Feature Extraction and Selection Methods for Motor Imagery EEG Signals : A Review

Rijuta V. Wankar
Department of Electronics and Telecommunication
Sardar Patel Institute of Technology
University of Mumbai
Maharashtra, India
(email id: rujuta.w1992@gmail.com)

Payal Shah
Department of Electronics
Sardar Patel Institute of Technology
University of Mumbai
Maharashtra, India
(email id: payal_shah@spit.ac.in)

Rajendra Sutar
Department of Electronics
Sardar Patel Institute of Technology
University of Mumbai
Maharashtra, India
(email id: rajendra_sutar@spit.ac.in)

***Abstract* : Extraction and selection of electroencephalography (EEG) features is a pivotal task. The brain–computer interface (BCI) for motor imagery (MI) task is analysed with respect to the classification accuracies in following described papers. The paper gives a brief discussion on various feature extraction and selection techniques that has been studied for different motor imagery functions. The comparison table is made with respect to the features extraction methods, selection methods, EEG data used for analysis, number of electrodes for data acquisition, computation time and method implemented. Different techniques such as JayaNFCSSCGLH , LPSVD, sparse weighted CSP, IMF, CBN,SBCSP are discussed. Flowcharts for every method is discussed. The techniques determines the defining characteristic in the procedure that helps in producing better signal for analysing and differentiating brain signal at it utmost probability. Lastly the discussion is made as to which technique outperformed when motor imagery task is taken into consideration for the (BCI) brain-computer interfacing mechanism. To clarify better the classification accuracies of studied methods are compared.**

***Index Terms* : brain-computer interface, feature extraction, feature selection, motor-imagery tasks.**

I. Introduction

BRAIN-PC interface once in a while called a direct neural interface, is an immediate correspondence connect between a working human mind and the outside world. BMI utilizes cerebrum movement to charge, control, impel and convey with the world specifically through cerebrum coordination with fringe gadgets and frameworks. The field of BCI has developed for the most part toward neuroprosthetics applications that go for re establishing harmed hearing, sight and development. Fundamental guideline behind this interface is the bioelectrical action of nerves and muscles. Cerebrum is made out of a large number of neurons. At the point when the neuron fires, or actuates, there is a voltage change over the phone which creates motions on the surface of the mind. By observing and investigating these signs we can comprehend the working of cerebrum. Electroencephalography (EEG) is estimation of electrical movement delivered by cerebrum as recorded from cathodes set on the scalp. The EEG is recorded with cathodes, which are set on the scalp. Anodes are little plates, which direct power. They give the electrical contact between the skin and the EEG recording contraption. The flag preparing segment is separated into three classifications ; a)Multichannel Acquisition Systems: at this area enhancement, starting sifting of EEG flag furthermore, conceivable curio expulsion happens. b) Spike Detection : It will enable the BMI to transmit just the action potential waveforms and their individual entry times rather than the inadequate, crude flag completely, c) Signal Analysis: In this arrange, certain elements are extracted from the preprocessed and digitized EEG flag which are contribution to the classifier. Classifier perceive distinctive mental undertakings.

Numerous analysis methods have been proposed for the feature extraction and the classification of two or more MI classes such as CSP and its various extensions. The time-frequency domain analysis such as short time fourier transform (STFT) ,Discrete Wavelet transform, Various discriminating algorithm such as self regression model , neural networks gaussian process and Bayesian processes to increase the robustness and accuracy of the BCI. Table I describes various methods involved in the extraction process.

The paper speaks about the feature extraction techniques that gives the precise evaluation of the features taken into consideration, the extraction and the classification methods used, the dataset required, electrodes used and the accuracy related to each method respectively.

II. Feature Extraction methods

A. CSP approach

The common spatial pattern (CSP) is most common and effective approach in feature extraction of biomedical signals.

The persuance of these spatial filters depends upon the frequency band in which they are active. The flowchart as shown in fig 1 gives the insight of the method proposed in the [1]. Shang-Lin Wu et al. proposed the use of sub-band CSP (SB-CSP) method [2] since the use of conventional CSP makes it a time consuming process to search for optimum frequency range. The decisions for sub-band are obtained from LDA classifier. Since the decisions are derived from different classifier fusion algorithm is applied that attempts better uncertainty identification [1]. The fuzzy –fusion aaproach [3] acquiesce the ambiguous, imprecate and partial information from EEG to process using the concept of Sugeno integral [4] and Choquet integral [5]. The EEG signals acquired from EEG device [6] with the specification mentioned in the paper. The filter bank extract frequencies from 1 to 30Hz. W is the matrix that maximize the deviation of the filtered data. LDA [7] determines the mean vector and covariance matrices of independent classes and find the linear combination of features that maximize the difference among different classes. Later decisions from LDA classifier are fed to the fuzzy integral. Here the combination of classifier is done as it is believed to enhance the decision-making process as well as increase the robustness and reliability of BCI system. The LDA decision is formulated applying five base classifier ie, delta, theta, alpha , beta and all band classifier. The confidence level of the classifier is ameliorated by partial swarm optimization (PSO) algorithm [8]. In this method the a partical finds a better position compared to the previous position and when the new position is found it is replaced with the previous positon . This updating the data continues to conduct until the best position is attained labelled as global best position. The mechanism exhibits until the predefined iterations are realized. This procedure makes possible to access optimized output during training period.

