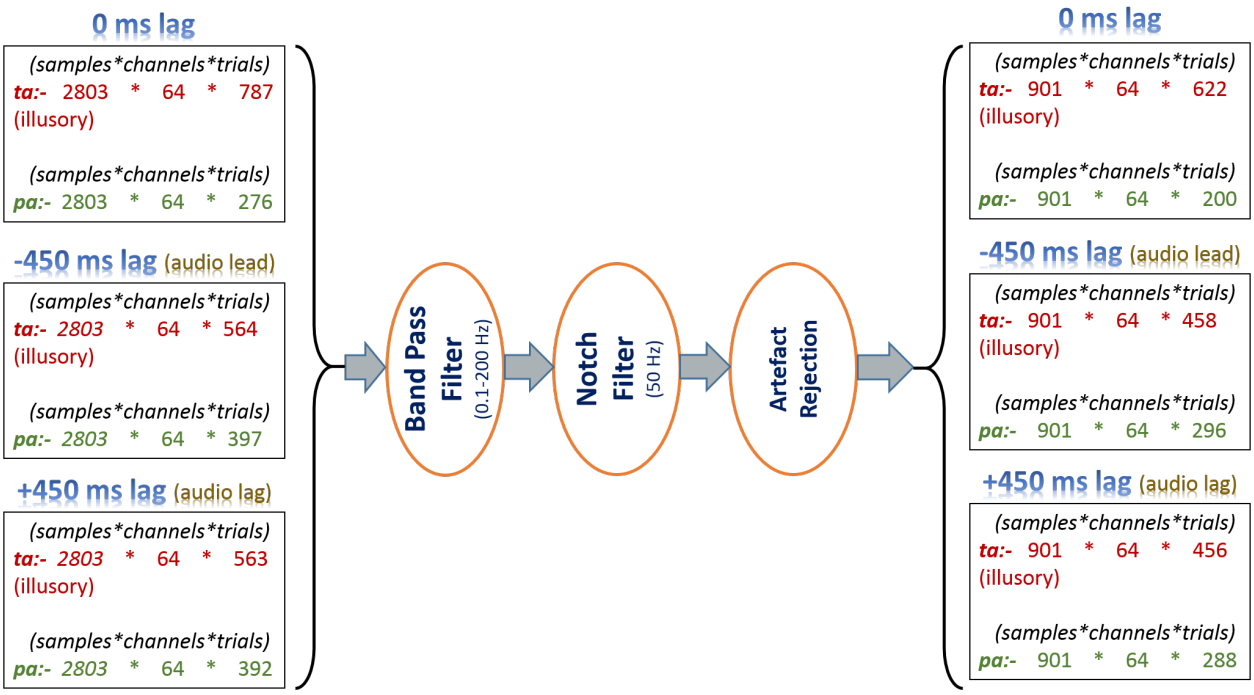



Fig. 2: A bird's eye view of the pre-processing steps applied on the EEG data.

the corrupted trials, we were left with the no. of trials as indicated in **Table 1** for further analysis. **Figure 2** sums up all the pre-processing steps described so far. Further details can be obtained in [8].

## 3 Methodology

Four EMD based algorithms proposed here are applied on electrodes separately, i.e. neither any multidimensional approach is applied, nor the feature vectors formed from different electrodes have been concatenated. Multidimensional scheme of classification has been avoided purposely in order to acquire the most relevant information from the minimum number of electrodes because researchers aim in reducing the number of electrodes without compromising on the quality of information gathered and thereby enhancing the portability of EEG sensors. Again, if the feature vectors from different electrodes have been concatenated, then the dimension of the feature vectors would have increased for the same number of trials and thereby intensifying the *'curse of dimensionality'*[12]. For these reasons, the algorithms discussed in this study describe the process of obtaining feature vectors for a single electrode of a trial. Separate RFs are trained for the different electrodes and after obtaining results from all the RFs (of all the electrodes), voting is done by the best electrodes to obtain a final label. From now on, in all the algorithms proposed, *trail* refers to a **single electrode** of the trial. While algorithm I is inspired by [7], algorithm II to IV find their inspiration from the temporal characteristics of cross-modal speech perception [8].

