International Journal of Innovation Engineering and Science Research
(ISSN: 2581-4591)

www.ijiesr.com

A Survey of Eeg Signals Prepossessing and Classification for Imagined Speech Application

Areej H. Al-Anbary
College of Information Engineering
Al-Nahrain University
Baghdad, Iraq

Prof. Dr. Salih Al-Qaraawi
Department of Control and Systems
University of Technology
Baghdad, Iraq

ABSTRACT

Recent studies in neuroscience, rehabilitation, and machine learning have concentrate on the Electroencephalography (EEG) Brain Computer Interfaces (BCI) as an important domain of research. So, the main purpose of the BCI is to restore communication in the severely paralyzed. Speech is mostly the normal way of communication for individuals; however, in circumstances where audio speech is not ready since the disability or adverse environmental condition, people may go about alternative ways such as augmented speech, that is, audio speech substituted by other forms, such as audiovisual speech, or Cued Speech. Accordingly, BCI for speech communication has been required in several non-medical domains. Selection of the processing technique of the EEG signals at each processing stage has a significant role in determining the success operation of BCI systems. In this survey, we outline a review for ten years ago of the most related speech techniques for pre-processing, feature extraction and classification developed to analyze the EEG signals in order to draw a guide line for researchers to help disabled (speechless) people to use some electronic devices so that such people can communicate with other people like normal persons to reduce their suffer during their daily life.

Keywords—*BCI, EEG, EMG, Imagined Speech, SSI, ICA, SVM, HMM.*

I. Introduction

Brain Computer Interface (BCI) is a system that connects human brain signals with appliances or devices without requiring of any physical contact, it has been seen as a new way for communication, where the brain activity has been used as a reflected form by electric, magnetic or hemodynamic brain signals to manage external system such as computers, wheelchairs, switches, or neuro prosthetic extensions[1] [2] [3]. BCIs are very useful tools for paralysis persons so there are two types of BCI systems, invasive or non-invasive BCI depending on the measurement method of the brain activity within BCI. If the place of sensors used for the measurement are placed inside the brain, i.e., under the skull, it is an invasive BCI. While, when the sensors are put on the scalp, it is a non-invasive BCI, they avoid injury risks and associated ethical concerns [4]. The processing stages of non-invasive BCI system are: data acquisition, data pre-processing, feature extraction, classification, device controller and feedback[5]. In invasive BCIs, electrodes or a multiunit electrode array will be placed directly inside the cortex to register electrical potentials for subsequent analysis of the electrocorticogram (ECoG). The resultant brain signals have a high signal-to noise ratio, requiring little user training, and are suitable for the rehabilitation motor functions in disabled patients, while the noninvasive BCIs, have deferent kinds of techniques for brain signals imaging, such as, Magnetoencephalography (MEG) electroencephalography (EEG), Electromyographic signals EMG[6] [7].

II. EEG Applications

Nowadays, EEG have been applied in different fields such as, monitoring alertness, coma, epilepsy, cognitive engagement and brain death, determining the damaged parts after head injury, stroke, cancer, physiology examining and sleep disorder [8]. Person identification[7], and controlling silent speech interfaces[9].

III. RELATED WORKS

In imagined speech recognition field, many researchers had worked on EEG preprocessing, analyzing and classifying. The most related works are reviewed then classified according to the technology used in acquiring the brain signals to three fields: researches with ECoG, with EMG, and with EEG.

*A. **Electrocorticogram ECoG:***

- F. Guenther and J. Brumberg [10] reporting two studies including BMI. The aim of those BMIs was to supply close immediate sound input from a discourse synthesizer to the BMI client. In one study, an intracranial electrode was used to record the neural signal by implanting them in left part of the brain, region concerns with speech, of a patient suffering from paralysis. Those signals were wirelessly transmitted over the scalp and to drive a formant synthesizer, enabling the client to create vowels. The second one, is a pilot study, a healthy volunteer had the ability to drive the formant synthesizer with envisioned developments distinguished utilizing electroencephalography. The outcomes showed the possibility of neural prostheses that can possibly give synthetic speech (near-conversational) for speechless persons.
- C. Herffet. al. [11] a synthesized speech from ECoG activity at the temporal regions at real time had been done. The spectrogram of the audio magnitude, had reconstructed from the neural activity, then the audio waveform from them was built. There was a considerable correlation between the base signal and the reconstructed signal. While audible form of spoken speech was used in the modeling, it considered as first stage to synthesize speech from thoughts.