B. Empirical mode function (EMD)

EMD is an time-frequency based method thus efficient for time frequency analysis of non-stationary signals. EEG signals are more prominently non-stationary signals thus EMD method is one of the efficacious technique for feature extraction [9]. EMD is a data- dependent method in which signal is disintegrated into number of intrinsic mode functions (IMF) that are oscillatory components. The mean frequency of IMFs are used for EEG classification [10]. The IMF is obtained by following two basic conditions are a) the total number of extremas and zero crossing should be equal to or vary by the maximum of one ; b) at any given instance, the standard value of envelop defined by local maxima and local minima, should be equal to zero [11]. Farhan Riaz et al. proposed the third order temporal moments and spectral features of the IMFs for feature extraction from EEG signals. The flowchart given in fig 3 shows the steps for calculation of IMFs[9]. the envelop of local maxima and minima is defined by $e_1(t)$ and $e_m(t)$. $c_1(t)$ is the obtained IMF. The analytical representation is obtained as it removes the DC offset from spectral component. $H\{c_m(t)\}$ is the Hilbert Transform of $c_m(t)$ is the math IMF of the signal x(t) [12]. The spectral analysis of signal is done using instantaneous frequencies (IF) but it is feasible only for mono-component signal [13]. Thus features are calculated by PSD in paper [9]. The spectral centroid , variation coefficient and the spectral skew are the parameter extracted from IMFs which form distinct group when supervised clustering is applied on the signals [14]. These features obtained are used for classification using support vector machine (SVM). SVM maximizes the separation margin in two different classes [15]. The training consist N labelled samples. In the decision function K(.) is the kernel function, F is input vector, a_i is the coefficient from training steps.

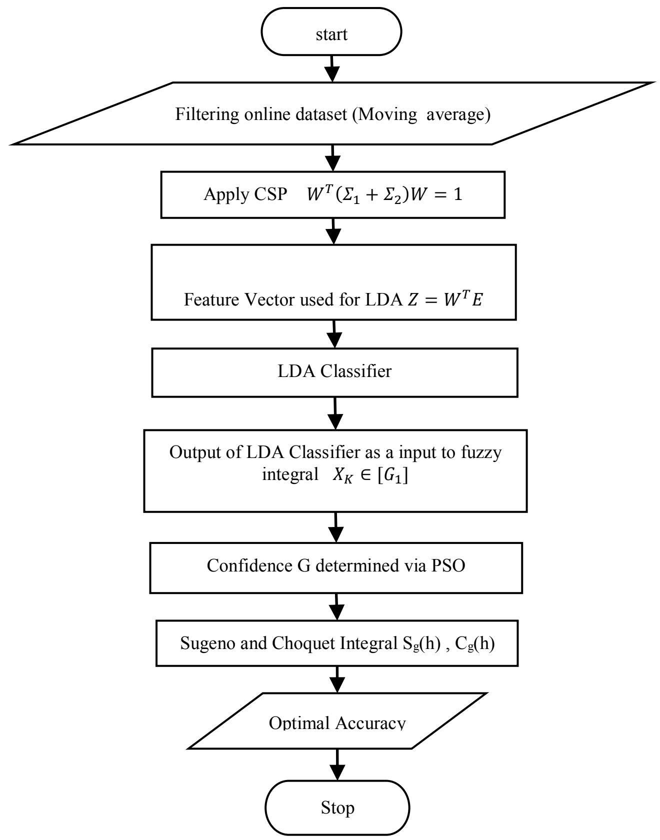

Fig 1: Flowchart showing CSP approach for feature extraction [1]

C. Bayesian network (CBN)

The Common Bayesian Network works on the baye's rule. This method is used to study the likeliness between the activation areas in the motor imagery tasks [16]. There are three steps for the framework of CBN ie. Bayesian Network (BN) construction, CBN construction feature discrimination. Lianghua He et al. has proposed a novel method of where gaussian mixture network GMM [17] is used to learn EEG data of every BN node. The flowchart regarding the novel

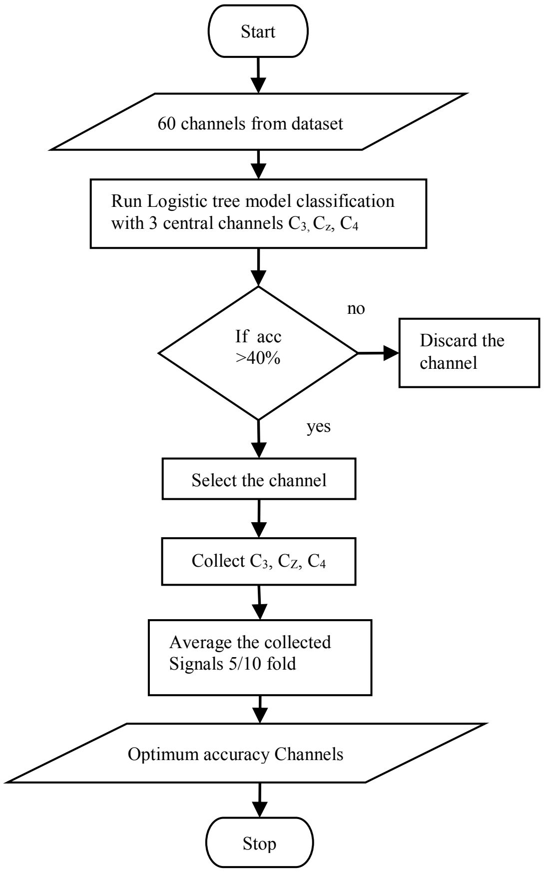

Fig 2a: Flowchart showing feature extraction method for transformed based feature extraction technique [22]