### 3.1 Algorithm I

The algorithm described in this section has been applied on P300 data in [7], another oddball paradigm like the McGurk effect and provided quite satisfactory results. The steps are discussed below:

1. All trails corresponding to /**ta**/ responses and /**pa**/ responses were ensemble averaged separately.

$$\mathrm{pa}_{\mathrm{avg}}(k) = \frac{1}{M_1}\sum_{i=1}^{M_1} x_{1,i}(k), \qquad \mathrm{ta}_{\mathrm{avg}}(k) = \frac{1}{M_2}\sum_{i=1}^{M_2} x_{2,i}(k) \tag{1.1}$$

   where $M_1$ and $M_2$ are the number of trials corresponding to /**pa**/ ($x_{1,i}(k)$) and /**ta**/ ($x_{2,i}(k)$) respectively.

2. EMD is applied to the 2 averaged signals, obtaining their IMFs:

$$\mathrm{pa}_{\mathrm{avg}}(k) = \sum_{i=1}^{N_1} c_{1,i}(k), \qquad \mathrm{ta}_{\mathrm{avg}}(k) = \sum_{i=1}^{N_2} c_{2,i}(k) \tag{1.2}$$

   where $N_1$ and $N_2$ are the number of IMFs obtained on decomposing $\mathrm{pa}_{\mathrm{avg}}$ and $\mathrm{ta}_{\mathrm{avg}}$ respectively, $c_{1,i}(k)$ and $c_{2,i}(k)$ being their respective IMFs.

3. Any trail $x(k)$ can be expanded on the basis formed by the IMFs as:

$$x(k) \simeq \sum_{i=1}^{N_1} b_{1,i}c_{1,i}(k) = \hat{x}_1(k), \qquad x(k) \simeq \sum_{i=1}^{N_2} b_{2,i}c_{2,i}(k) = \hat{x}_2(k) \tag{1.3}$$

   where $b_{1,i}$ and $b_{2,i}$ are the expansion coefficients based on the IMFs of averaged trials $\mathrm{pa}_{\mathrm{avg}}$ and $\mathrm{ta}_{\mathrm{avg}}$ respectively. The coefficients are calculated by the pseudo-inverse method and least square constraint as follows:

$$A_1 b_1 = \hat{x}_1, \qquad A_2 b_2 = \hat{x}_2 \tag{1.4}$$

$$\begin{aligned} A_1 &= \begin{bmatrix} c_{1,1}(0) & c_{1,2}(0) & \cdots & c_{1,N_1}(0) \\ c_{1,1}(1) & c_{1,2}(1) & \cdots & c_{1,N_1}(1) \\ \vdots & \vdots & \ddots & \vdots \\ c_{1,1}(K-1) & c_{1,2}(K-1) & \cdots & c_{1,N_1}(K-1) \end{bmatrix} \\ A_2 &= \begin{bmatrix} c_{2,1}(0) & c_{2,2}(0) & \cdots & c_{2,N_2}(0) \\ c_{2,1}(1) & c_{2,2}(1) & \cdots & c_{2,N_2}(1) \\ \vdots & \vdots & \ddots & \vdots \\ c_{2,1}(K-1) & c_{2,2}(K-1) & \cdots & c_{2,N_2}(K-1) \end{bmatrix} \end{aligned} \tag{1.5}$$

   where $K$ is the length of the signal.

4. From $A_1$, $A_2$ and $x$, the vectors of expansion coefficients can be constructed as:

$$\begin{aligned} b_1 &= [b_{1,1} \quad b_{1,2} \quad \cdots \quad b_{1,N_1}] = (A_1^T A_1)^{-1} A_1^T x \\ b_2 &= [b_{2,1} \quad b_{2,2} \quad \cdots \quad b_{2,N_2}] = (A_2^T A_2)^{-1} A_2^T x \end{aligned} \tag{1.6}$$

5. The vectors $b_1$ and $b_2$ are concatenated together to form the feature vector.

– As EMD is not applied on each of the trials individually, the above algorithm can be implemented replacing EMD with its noise assisted version, ICEEMDAN.

