Dynamic Initiation and Dual-Tree Complex Wavelet Feature-based Classification of Motor Imagery of Swallow EEG Signals

Huijuan Yang, Cuntai Guan, Kai Keng Ang, Chuan Chu Wang, Kok Soon Phua and Juanhong Yu
Institute for Infocomm Research,
Agency for Science, Technology and Research (A*STAR), Singapore 138632.
Email: {hjyang, ctguan, kkang, ccwang, ksphua and jyu}@i2r.a-star.edu.sg

***Abstract*—The use of motor imagery-based brain computer interface has recently been shown to have potential for rehabilitation. This paper proposes a novel scheme to detect motor imagery of swallow from electroencephalography (EEG) signals for dysphagia rehabilitation. The proposed scheme extracts features from the coefficients of dual-tree complex wavelet transform (DT-CWT). A novel sliding window-based peak localization scheme is proposed to dynamically locate the initiation of tongue movement from Electromyography (EMG) signal. Subsequently, effective time segments are extracted from EEG signal for classification based on the detected dynamic initiation location. Comparisons are made between our proposed scheme with that of the three existing approaches. The results based on six healthy subjects show that an increase in averaged accuracy of 9.95% is achieved. Further, an increase in averaged accuracy of 8.02% is resulted comparing our proposed scheme by using and not using the dynamic initiation to extract the time segments. Classification results using EMG data confirm that our results are not due to movements artifacts. Statistical tests with 95% confidence to estimate the accuracy on the respective action at chance level show that five out of six subjects performed above chance level for our proposed dynamic initiation and wavelet feature-based approach.**

***Index Terms*—brain computer interface, motor imagery of swallow, dynamic initiation localization, dysphagia.**

I. INTRODUCTION

Brain computer interface (BCI) has opened an effective way of communication between the patients and outside world, which has shown potential applications in the rehabilitation of patients [1][2]. The rationale is to harness plasticity to improve the recovery of human function after brain injury. The use of motor imagery for stroke rehabilitation is based on the assumption that motor imagery activates similar pathways as that of executed movements. It may improve functional recovery after stroke [3]. Detection of motor imagery of tongue movement [4] may be used for dysphagia rehabilitation considering the facts that tongue movement is an integral part of the complex swallowing process. Dysphagia or difficulty in swallowing, is the inability to swallow or difficulty in swallowing which is caused by stroke or other neuro-degenerative diseases. It occurs in approximately 30-42% of acute stroke patients requiring hospital admission [5]. Timely identification of dysphagia is essential to ensure timely treatment. The normal swallowing process starts from the mouth, passing through pharynx and esophagus, and finally reaching the stomach. In general, swallowing consists of four stages: the oral preparatory phase, oral transfer phase, pharyngeal phase (including swallow reflex initiation and pharyngeal function, and clearance) and esophageal phase [5]. Swallowing is a complex process that requires the integration of the respiratory center and motor function of multiple cranial nerves. Methods for treating swallowing disorders are: dietary changes, thermal stimulation, food and position changes, tongue strengthening exercises, pharyngeal maneuvers and neuro-muscular stimulation such as VitalStim [6]. Recently, transcranial magnetic stimulation is used to reorganize the human motor cortex based on the frequency, intensity and duration of the stimulus [7]. Accurate classification of the motor imagery from idle state can greatly facilitate the rehabilitation of the patients whose limbs are paralyzed. Further, combining motor imagery with passive way of moving can provide a quality life to stroke patients. However, to the best of our knowledge, there is currently no work on the detection of motor imagery of swallow from EEG signal for dysphagia rehabilitation. Hence, this paper firstly proposes a novel algorithm to detect the motor imagery of swallow using classification techniques. Secondly, we propose a novel technique to dynamically locate the initiation of tongue movement using EMG signal, which is eventually used to extract effective time segments from the EEG signals to boost the classification performance. Thirdly, we investigate how to extract discriminate features from the EEG signals using the dual-tree complex wavelet transform.