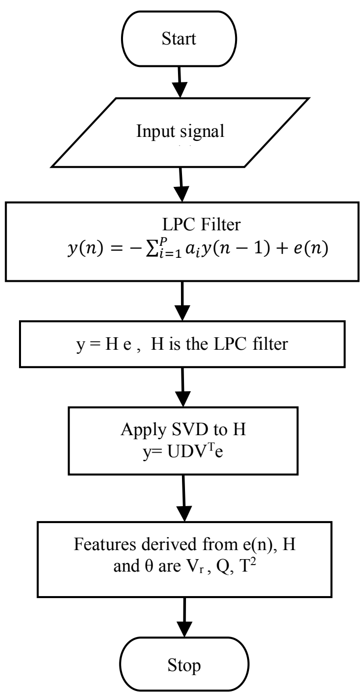

Fig 2b: Flowchart showing channel selection method for transformed based feature extraction technique [22]

method of CBN construction and its use in feature extraction is shown in fig 4. The signal is collected from the online data set as mentioned in the classification accuracy table below. The learning task in BN is separated into two subtask, structure learning and the parameter estimation. An EEG signals are obtained from different brain areas. Different brain areas leads to different cognition function, resulting into different cognition task. Thus it is concluded that different cognitive task results in different brain networks [18]. The BN network construction is partite into two steps ie. structure learning using GMM and conditional densities using GMM. The primary task is to fine the vector parameter θ which is the GMM model for one channel signal. Estimation-Maximization Algorithm is used to find θ [19]. In Conditional densities, two nodes are considered. Their condition probability is the main attribute to estimate the structure. for each node the GMM is learned , every GMM having covariance matrix diagonal and the relationship among different Gaussian component is independent. BIC score is used to evaluate the BN structure. The structure having minimum BIC score is used further for CBN construction. This scoring is done through greedy search algorithm [20]. The features extracted from BN has two properties: it has discriminating information for different kinds of motor imagery task; it has stable position to make calculation possible. Due to low SNR the learned structure changes from time to time thus CBN construction is required for the stable structure evaluation. It consist of two concepts Edge Common Rate and Node Variation Rate. The feature extracted in the most recognition task shows the common property for MI task. The experiments performed in [16] shows that edges played a major role in differentiating the channels. so it is concluded that edge itself is a BN structural feature. In Node variation rate, node describes the differentiating factor in every MI task. Signal collected from every EEG channel is considered to be node. This node discrimination is done by Fisher's rule. The node variation rate is given by

$$\boldsymbol{f = \sum_{i=1}^{N_i} \alpha_i} \tag{1}$$

Here the value of f shows that larger the f large is the difference between interclass and intraclass of MI task. The feature extraction is done using SVM method [21]

D. *Transformed based Feature extraction*

The aim of transformed based feature extraction method is to detect a compressed lower-dimensional representation where utmost data energy is packed in less number of uncorrelated coefficients. These method extracts effective feature by elimination irrelevant features while preserving the general performance of the technique and decreasing the computational complexity in the classification stage [22]. ie fewer spectral coefficients contains most of the signals energy.

Hamza Baali et al. has proposed this method, it performs the combination of feature extraction and channel selection.

Wrapper-based method is used in channel selection [23]. The author proposed the LP-SVD transform method for feature extraction. it is a two step process which includes estimation and computation of LPC filter. y(n) is the input signal given to the LPC filter. Linear Prediction is the time series analysis method used , the linear combination of the past samples is calculated and given as the current value of the signal ie.y(n). αi is the linear prediction coefficients, H is the impulse response matrix or the synthesis filter. Singular Value decomposition (SVD) is applied to the H and thus feature vector θ is obtained. The feature vector converts (decorrelates) the signal energy into few transform coefficients. The prediction variance (Vr) is estimated using α_i (LP coefficient), the square prediction error (SPE) or Q-statistics is another parameter calculated as the error between the vector **y** and its reduced version $\hat{\boldsymbol{y}}$ where **y** is LP-SVD transformed feature , the Hotelling's T^2 measures the sum of α_i bye their corresponding singular values σ_i. The features extracted from these three features overall helps in feature reduction. Channel selection is followed using wrapper based technique using Sequential Forward Selection (SFS) algorithm used for feature selection [24]. The classifier used is logistic model tree, the most correlated variable with the current reside is selected the simple regression coefficient is applied on it and added to current coefficient for that variable [25]. The maximum number of iterations depends upon the cross-validation ie five fold or ten fold. The flowchart regarding the method is mentioned in fig 2a and in fig 2b.

E. Jaya based ANFIS

The neuro-fuzzy classifier (NFC) have the capability of reasoning and learning ability. There adaptiveness, unsupervised learning and the population search capabilities makes improve the fuzzy rule into more appropriate selection process [26]. Jaya based k-means is applied to segment the feature set into two mutually exclusive cluster and impel the fuzzy rule [27]. K-means technique uses k clusters and N objects as input k < N, the fitness function is carried out to find and specify the position of ach particle in the cluster. Each particle is associated with its nearest cluster C_j and m_i is the center. Updation of the cluster takes place given by the equation described in [26]. The equation gives the best value ie. The minimum value of the particle as well as worst value of the particle so that the worst value can be discarded. The equation has the impulse to suppress the lowest distance particle by avoiding it and move the solution to the better position among the particles. The linguistic Hedge (LH) is the function that remodel the fuzzy set into another set by modifying the shape of the membership value of any feature is around one then, the feature is tending to give better results and are considered as the relevant feature [28]. Every node forms a signal converted into sample, the weights obtained are taken as the outputs and are normalized so that average weighted output is obtained. Speeding up scaled conjugate technique is applied (SSCG) in the training algorithm to adjust the parameters of NFC [29].The technique is described in fig 5.