- G. Anumanchipalliet. al. [12], showed the possibility of producing artificial speech from neural signals of the brain. ECoG signals had been recorded for five volunteers, who submitted for monitoring the intracranial for treatment of epilepsy. Cortical signals had been decoded by recurrent neural network with a clear impersonation for the articulatory dynamics to obtain audible speech synthesize output.

B. ***Electromyographic signals EMG:***

- M. Cleret. al. [13] recorded the facial muscles by using surface electromyography (sEMG) to develop their system to dominate phonemic link and voice synthesizer then tested the system in healthy individuals. They computed the mean rates of transferring the information for a selection of phonemes (ITRs), which was 59.5 bits/min. For the orthographic systems, they also computed ITRs depending on the number of letters needed for spelling the selected word, and reached the results of having mean ITR value of 70.1. So, a comparison for the consequences was made to apply their system on more individuals.

1.Trans cranial Magnetic Stimulation (TMS)

A.Ausilioet. al.[14], TMS had been used on motor area of tongue and lips by applying event-related double-pulse TMS on them. The suggested data showed that an important role may played by the motor system in noisy, surroundings, for speech signal recognition.

2. Electro-Magnetic Articulography (EMA)

- *P*. Heracleous et al. [15], introduced a communication using augmented speech based on EMA. Movements of jaw, tongue, and lips, were traced by EMA and were considered as features to build HMMs. The possibility of recognizing speech (without any audio information) had been examined by conducting the experiences of articulation automatic phoneme discrimination. Outcomes confirmed that phonetic features describing articulation are as discriminating as those characterizing acoustics (except for voicing). Experiments were described and conducted in noisy surroundings using EMA parameters and fused audio. Their results showed that, when EMA parameters were combined with noisy audio speech, the rate of discrimination was better than when applying just a noisy speech.
- M. Wand and T. Schultz [16] used Surface Electromyography (EMG) as a basis for Silent Speech Interface(SSI), the electric activity produced from the articulatory muscles was picked up from the face of user, by the electrodes to decode underlying speech, so the speech would be distinguished even when there were no sound was produced or heard. They used unsupervised session adaptation where a system was first trained with data set from different recorded session and then it was adapted with the required recorded data. They got a great level of accuracy improvements so that their technology may be used in future applications in real-life of SSI.

- Y. Ji et al. [17], updating the Silent Speech Interface (SSI) by using strategy of Deep Learning. A Word Error Rate had been minimized from 17.4% to 6.4%, also the data dimensionality had been reduced by using auto-encoder features. The module applied to two distinctive languages.