II. OUR PROPOSED SCHEME

A. Dynamic Initiation (DI) Localization

In synchronous BCI, subjects perform required tasks in synchronization with the given cues [8][9][10][11][12]. However, the actual response from different subjects are different even if the cue is given in a fixed time interval. If the actual initiation time point can be localized, the exact time segment can be extracted for the classification to boost the performance. The dynamic initiation is located using the EMG signals recorded in tongue movement, which are firstly extracted from the entire time segment for one trial. For example, a time segment of [-6 7]s from the onset of the action is used. The mean of the signals across trials, and a baseline from 300ms to 50ms before the onset of the visual cue are deducted from the

signal. The complex morlet wavelet transform [13] centered at frequency from 6Hz to 30Hz (with an increment of 2Hz each time) is computed, which is subsequently convolved with the signal to obtain the time-frequency power. These centering frequency bands are chosen due to the effectiveness in detecting event-related desynchronization (ERD) and event-related synchronization (ERS) during motor preparation, execution and imagery in brain signals. Five seconds from the beginning and two seconds from the ending of one trial are skipped due to their instability, namely, "*skipped segments*". Skipping these time segments is based on the assumption that the actual initiation is located near the visual cue. Hence, there is no need to consider the time segments that lie far from the cue.

A sliding window of length w_s=$\alpha \times f_s$ is used to identify the local peaks, where f_s is the sampling rate, α is the parameter to control the window size. There is a tradeoff between the size of sliding window and the quality of the detected peaks. If the window size is chosen to be too large, some small peaks may be missed out. On the other hand, if the window size is chosen to be too small, some false small peaks will be detected. In the experiments, α=1 is chosen to achieve good detection results. The power energy is averaged across channels and the centering frequency bands. The distinctive peaks are identified by our proposed "*double thresholding*" technique, which is described as follows.

The global threshold T_g is defined as the mean power of the chosen samples.

$$T_g = \frac{1}{N_g}(\sum_{j=1}^{N_g}(P_w(j))) \tag{1}$$

where N_g is the total number of samples in the current trial by excluding the *skipped segments*; $P_w(j)$ is the power energy of the jth sample. T_g is generally larger since the dominant peaks are much higher than that of the other samples. In order not to miss out any peaks, a lower threshold T_l is calculated by

$$T_l = \frac{1}{N_s}\sum_{i=1}^{N_s}(P_w(i)|P_w(i) < T_g) \tag{2}$$

where N_s is the number of samples satisfying the condition; The status of each sample, i.e., $S_s(i)$, is given by

$$S_s(i) = \begin{cases} 1 & \text{if } mean(P_w(i+k)) \geq T_l \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $k \in [-\frac{w_s}{2}\ \frac{w_s}{2}]$ and $mean(X)$ is the sample mean of set X. The qualified samples are those whose sample means in the window centered at the candidate are higher than T_l. Finally, the peak in the window is identified as

$$\hat{k} = \underset{k,k\in[-\frac{w_s}{2}\ \frac{w_s}{2}]}{\arg\max}(P_w(i+k)|S_s(i)=1) \tag{4}$$

$$i = \hat{k} \tag{5}$$

where w_s and i are the window size and index of the candidate sample, respectively. In this case, the candidate sample is the local maxima in the window. The first peak is chosen as the dynamic initiation location for the current trial. Finally, the median location of the subject-specific dynamic initiation is used to extract effective time segments for the classification of motor imagery of swallow versus the background state: idle.

B. Feature Extraction in Dual-tree Complex Wavelet Transform Domain

The shift invariance and rich directionality of DT-CWT increase the flexibility in designing spatially adaptive filters without introducing undesirable aliasing artifacts [14]. DT-CWT is a redundant, non-separable and complex transform. The redundancy of the DT-CWT is a factor of 2^d, where d is the dimension of the signal. In this paper, we investigate how to extract features from the coefficients of DT-CWT of the EEG signals. Assume the signal is decomposed into J=4 levels and the sampling frequency of the EEG signal is 250Hz. The data is firstly appended to the integer times of 2^J. The frequency ranges represented by the wavelet coefficients and the corresponding EEG rhythm of the event-related desynchronization/synchronization (ERD/ERS) are shown in Table I.