III. Discussion

In [1] the application of PSO increased the accuracy results in both Sugeno integral and Choquet integral. The accuracy is stated in the classification table given in Table II. The experiment was performed by comparing with existing fusion technique like LDA and SVM, which demonstrates that the area under curve(AUC) was augmented and the standard deviation was depreciated implying better stability. The robotic arm used was six- axis robotic manipulator with three-fingered hand. The fuzzy fusion technique was also applied on online BCI data and the accuracy rate results out to be 86% approximately. The method was proposed to reduce the uncertainty of the data coming from multiple decision of multiple sources. The method is applicable to any EEG data acquired from healthy subject on the account of basic motor imagery tasks carried.

Paper [9] discussed the EMD technique used four different classifiers ie., 1-nearest neighbour(1NN), decision tree, artificial neural networks(ANN), and Support Vector Machine(SVM) and compared with feature extraction techniques: EMD, IF , BW and Wavelets. The data was grouped into five sets. All sets were further grouped forming healthy class, interical class and seizure class and the combination of these sets and class was categorized into cases which were then classified according to the feature extraction methods and classifiers. For each case the proposed method in [9] outperformed the other feature extraction methods in every class. The best performance obtained in combination with decision tree with 96% classification accuracy. The method discussed in the paper used the weighted frequencies from IMFs to identify the seizures classification. Thus is applicable on heathy and presurgical diagnosed subjects for the detection of the disease.

In Common Bayesian Network (CBN) [16] two sets of online dataset was used. On one set data was collected from three subjects and for other set nine subjects were used for data collection. Kappa coefficient evaluation was experimented [36] as it is considered to be robust evaluation compared to the simple percentage. The (CBN+SVM) method was compared with CSP, LDA and SVM methods. The novel method has shown 11 to 22% increase in accuracy rate varying between the range within every subject and the kappa coefficient has the accuracy of 66%. The technique can be applied on any online or real time EEG data. Since data acquisition is a lengthy process the signal quality can be affected over a period of time reducing the efficiency of the method, for this purpose shortest period data collection is

TABLE I. VARIOUS FEATURE EXTRACTION AND FEATURE SELECTION METHOD STUDIED

Author	Year	EEG fetures	Method/ Description	Electrodes	Feature Selection	Classifier
Shang-Li Wu, Yu-Ting Liu[1]	2016	Sensorimeter area (1-30Hz)	Spatial filtering	Dry electrodes	Heuristically using PSO	CSP and LDA
Farhan Riaz, Ali Hassan[9]	2016	Online database foe epilepsy or seizures	EMD	-	Linear kernel	SVM
Lianghua He, Die Hu, Meng Wan et.al [16]	2016	Online database BCI competition II, III	CBN	K_3, K_6,L_1 (C_3,C_z ,C_4)	-	SVM
Hamza Baali, Aida Khushidtabab[22]	2015	Primary motor cortex	LPSVD paired with DCT and AAR	C_3,C_z ,C_4	SFS wrapper	Logistic tree based model
Suraj, Rakesh kumar Sinha[27]	2017	μ(7-14Hz) β(17-26Hz)	JAYA based k-means using NFC	C_3,C_z ,C_4	Neuro Fuzzy Classifier using SCG and SSCG	NFC
Peter Horki, Daniela Kiobassa[31]	2015	ERS and EDS	3rd order Butterworth filter	AF_2,F_3,F_1,Fz,F_2,F_4,FC_2, FCz,FC_4,C_5,C_3,C_1,C_2,Cz, C_4,C_6,CP_3,CP_1,CP_2,CP_4,P_3 ,P_1,Pz,P_2,PO_2	Bonferroni’s corrected pair test	BLDA nested cross validation
Naoki Tomida ,Toshihisa Tanaka[34]	2015	7-30 Hz	Butterworth LPF	F_3,F_1,Fz,F_2,F_4,FC_2, FCz,FC_4,C_5,C_3,C_1,C_2,Cz, C_4,C_6,CP_3,CP_1,CP_2,CP_4,P_3 ,P_1,Pz,P_2	Sparsity aware method	CSP with variants of CSP
Qiang Wang, Ogla Sourina[32]	2013	PCA method 4 fold cross validation	GHFDS		SVMRBF kernel	SVMRBF
Pega Zarjam, Julian Epps[33]	2015	Wavelet based feature of best estimated coeff. FL,FR,OC channel	DWT RECT window T=5sec		Weighted energy, Weighted entropy,wighted SD	MLP ANN’s classifier
Shouyi Wang. JAcek Gwizdka[35]	2016	η-back test	DWT daubiechies wavelet	F_3,F_4,AF_3,AF_4	m-RMR	Proximal SVM

The transform based method [22] compared with the similar state-of-art extraction methods that are based on signal modelling and orthogonal transform. The techniques compared are discrete cosine transform (DCT) and adaptive autoregressive model (AAR). Three subjects L1b, K3b and K6b were taken from the online dataset and the test was performed. It shows that the LP-SVD method with four coefficients, AR model and error variance shows 25% better accuracy when compared to DCT and 6% better accuracy when compared to AAR model. The method was applied on online dataset and can also be applied on real time data making it an effective technique considering the extraction and selection processes.

The Jayabased ANFIS system [27] showed that the LH based feature extraction system increased the classification accuracy rate in the experiments. The SSCG training method improved the accuracy and computation time effectively. The classification was compared with SVM, radial basis SVM and LDA. Nine subjects were taken as a data for training the classifier. The performance of JayaNFCSSCGLH was better than all the other classifiers. Friedman’s Test is a non-parametric test which was also carried out to check the implication and counterfeit of the result. The method has been data considering the primary motor imagery task of imaging the left, right hand movements or left, right foot movement. Thus all the above methods are quie practical and equally applicable for all the available online datasets as wllw as real time data acquired from the EEG acquisition model.