3. Imagined speech using EEG:

- J. Brumberg and F. Guenther [18], reviewed many methods for rehabilitation of communication by BCI for persons having severe cases of paralysis, also the dissimilarity between spelling devices and speech prosthesis or direct speech prediction .
- B. Denby et al. [19], condensed the development of the silent speech interface(SSI) from the domains of; automatic speech processing, speech production, speech pathology research, and telecommunications privacy issues. This work followed by the description of experimental systems based on seven diverse kinds of technologies. Pros and cons had been presented for each method.
- A. Riaz et al. [20] considered the state of envisioned and mouthed non-discernible speech, recorded with EEG terminals. They broke down various feature extraction strategies, for example, “Mel Frequency Cepstral Coefficients” (MFCCs), log fluctuation Auto Regressive (AR) coefficients. a pairwise arrangement of vowels was made by utilizing three diverse grouping models dependent on “Support Vector Machine” (SVM), Hidden Markov Models (HMM) and KNN classifier. The proposed procedure was applied on four unique informational indexes with some preprocessing systems, such as, “Common Spatial Pattern” (CSP) separating. The objective of this investigation was to play out a bury examination of various order models and related highlights for pairwise vowel symbolism.
- E. González-Castañeda et al. [21], utilized a strategy of Auditory display, sonification, on EEG signals to get better classification level for EEG signals for imagined speech, was utilized, which enables the describing of EEG signal as a sound sign. They compared the results of sonication processing EEG signals, then observed an improvement in the normal precision rates for signals, it had risen from 48.1% to 55.88%, so the characterization rates improved somewhat.
- K. Mohanchandra and S. Saha [22], concentrated in their work on specking in subvocalized manner, and that was the first trail in using sub-vocal words in EEG with imagined speech. The EEG signals then processed, to synthesize speech from them, with feedback returned to the user to confirm the results. They based on the assumption that, if the speech is undisguised or ulterior it will produce in the brain. The results showed the prediction possibility of the imagined speech. A pairwise correlation was used to minimize the data size and a multiclass SVM was used in classification process of EEG for five words obtained from electrodes.
- K. Brigham and B. Kumar [23], assessed the possibility of individuals’ identification using EEG signals during imagined speech (imagining syllables, /ba/ or /ku/). Noise and artifact effect

reduction was done by preprocessing EEG; feature extraction process was done by Autoregressive (AR) coefficients from each channel, and classified with linear SVM. The accuracy of identification was 99.76 % which reflected the possibility of utilizing envisioned discourse EEG information for biometric discrimination because of its solid variety among subjects.

- T. Schultz et al. [24], published a paper that gave an outline of the different ways, inquire about methodologies, and targets for using brain signal for communication via speech.
- P. Kumar et al. [25], proposed a discrimination of envisioned speech from EEG signals. A random forest algorithm was used in features classification coarse level, to classify them either non-text or text classes, then recognition of a finer-level envisioned speech from those classes had been applied. They got an accuracy of recognition about 85.20 at coarse classification, while the accuracy for fine level classification was 67.03%.
- J. S. Brumberg et al. [26], performed a research on using EEG in controlling the synthesize of speech, for the sounds of vowels(/i/, /A/, and /u/). They used three kinds of feedback to the user splitting them to three groups, such that, feedback of unimodal auditory for the speech synthesize, feedback of unimodal visual for the formant frequencies and the other feedback of multimodal. The results showed that the feedback audio-visual type improved the accuracy of the performance.

IV. DISCUSION

After exploring the most related BCI technologies used in acquiring, processing and classifying of the brain signals for inferring the imagined speech; so the following question has to be answered, "What practical techniques can be used in imagined speech to better detection of what is exactly the patient needs to say?". Table (1) gives the appropriate answer, it identifies parameters, degree of complexity, flexibility for some reviewed techniques.

Table (1) Summary of Some Related Researches: Parameters, Complexity and Results

Ref. No..	Parameters	Acquiring Method of brain activity	Classification Algorithm	Results
11.	The extraction of broadband gamma for hearable speech was done by many preprocessing steps like, filtering and down sampling	ECoG	A linear model was applied; since this study was a pilot one.	The study represented the first step for synthesizing imagined speech from neural brain signals
13.	Using the face muscles activities to control the selection of a phonemic interface and voice synthesizer	EMG	Calculating root mean square (RMS) for sEMG signals from eachelectrode every 100 ms and comparingwith the thresholds.	Producing "mean information transfer rates (ITRs) "of 70.1. bits/min i.e., the number of