### 3.2 Algorithm II

Since, for synchronous incongruent AV stimuli, increased global gamma-band coherence and decreased alpha and beta band coherence are observed around the duration of 300-600 ms after the presentation of the AV stimulus, in case of multisensory perception rather than unisensory perception as reported by Vinodh et al. in [8], it is evident that the features for classification between 2 classes will remain embedded in the spectral domain around the said temporal window. Based on this finding, the following algorithm is proposed.

1. A trial is divided into 3 sections: 1-300 ms, 301-600 ms and 601-900 ms.
2. Single channel EMD is applied separately on each of the three sections.
3. As in most of the trials, gamma band (30-50 Hz) is found in abundance in the 1st IMF and alpha (8-13 Hz) and beta (13-30 Hz) band frequencies are prevalent mostly in the $2^{nd}$ IMF, first 2 IMFs of each section are considered.
4. Apply HHT (Hilbert-Huang Transform) to the pair of IMFs obtained from each section and obtain the instantaneous frequencies.
   Now, for a particular trial, we have 6 vectors of instantaneous frequencies (*2 from each section*).
5. From the instantaneous frequency vectors of the $1^{st}$ IMFs, the region of occurrence of gamma band is found out and its corresponding average timing is noted. At places, where gamma is present, corresponding energy of the signal is also obtained.
6. In this same way, region of occurrence of alpha and beta bands were sought after in the $2^{nd}$ IMF and the time of occurrence and average energy of alpha and beta bands are also considered as features.

In this algorithm, EMD couldn't be replaced with any of the improved versions because on a particular trial, EMD is applied 3 times and applying improved versions of EMD will shoot up the time of computation sky high.

### 3.3 Algorithm III

It is already well established in literature that EMD generates some unwanted frequencies towards the end of a signal, which implies that there is a very high chance of occurrence of artefacts towards the edges. This problem may become prominent while performing EMD separately on three different sections after trisecting a particular trial (*as described in algorithm II*). To avoid this, and also to retain the focus on the power of different frequency bands at three different sections, EMD can be applied on the trials before segmenting it and the power of different frequencies can then be observed at various portions of the signal. Steps of the algorithm are described below:

1. Let N be the number of IMFs that are obtained on applying EMD to the $i^{th}$ trial.
2. Select the $1^{st}$ IMF and apply HHT (Hilbert Huang Transform) to obtain the instantaneous frequency vector.
3. Divide the frequency vector into three sections 1:300, 301:600 and 601:900.
4. Locate the presence of gamma band in all of the three sections.
5. Note the presence (average timing) of the gamma band and it's corresponding energy in all the three sections.
6. Select the $2^{nd}$ IMF and repeat steps 2 and 3.
7. Detect the alpha and beta bands from the three sections.

8. Location of occurrence (average timing) of the alpha and beta bands and their corresponding energies form the last dozen ($2 \times 2 \times 3 = 12$) of features of the feature vector.

In the above algorithm also, updated versions of EMD couldn't be applied on the trials because of the same reason of high computation time.

### 3.4 Algorithm IV

Although most of the information for multisensory speech integration of synchronous AV stimuli is found to accumulate within the top three frequency bands, no such unanimous findings were reported in the literature for stimuli with lags, both positive and negative. Keeping this in consideration, this algorithm is proposed so that it can be applied on EEG data for synchronous as well as asynchronous AV stimuli. Here also, only the first 2 IMFs are considered as many trials failed to generate more than a pair. The steps are documented below:

1. Apply EMD to a particular trial and select the $1^{st}$ pair of 'oscillatory modes'.
2. Concatenate the IMFs one after another to form the feature vector.