IV. CONCLUSION

In this paper, we have discussed the brief study of various feature extraction method. The methods described are better than the conventional feature extraction and classification methods ie,. CSP, SVM, LDA, PSD. All the methods are feasible to implement in real-time EEG analysis for the motor imagery functions such as robotic arm control, controlling the computer task or wheelchair control. Most of the methods shows the spectral features domination except for the transform based method which is temporal based method. This shows that frequency domain method shows the better computation time as well as accuracy improvements. In Fuzzy integral technique the PSO are the predominate factors that results in better accuracy. EMD based signals have better flexibility over non-stationary signals which is the case of EEG signals and the spectral domain was considered as it shows good PSD discriminating powers in the classification. The two component GMM model and the learning BN structure proved to be a explicit option for the feature extraction compared to gaussian model in the CBN model. The LP-SVD based feature extraction in Transform based extraction method results into accuracy improvement. The Linguistic hedge based feature selection, use of NFCs that reduce the computation time and the SSCG training algorithm that further lessened the computational complexity without compromising accuracy are the features that makes the Jayabased ANFIS system a better option for the feature extraction technique in EEG classification for the motor imagery based BCI system. Thus all the five method showed their ethnicity in the methods yet providing good accuracy in the extraction and therein real time BCI control.

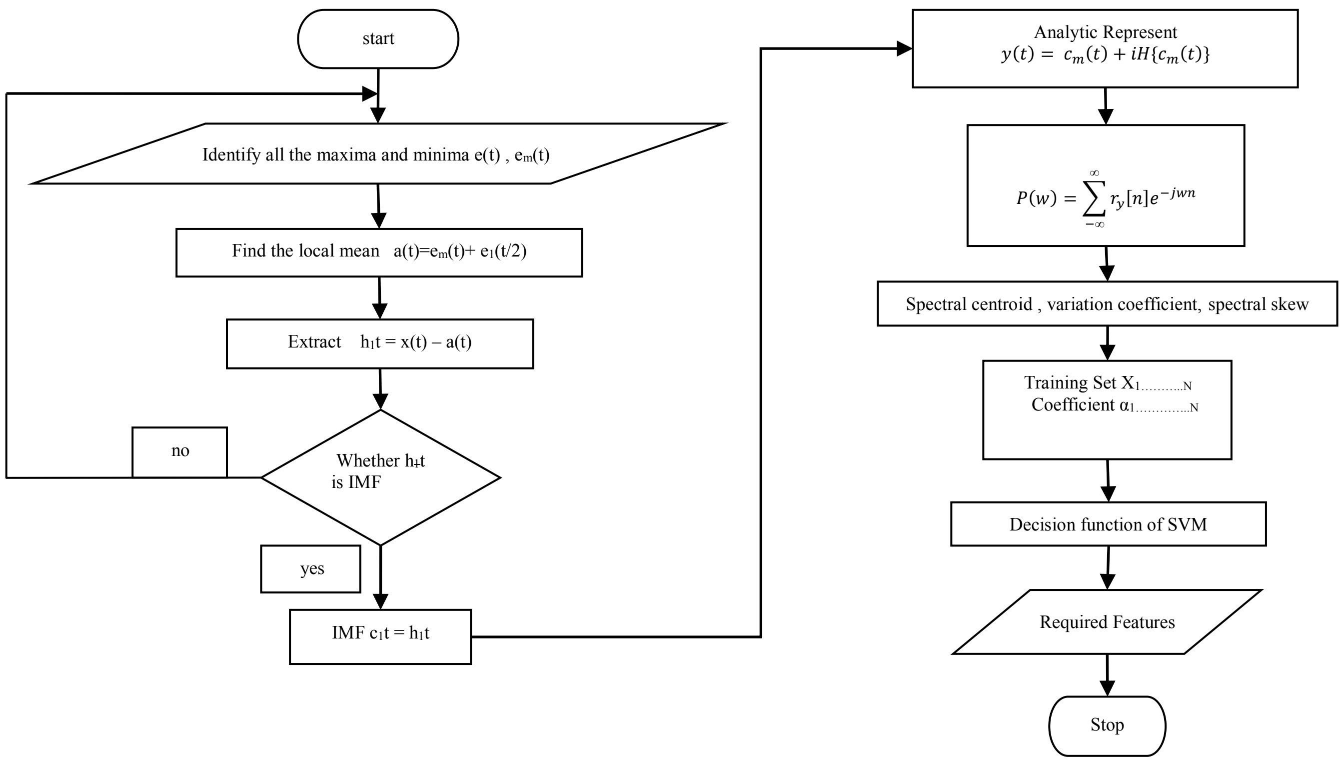

Fig 3: Flowchart showing Feature extraction technique for Empirical Mode Function [9]