Ref. No..	Parameters	Acquiring Method of brain activity	Classification Algorithm	Results
				selected phonemes per minute
15.	Movements of jaw, tongue, and lips, were traced by EMA and were considered as features to build HMMs.	EMA	Using of Hidden Markov models (HMMs)	Combining EMA parameters with noisy audio speech, enhanced the rate of discrimination.
16.	Offline Silent Speech Interface (SSI) model.	EMG	Unsupervised neural network for training on the recorded signals.	Good accuracy for testing phase.
17.	Updating Silent Speech Interface (SSI) model	EMA	Deep learning	Reducing Word Error Rate of by 6.4%
20.	Classification for vowels using pairwise by selecting three models of classification based on Hidden Markov Models (HMM), Support Vector Machine (SVM), and k-nn classifier.	EEG	Auto Regressive (AR) coefficients, Mel Frequency Cepstral Coefficients (MFCCs) and log variance	Classification of five vowel sounds
22.	BCI for imagining subvocalized words	EEG	Using of multiclass SVM for the extracted features from scalp electrodes.	Five subvocalized words were classified.
23.	Using speech imagining, for two syllables, /ba/ and /ku/, at different rhythms, to identify persons	EEG	Applying linear model of SVM classifier.	The results showed an accuracy of 99.76% in identification.
25.	Envisioned speech	EEG	Random forest classifier.	Accuracy of recognition was 85.20 %.
26.	Generating formants for the vowels"/i/ /u/ and /A/", by exploiting EEG of MI activity to control formant vector or synthesized in real-time for immediate auditory feedback.	EEG	Using "Hilbert transform" for training and "Kalman filter" for decoding.	Enhancing the performance by using meaningful multimodal feedback.

V. CONCLUSION AND FUTURE GUIDE LINE

In this paper the most related researches with the field of using BCI in inferring the imagined speech has been reviewed. Many methodologies had been used in extracting brain neural activities like; ECoG, EMG, TMS, EMA, and EEG. The researches' results showed that, ECoG had a reasonable accuracy in discriminating the envisioned speech because of its high SNR. Most of the studies worked on using EEG and EMG for recognizing imagined vowels or syllabus of speech, while very few of them worked on using EEG, EMG for recognizing single words, such as (Yes, No, Left, Right). As a proposal in the next step in this field, we will use EEG signals in designing a smart system as an assist device for paralysis or speechless persons. Fig. 1, shows the processing stages for the proposal, beginning with acquiring EEG signal, second preprocessing them (such as, filtering for the unwanted signals, extracting the most relevant features), third classifying the processed EEG signals are performed to their classes, fourth building a neural network model and training it with deep learning on the required classes, finally processing the more accurate event on a specific program to produce the required speech. The reason of preferring EEG signals for imagined speech research field because of their ease of use, low cost of set-up, good temporal resolution, and portability.