## 4 Results

Features, as generated from the proposed algorithms, are classified as depicted in **Figure 3** and labels from the top three electrodes exhibiting maximum accuracy are taken into consideration for prediction of the final label of the test data. Like most of the non-linear classifiers, random forest also requires fine tuning of a few parameters, like increase in the number of trees causes a rise in the accuracy (*though at the cost of computational complexity and time*) which saturates after reaching a certain number. In this case, the accuracy gets saturated when the number of trees in the RF reaches 200. A comparative study among the classification accuracies of different feature extraction techniques, presented in **Table 2**, reveals algorithm I to be more effective than others in all of the 3 cases of AV lags.

It turns out that, owing to the very high dimension of the feature vectors in algorithm IV, it took considerable amount of time in classifying the test data but the accuracy didn't improve appreciably. Whereas, algorithm I when modified with Improved Complete Ensemble EMD (ICEEMDAN) [13], the time required for classifying with this enhanced version of EMD was not much high with respect to the satisfactory classification accuracy it achieved. The parameters chosen for the training of the Random Forest classifiers in this study is as follows [14]:

- No. of samples for constructing a decision tree:- a randomly sampled number between 100 and 150.
- No. of trees in the forest:- 200.

The parameters for ICEEMDAN are:

- $\beta_0$ is kept constant at 0.2.
- Number of realizations of noise:- 800.
- Maximum number of iterations before terminating:- 200.

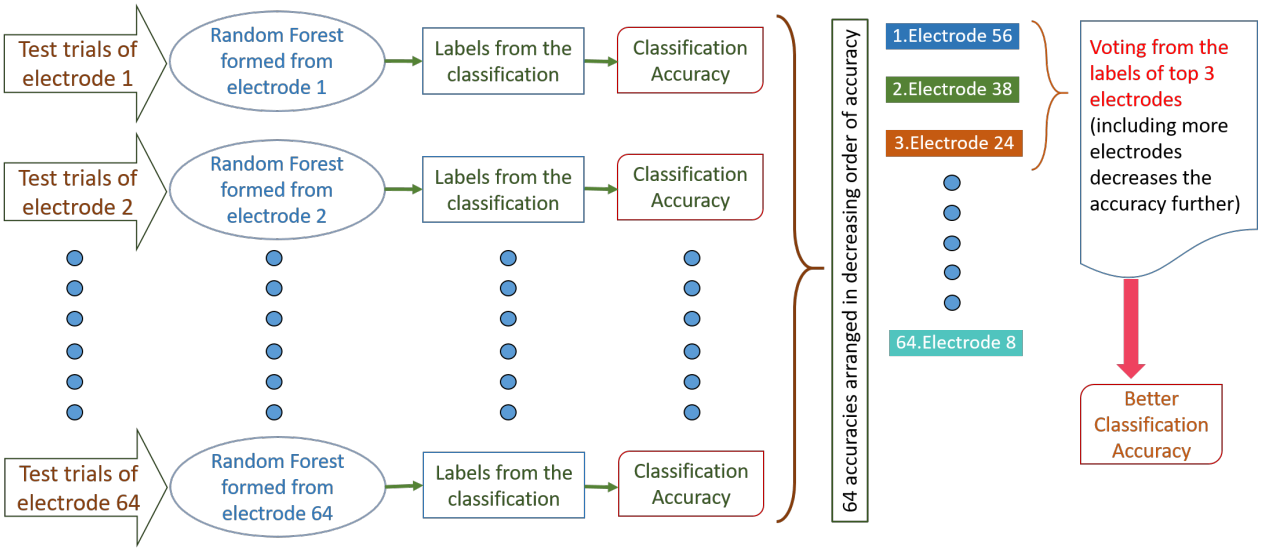



Fig. 3: Feature vectors of individual electrodes are classified through separate random forests and top 3 are considered for final labels.

### 4.1 Most informative frequency band

This study also aims at finding the frequency band which plays the most vital role in audio-visual integration. EMD, being an entirely data-driven approach, can't guarantee the presence of a particular frequency band in a specific IMF. So, for extracting distinct frequency bands from the combination of a wide range of frequencies, we applied the traditional filter bank technique.