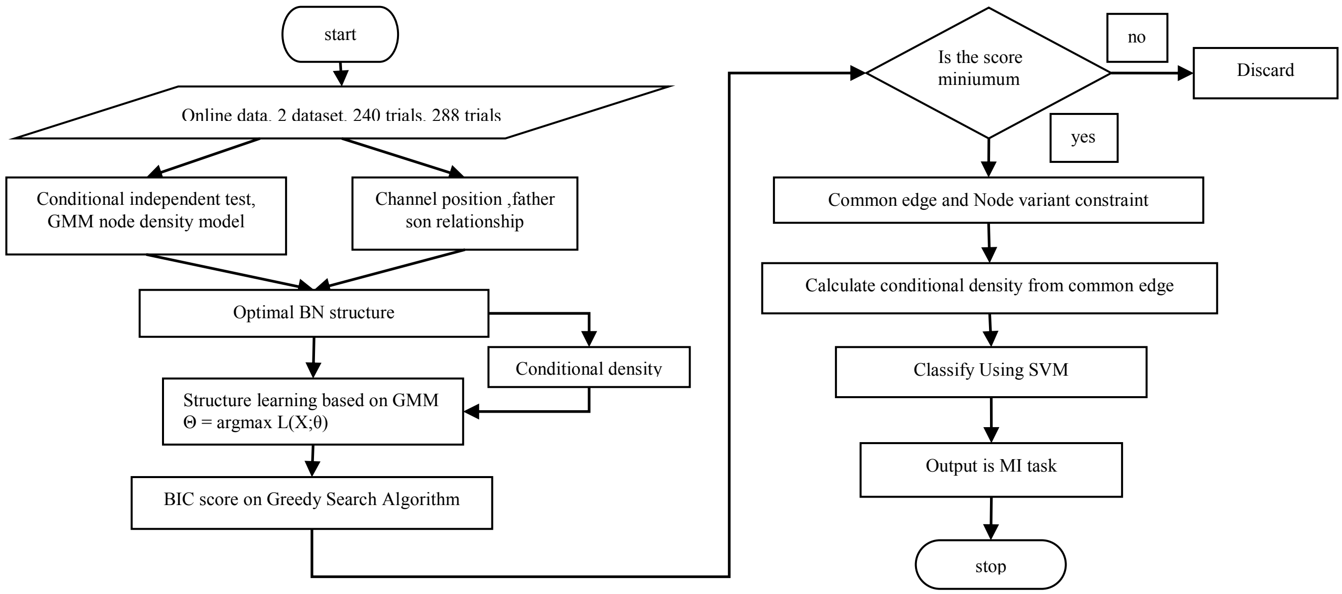

Fig 4: Flowchart showing feature extraction technique for Common Bayesian Network [16]

TABLE II . CLASSIFICATION ACCURACIES STUDIED

<table>
<tr><th>Paper</th><th>Data set/ subject</th><th>Proposed technique</th><th>Techniques compared</th><th colspan="5">Classification accuracy</th></tr>
<tr><td rowspan="3">[1]</td><td rowspan="3">10 subjects(22-26 years),4 channels, left hand –right hand movements, 40 MI trials</td><td rowspan="3">Fuzzy fusion</td><td rowspan="3">Conventional method, single LDA</td><td colspan="2">Fuzzy fusion</td><td colspan="2">with pso</td><td>w/o pso</td></tr>
<tr><td colspan="2">Sugeno</td><td colspan="2">0.968±0.063</td><td>0.998±0.040</td></tr>
<tr><td colspan="2">choquet</td><td colspan="2">0.992±0.014</td><td>0.998±0.003</td></tr>
<tr><td>[9]</td><td>5 subjects (2 cases healthy volunteer and seizure diagnosed volunteer), 100 single channel,opening and closing of eyes for 23.6 sec each ,4097 samples</td><td>Intrinsic mode function of EMD

Spectral and temporal features</td><td>Features compared

(EMD,IF,BW, wavelets)</td><td colspan="5">Techniques used : kNN (93%), Decision tree(96%), ANN(94%), SVM(93%)</td></tr>
<tr><td>[16]</td><td>Dataset III2a in BCI compitition III(3 subjects,40 trails 6 runs) and dataset 2a in BCI competition IV(9 subjects, 48 trials 6 runs)</td><td>Common Bayesian network (CBN)

CBN+ SVM</td><td>CSP+SVM

CSP+SVM+kNN+LDA

PCA+ICA+SVM</td><td colspan="5">The proposed method have kappa coefficient as 0.9

The 3 subjects as in DS2a K3, K6,L1 givs the accuracy of 0.98,0.88,0.82 respectively</td></tr>
<tr><td rowspan="2">[22]</td><td rowspan="2">Dataset IIIa BCI competition III, 60 channels , left-right hand tongue foot, 9runs(K3b) 360 trials, 6 runs (K6b,L1b) 240 trials</td><td rowspan="2">LPSVD</td><td rowspan="2">DCT ,AAR</td><td>subject</td><td>DCT</td><td>AAR</td><td>LPSVD 4 coeffi-cient</td><td>LPSVD Error variance</td></tr>
<tr><td>Avg of 3 subjects</td><td>42.54</td><td>61.16</td><td>58.19</td><td>67.35</td></tr>
<tr><td rowspan="2">[27]</td><td rowspan="2">9 subjects(maximum accuracy found in 5 subjects)</td><td rowspan="2">JAYA based NFCSSCGLH</td><td rowspan="2">SVM, SVM rbf, LDA</td><td>Subject</td><td>Proposed</td><td>SVM</td><td>SVM rbf</td><td>LDA</td></tr>
<tr><td>AVG of 9 subjects</td><td>76.47</td><td>67.22</td><td>69.02</td><td>66.97</td></tr>
</table>