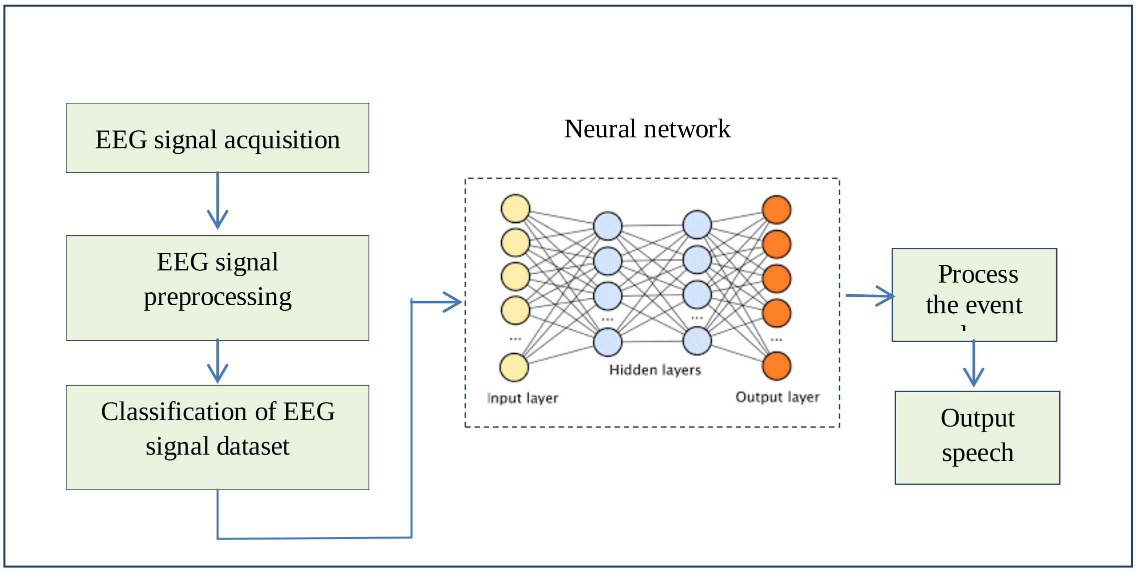

Fig. 1: The processing stages for the proposal

REFERENCES

[1.] G. Schalk, D. J. Mcfarland, T. Hinterberger, N. Birbaumer, J. R. Wolpaw, and A. B. I. B. C. I. Technology, "BCI2000 : A General-Purpose Brain-Computer Interface (BCI) System," vol. 51, no. 6, pp. 1034–1043, 2004.

[2.] R. Arulmozhiyal and K. Baskaran, "Implementation of a Fuzzy PI Controller for Speed Control of Induction Motors Using FPGA," Journal of Power Electronics, vol. 10, pp. 65-71, 2010.

[3.] F. Lotte et al., "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," J. Neural Eng., vol. 15, no. 3, p. aab2f2, 2018.

[4.] J. Freitas, A. Teixeira, M. S. Dias, and S. Silva, An Introduction to Silent Speech Interfaces. 2016.

[5.] E. R. Miranda, Guide to Brain- Computer Music Interfacing. .

[6.] B. Interfaces, "A Tutorial on EEG Signal Processing Techniques for Mental State Recognition in Brain-Computer Interfaces," Guid. to Brain-Computer Music Interfacing, no. August, 2014.

[7.] R. Benca, "Fundamentals of EEG Technology, Volume 2: Clinical Correlates," J. Neuropsychiatry Clin. Neurosci., vol. 1, no. 3, pp. 335–335, 1989.

[8.] P. Nguyen, D. Tran, X. Huang, and D. Sharma, "A proposed feature extraction method for EEG-based person identification," Proc. 2012 Int. Conf. Artif.Intell.ICAI 2012, vol. 2, pp. 826–831, 2012.

[9.] S. U. C. M. Ehrotra, Introduction To Eeg- and Emotion Recognition. 2018.

[10.] G. N. Jayabhavani and N. R. Raajan, "Silent speech generation using brain machine interface," 2013 IEEE Int. Conf. Emerg. Trends Comput.Commun.Nanotechnology, ICE-CCN 2013, no.Iceccn, pp. 630–633, 2013.

[11.] F. H. Guenther and J. S. Brumberg, "Brain-machine interfaces for real-time speech synthesis," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, pp. 5360–5363, 2011.

[12.] C. Herff, G. Johnson, L. Diener, J. Shih, D. Krusienski, and T. Schultz, "Towards direct speech synthesis from ECoG: A pilot study," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2016-Octob, pp. 1540–1543, 2016.

[13.] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," Nature, vol. 568, no. 7753, pp. 493–498, 2019.

[14.] M. J. Cler, A. Nieto-Castanon, F. H. Guenther, and C. E. Stepp, "Surface electromyographiccontrol of speech synthesis," 2014 36th Annu.Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC 2014, pp. 5848–5851, 2014.

[15.] A. D'Ausilio, I. Bufalari, P. Salmas, and L. Fadiga, "The role of the motor system in discriminating normal and degraded speech sounds," Cortex, vol. 48, no. 7, pp. 882–887, 2012.

[16.] P. Heracleous, P. Badin, G. Bailly, and N. Hagita, "A pilot study on augmented speech communication based on Electro-Magnetic Articulography," Pattern Recognit. Lett., vol. 32, no. 8, pp. 1119–1125, 2011.

[17.] M. Wand and T. Schultz, "Towards real-life application of EMG-based speech recognition by using unsupervised adaptation," Proc. Annu.Conf. Int. Speech Commun.Assoc. INTERSPEECH, no.September, pp. 1189–1193, 2014.