EEG signal from a particular trial is passed through 5 band-pass filters yielding filtered signals corresponding to 5 different frequency bands namely: gamma (30 Hz to 50 Hz), beta (13 Hz - 30 Hz), alpha (8 Hz - 13 Hz), theta (4 Hz - 7 Hz) and delta (0.5 Hz - 4 Hz) bands. After that, each trial is represented by $64 \times 5 = 320$ vectors (5 frequency bands for each of the 64 electrodes). Classification is done on these vectors separately and arranged in the order of decreasing accuracy as shown in **Figure 4**. On completion of that, starting from the band of the highest accuracy, groups of 10 are formed with decreasing accuracy i.e. $10^{\text{th}}$ member of group 32 will have the least accuracy. In each of these 32 groups, the number of gamma, beta, alpha, theta and delta bands are noted separately. After that, a cumulative sum is performed for each of these frequency bands and plotted against group no.s.

| | **Single Electrode** | | | **Top 3 electrodes** | | |
|---|---|---|---|---|---|---|
| **AV lags** $\rightarrow$ | *-450 ms* | *0 ms* | *+450 ms* | *-450 ms* | *0 ms* | *+450 ms* |
| **Algorithm I (EMD)** | **62.8** | 63.66 | **60.14** | **66.13** | 75.85 | **64.59** |
| **Algorithm I (ICEEMDAN)** | 58.27 | **65.37** | 59.59 | 64.13 | **76.95** | 63.92 |
| **Algorithm II** | 59.87 | 62.56 | 57.30 | 62.27 | 75.37 | 61.35 |
| **Algorithm III** | 58.4 | **65.37** | 54.05 | 58.67 | 74.75 | 61.08 |
| **Algorithm IV** | 56.53 | 64.39 | 56.22 | 62.4 | 72.20 | 60.68 |

Table 2: Classification accuracies for different feature extraction techniques, on being classified with RF containing 200 trees each.

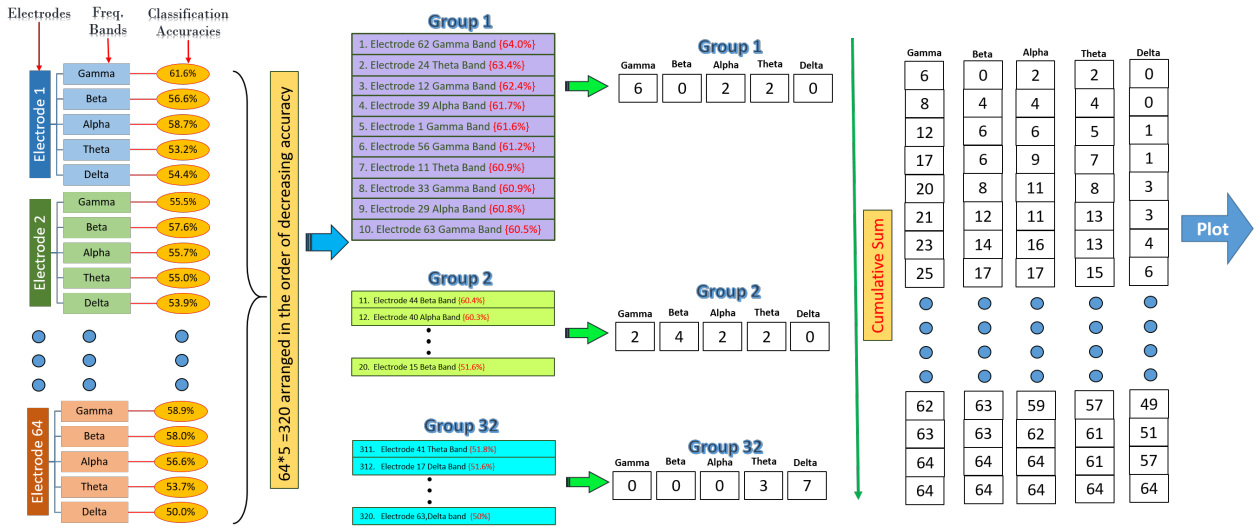
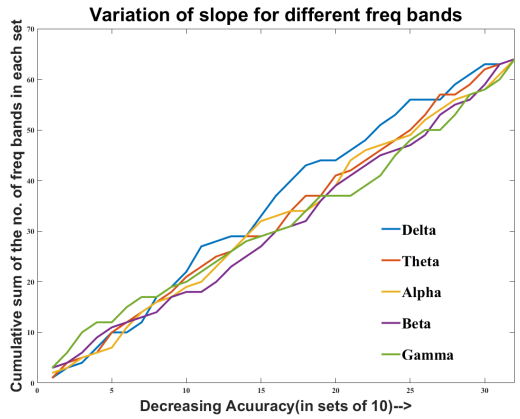
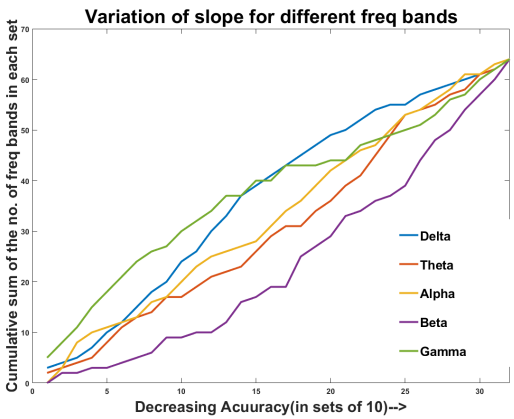