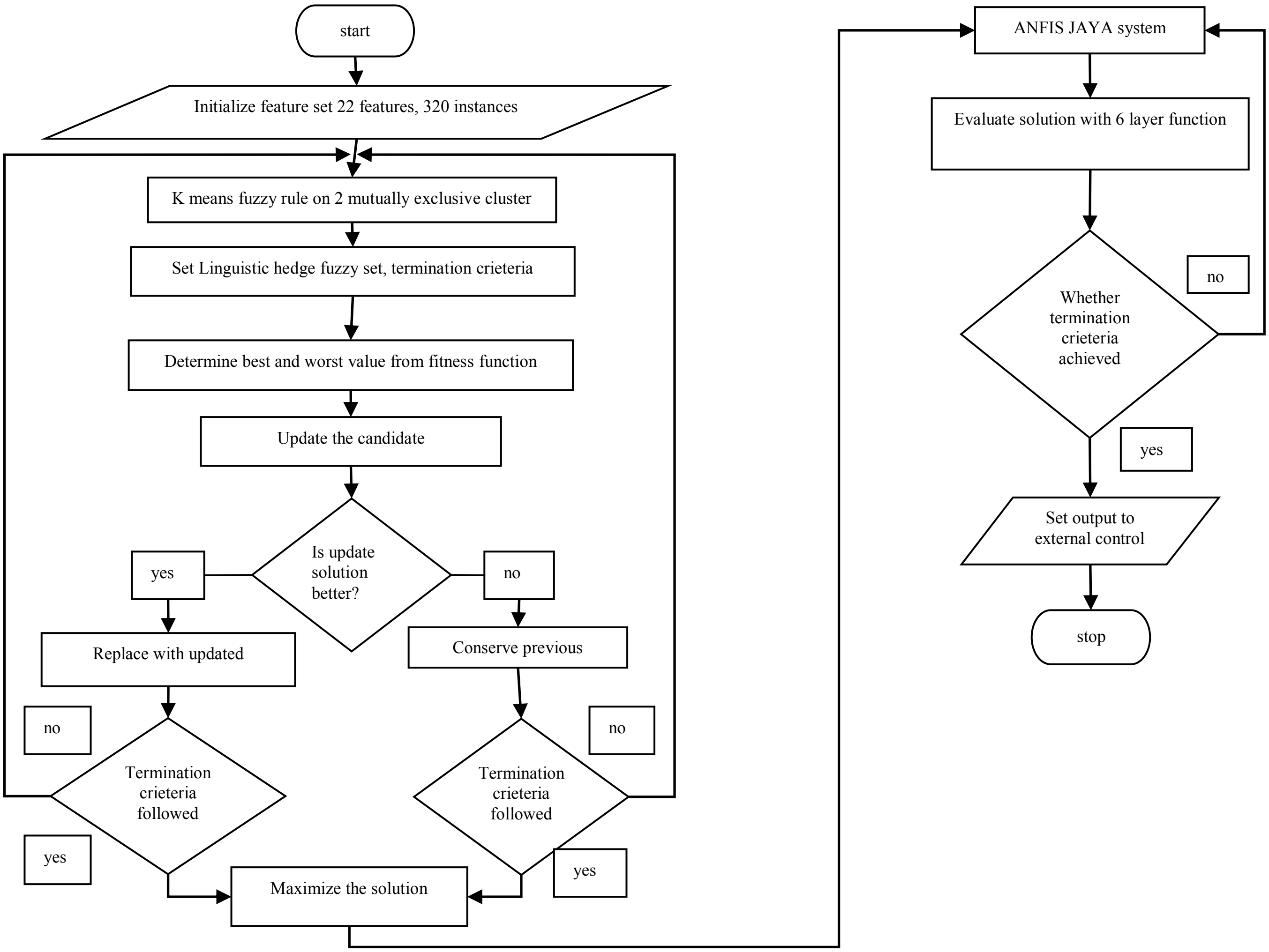

Fig 5: Flowchart showing feature extraction technique for Jayabased ANFIS [27]

REFERENCE

[1] Yu-Ting Liu, Tsung-Yu Hsieh, Yang-Yin Lin, Chih-Yu Chen, Chun-Hsiang Chuang and Chin-Teng Lin* Shang-Lin Wu, "Fuzzy Integral with Particle Swarm Optimization for a Motor-Imagery-based Brain-Computer Interface," *IEEE Transactions on Fuzzy Systems 25*, pp. 21-28, 2016.

[2] G. Cuntai, T. H. Dat, and X. Ping Q. Novi, "Sub-band Common Spatial Pattern (SBCSP) for Brain-Computer Interface," *International IEEE/EMBS Conference on Neural Engineering*, pp. pp. 204-207, 2007.

[3] B. S. Yoo and J. H. Kim, "Fuzzy Integral-Based Gaze Control of a Robotic Head for Human Robot Interaction," *IEEE Trans Cybern*, vol. 45, pp. pp. 1769-1783, 2015.

[4] D. T. Anderson, and D. J. Wescott M. F. Anderson, "Estimation of adult skeletal age-at-death using the Sugeno fuzzy integral," *Am. J. Phys. Anthropol.*, vol. 142, pp. pp. 30-41, 2010.

[5] V. K. Madasu, S. Srivastava, and M. Hanmandlu M. Singh, "Choquet fuzzy integral based verification of handwritten signatures," *J. Intell. Fuzzy Syst*, vol. 24, pp. pp. 145-161, 2013.

[6] S. L. Wu, C. H. Liou, S. W. Lu, S. A. Chen, S. F. Chen, et al. L. D. Liao, "A Novel 16-Channel Wireless System for Electroencephalography Measurements With Dry Spring-Loaded Sensors," *IEEE Trans. Instrument. Measure.*, vol. 63, pp. pp. 1545-1555, 2014.

[7] H. P. Hsu and J. S. Shih, "Multi-channel surface acoustic wave sensors based on principal component analysis (PCA) and linear discriminateanalysis (LDA) for organic vapors," *J. Chin. Chem. Soc.*, vol. 53, pp. pp. 815-824, 2006.

[8] L. Bianchi, L. R. Quitadamo, and G. Saggio F. Cavrini, "A Fuzzy Integral Ensemble Method in Visual P300 Brain-Computer Interface," *J Math Anal Appl*, vol. 2016, p. 9, 2016.