[18.] Y. Ji, L. Liu, H. Wang, Z. Liu, Z. Niu, and B. Denby, "Updating the Silent Speech Challenge benchmark with deep learning," Speech Commun., vol. 98, pp. 42–50, 2018.

[19.] J. S. Brumberg and F. H. Guenther, "Development of speech prostheses: Current status and recent advances," Expert Rev. Med. Devices, vol. 7, no. 5, pp. 667–679, 2010.

[20.] B. Denby, T. Schultz, K. Honda, T. Hueber, J. M. Gilbert, and J. S. Brumberg, "Silent speech interfaces," Speech Commun., vol. 52, no. 4, pp. 270–287, 2010.

[21.] A. Riaz, S. Akhtar, S. Iftikhar, A. A. Khan, and A. Salman, "Inter comparison of classification techniques for vowel speech imagery using EEG sensors," 2014 2nd Int. Conf. Syst. Informatics, ICSAI 2014, no. Icsai, pp. 712–717, 2015.

[22.] E. F. González-Castañeda, A. A. Torres-García, A. Rosales-Pérez, C. A. Reyes-García, and L. Villaseñor-Pineda, "EEG Sonification for Classifying Unspoken Words," Model. Anal.vocalEmiss. Biomed.Appl. 9 th Int. Work., no. July 2016, pp. 63–66, 2015.

[23.] K. Mohanchandra and S. Saha, "A Communication Paradigm Using Subvocalized Speech: Translating Brain Signals into Speech," Augment. Hum. Res., vol. 1, no. 1, pp. 1–14, 2016.

[24.] K. Brigham and B. V. K. V. Kumar, "Subject identification from Electroencephalogram (EEG) signals during imagined speech," IEEE 4th Int. Conf. Biometrics Theory, Appl. Syst. BTAS 2010, 2010.

[25.] T. Schultz, M. Wand, T. Hueber, D. J. Krusienski, C. Herff, and J. S. Brumberg, "Biosignal-Based Spoken Communication: A Survey," IEEE/ACM Trans. Audio Speech Lang. Process., vol. 25, no. 12, pp. 2257–2271, 2017.

[26.] P. Kumar, R. Saini, P. P. Roy, P. K. Sahu, and D. P. Dogra, "Envisioned speech recognition using EEG sensors," Pers. Ubiquitous Comput., vol. 22, no. 1, pp. 185–199, 2018.

[27.] J. S. Brumberg, K. M. Pitt, and J. D. Burnison, "A Noninvasive Brain-Computer Interface for Real-Time Speech Synthesis: The Importance of Multimodal Feedback," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 26, no. 4, pp. 874–881, 2018.[33]

[28.] A. B. Usman, J. Gutierrez, and A. B. Bichi, 'A Neighbourhood-Based Trust Protocol for Secure Collaborative Routing in WirelessGutierrez Mobile D2D HetNets', Int. J. Comput. Sci. Inf. Secur., vol. 16, no. 4, pp. 1–9, 2018.

[29.] D. Liu, Z. Li, Z. Hu, and Y. Li, 'Distributed Reinforcement Learning for Quality-of-Service Routing in Wireless Device-to-device Networks', 2018 IEEE/CIC Int. Conf. Commun. China, ICCC Work. 2018, pp. 282–286, 2019.

[30.] S. Alwan, I. Fajjari, and N. Aitsaadi, 'Joint Routing and Wireless Resource Allocation in Multihop LTE-D2D Communications', Proc. - Conf. Local Comput. Networks, LCN, vol. 2018-Octob, no. 1, pp. 167–174, 2019.

[31.] N. Ben Halima and H. Boujemâa, 'Optimal routing and one hop routing for D2D communications in the presence of mutual interference', Springer Telecommun. Syst., vol. 71, no. 1, pp. 55–64, 2019.

[32.] V. Tilwari, K. Dimyati, M. Hindia, A. Fattouh, and I. Amiri, 'Mobility, Residual Energy, and Link Quality Aware Multipath Routing in MANETs with Q-learning Algorithm', Appl. Sci., vol. 9, no. 8, p. 1582, 2019.