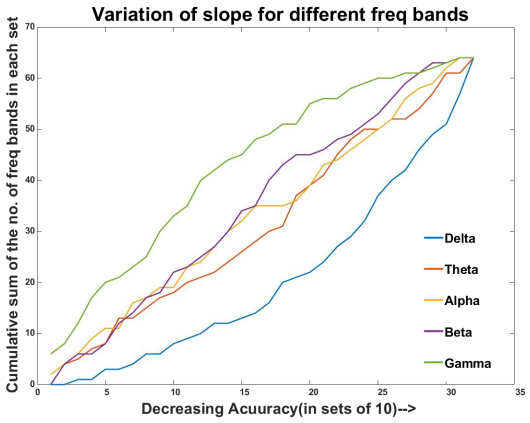

Fig. 4: The figure depicts the flow diagram for finding the relative activeness of different frequency bands. The data in the above figure is obtained when EEG signal for synchronous AV stimuli is classified with RF containing 200 trees each.



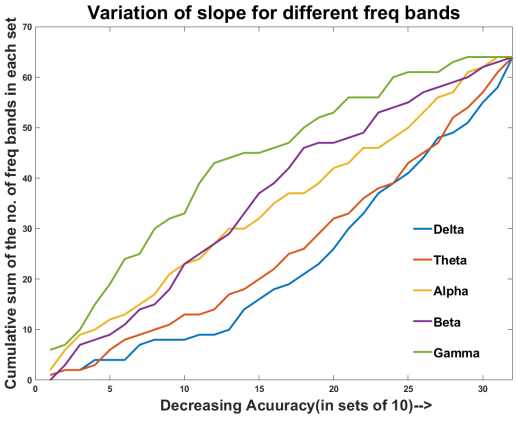

(a) For -450 ms lag AV stimuli




(b) For +450 ms lag AV stimuli




(c) For synchronous (0 ms lag) AV stimuli




(d) 300 ms - 600 ms of synchronous AV stimuli

Fig. 5: Comparison of participation from different frequency bands at different AV stimuli.

It is evident from **Figure 5c** that, for synchronous AV stimuli, gamma band holds the maximum information followed by alpha and beta bands. This pattern becomes more prominent when instead of all the 901 samples, only the data from 300 ms to 600 ms is considered (Refer **Figure 5d**). Both these findings contemplate the previous findings of [8]. The literature also suggested that no frequency bands were found to be exceptionally active with respect to others during the presentation of asynchronous AV stimuli which is also supported by **Figure 5a** and **5b**.