[9] Ali Hassan, Saad Rehman, Imran Khan Niazi, and Kim Dremstrup Farhan Riaz, "EMD-Based Temporal and Spectral Features for the Classification of EEG Signals Using Supervised Learning," *IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING*, vol. 24, pp. 28-35, january 2016.

[10] R. B. Pachori, "Discrimination between ictal and seizure-free EEG signals," *Res. Lett. Signal Process.,* vol. 2008, p. 14, 2008.

[11] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London. Ser. A, Math., Phys. Eng. Sci*, vol. 454, no. 1971, pp. 903–995, 1998.

[12] T. I. Netoff and S. J. Schiff, "Decreased neuronal synchronization during experimental seizures," *J. Neurosci.*, vol. 22, no. 16, pp. 7797-7307, 2002.

[13] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520-538, April 1992.

[14] R. J. Oweis and E. W. Abdulhay, "Seizure classification in EEG signals utilizing Hilbert-Huang transform," *Biomed. Eng. Online*, vol. 10, no. 1, p. 38, 2011.

[15] J. Weston, S. Barnhill, and V. Vapnik I. Guyon, "Gene selection for cancer classification using support vector machines," *Mach. Learn*, vol. 46, no. 1-3, pp. 389-422, 2002.

[16] Die Hu, Meng Wan and Ying Wen, Lianghua He, "Common Bayesian Network for Classification of EEG-Based Multiclass Motor Imagery BCI," *IEEE Transaction On Systems, Man, And Cybernatics: Systems*, vol. 46, no. 6, pp. 843-854, June VOL. 46, NO. 6, JUNE 2016.

[17] O. Fukuda, H. Ichinobe, and M. Kaneko T. Tsuji, ""A log-linearized Gaussian mixture network and its application to EEG pattern classification," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 29, no. 1, pp. 60-72, February 1999.

[18] R. Cabeza and Kingstone A., "Handbook of Functional Neuroimaging of Cognitive," *2nd ed. Cambridge, MA, USA: MIT Press*, 2006.

[19] M. C. Zhou, and G. Jiang, Z. Guo, "Adaptive optimal sensor placement and boundary estimation for monitoring mass objects," *IEEE Trans.Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 1, pp. 222-232, January 2008.

[20] D. M. Chickering, "Learning equivalence classes of Bayesian network Structures," *J. Mach. Learn. Res.*, vol. 2, pp. 445-498, March 2002.

[21] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1-27, 2011.

[22] AIDA KHORSHIDTALAB and MOSTEFA MESBAH HAMZA BAALI, "A Transform-Based Feature Extraction Approach for Motor Imagery Tasks Classification," *IEEE Journal of Translational Engineering in Health and Medicines*, vol. 3, p. 8, October 2015.

[23] D. Erdogmus, and R. Jenssen U. Ozertem, "Spectral feature projections that maximize Shannon mutual information with class labels," *Pattren Recognitiom.*, vol. 39, no. 7, pp. 1241-1252, 2006.

[24] A. Al-Ani and A. Al-Sukker, "Effect of feature and channel selection on EEG classification," *Proc. 28th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 2171-2174, Aug/Sept 2006.

[25] M. Fatourechi, R. K. Ward, and G. E. Birch A. Bashashati, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *J. Neural Eng.*, vol. 4, no. 2, p. R32, 2007.

[26] E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten M. Hall, "The WEKA data mining software: An update," *ACM SIGKDD Explorations Newslett.*, vol. 11, no. 1, pp. 10-18, 2009.

[27] RAKESH KUMAR SINHA AND SUBHOJIT GHOSH SURAJ, "Jaya Based ANFIS for Monitoring of Two Class Motor Imagery Task," *IEEE Access*, vol. 4, p. 10, November 2016.

[28] R. V. Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," *Int. J. Ind. Eng.Comput.*, vol. 7, no. 1, pp. 19-34, 2016.

[29] B. Cetisli, "The effect of linguistic hedges on feature selection: Part 2," *Expert Syst. Appl.*, vol. 37, no. 8, pp. 6102-6108, 2010.

[30] B. Ceti³li and A. Barkana, "Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classier training," *Soft. Comput.*, vol. 14, no. 4, pp. 365-378, 2010.

[31] Daniela S. Klobassa, Christoph Pokorny, and Gernot R. Muller-Putz Petar Horki, "Evaluation of Healthy EEG Responses for Spelling," IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 1, pp. 29-36, January 2015.

[32] Qiang Wang and Olga Sourina, "Real-Time Mental Arithmetic Task Recognition From EEG Signals," IEEE Transactions on Neural Systems and Rehabilitation Engineering , vol. 21, no. 2, pp. 225-232, March 2013.

[33] Julien Epps and Nigel H. Lovell. Pega Zarjam, "Beyond Subjective Self-Rating: EEG Signal Classification of Cognitive Workload," IEEE Transaction on Autonomous Mental Developement, vol. 7, no. 4, pp. 301-309, December 2015.

[34] Toshihisa Tanaka,Shunsuke Ono, Masao Yamagishi, and Hiroshi Higashi Naoki Tomida, "Active Data Selection for Motor Imagery EEG Classification," IEEE Transaction on Biomedical Engineering , vol. 62, no. 2, pp. 458-467, February 2015.

[35] Jacek Gwizdka, and W. Art Chaovalitwongse Shouyi Wang, "Using Wireless EEG Signals to Assess Memory Workload in the n-Back Task," IEEE Transaction on Human-Machine Systems , vol. 46, no. 3, pp. 424-435, June 2016.

[36] M. Capozzoli, L. McSweeney and D. Sinha M Banerjee, "Beyond kappa: A review of interrated agreement measures," Can.J.Stat, vol. 27, no. 1, pp. 3-23, March 1999.