## 5 Conclusion

Obtaining classification accuracy of around 60% and 70% from a single channel and three channel EEG data respectively will lead to design a fast and practical way of classifying EEG signals with McGurk paradigm. As most of the results of the classification obtained in this study resonate with the existing literature, these set of electrodes can reliably be used in developing real time BCI systems. Similarly, this approach can also be extended to other paradigms in order to reduce the amount of data needed to reach high precision as well as facilitate the set up with a less bulky hardware.

## Acknowledgement

This research was funded by NBRC core and the grants Ramalingaswami fellowship (BT/RLF/Re-entry/31/2011) and Innovative Young Bio-technologist Award (IYBA) (BT/07/IYBA/2013) from the Department of Biotechnology (DBT), Ministry of Science  Technology, Government of India to Arpan Banerjee. Dipanjan Roy is supported by the Ramalingaswami fellowship (BT/RLF/Re-entry/07/2014) and DST extramural grant (SR/CSRI/21/2016). Neural Engineering Lab, IIT Guwahati is supported by Startup Grant, IIT Guwahati and NECBH grant sponsored by DBT.

## References

1. H. McGurk and J. MacDonald, “Hearing lips and seeing voices,”Nature, vol. 264, no.5588, p. 746, 1976.
2. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 454, no. 1971, pp. 903–995, 1998.
3. M. Molla, K. Islam, M. S. Rahman, A. Sumi, and P. Banik, “Empirical mode decomposition analysis of climate changes with special reference to rainfall data,” Discrete dynamics in Nature and Society, vol. 2006, 2006.
4. C.-H. Loh, T.-C. Wu, and N. E. Huang, “Application of the empirical mode decomposition-hilbert spectrum method to identify near-fault ground-motion characteristics and structural responses,” Bulletin of the seismological Society of America, vol. 91, no. 5, pp. 1339–1357, 2001.
5. D. Looney and D. P. Mandic, “Multiscale image fusion using complex extensions of emd,” IEEE Transactions on Signal Processing, vol. 57, no. 4, pp. 1626–1630, 2009.
6. A. D. Veltcheva, “Wave and group transformation by a hilbert spectrum,” Coastal Engineering Journal, vol. 44, no. 4, pp. 283–300, 2002.
7. A. Arasteh, M. H. Moradi, and A. Janghorbani, “A novel method based on empirical mode decomposition for p300-based detection of deception,” IEEE Transactions on Information Forensics and Security, vol. 11, no. 11, pp. 2584–2593, 2016.

8. G. V. Kumar, T. Halder, A. K. Jaiswal, A. Mukherjee, D. Roy, and A. Banerjee, “Large scale functional brain networks underlying temporal integration of audiovisual speech perception: An eeg study,” Frontiers in psychology, vol. 7, p. 1558, 2016.
9. J. Hocking and C. J. Price, “The role of the posterior superior temporal sulcus in audiovisual processing,” Cerebral Cortex, vol. 18, no. 10, pp. 2439–2449, 2008.
10. V. G. Kumar, S. Dutta, S. Talwar, D. Roy, and A. Banerjee, “Neurodynamic ex-planation of inter-individual and inter-trial variability in cross-modal perception,”bioRxiv, p. 286609, 2018.
11. M. Teplanet al., “Fundamentals of eeg measurement,”Measurement science review,vol. 2, no. 2, pp. 1–11, 2002.
12. F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of clas-sification algorithms for eeg-based brain–computer interfaces,”Journal of neural en-gineering, vol. 4, no. 2, p. R1, 2007
13. M. A. Colominas, G. Schlotthauer, and M. E. Torres, “Improved complete ensembleemd: A suitable tool for biomedical signal processing,”Biomedical Signal Processingand Control, vol. 14, pp. 19–29, 2014.
14. Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5-32.