TABLE I
FREQUENCY RANGES REPRESENTED BY WAVELET COEFFICIENTS.

Coefficients	frequency ranges (Hz)	EEG rhythm
$C_{1,d,r}$	64-128	high γ wave
$C_{2,d,r}$	32-64	γ wave
$C_{3,d,r}$	16-32	α wave
$C_{4,d,r}$	8-16	β wave
$C_{5,a,r}$	0-8	δ, θ wave

It can be seen from the table, θ and δ are represented by coefficients $C_{J+1,d,r}$; β, α, γ and high γ waves are represented by $C_{s,d,r}$, where s ($s \in \{4,3,2,1\}$), d ($d \in \{1,2\}$) and r ($r \in \{1,2\}$) denote level of decomposition, directions, and real and imaginary parts of the complex coefficient $C_{s,d,r}$, respectively. Hence, the features are actually extracted from the coefficients of a wide range of EEG rhythms, e.g., ranging from δ to high γ, which consists of three parts. Firstly, the features consist of the energy of coefficients ($F^e_{s,d}$). This is based on the facts that the energy of EEG signals in different frequency band is important in detecting the event-related synchronization and event-related dysynchronization. Secondly, the features consist of the phase in each level and direction ($F^p_{s,d}$), which are important in discriminating different tasks. Thirdly, the features consist of the coarse representation of the coefficients ($F^a_{s,d}$). These features are calculated by

$$F^e_{s,d} = \sum_{s=1}^{J}\sum_{d=1}^{2}((C_{s,d,1})^2 + (C_{s,d,2})^2) \tag{6}$$

$$F^p_{s,d} = \sum_{s=1}^{J}\sum_{d=1}^{2} atan(\frac{C_{s,d,2}}{C_{s,d,1}}) \tag{7}$$

$$F^a_{s,d} = C_{J+1,d,1} \tag{8}$$

where $atan(x)$ is the arctangent of element x. The features further consist of the median and variance of $F^e_{s,d}$ and $F^p_{s,d}$,

and the coarse representation is corrected by its mean, which are given by

$$F^e_{s,d} = F^e_{s,d}||med(F^e_{s,d})||var(F^e_{s,d}) \quad (9)$$
$$F^p_{s,d} = F^p_{s,d}||med(F^p_{s,d})||var(F^p_{s,d}) \quad (10)$$
$$F^a_{s,d} = F^a_{s,d} - mean(F^a_{s,d}) \quad (11)$$

where "||" is the "concatenation" operation; $med(X)$, $mean(X)$ and $var(X)$ give the median, mean and variance of the elements in set X, respectively. Normalization is then applied to $F^k_{s,d}$, where $k \in \{e,p,a\}$ by the maximum value in each feature set, which are given by

$$\tilde{F}^k_{s,d} = F^k_{s,d}/max(abs(F^k_{s,d})) \quad (12)$$

where $max(X)$ gives the maximum values in set X. Finally, the wavelet-based feature vector (F_w) is given by

$$F_w = \tilde{F}^e_{s,d}||\tilde{F}^p_{s,d}||\tilde{F}^a_{s,d} \quad (13)$$

Note that the time series is appended to integer times of 2^J, which is 256 in the implementation. Hence, the total number of features for each channel is: N_{fc}=$256/(2^J)$+$4(J+2)$, which is 40 for J=4.

III. EXPERIMENTAL RESULTS

A. Experimental Paradigm

Six healthy subjects without any history of respiratory, swallowing or neurological disorder, and eating or nutrition problems participated in the experiments. The experiments consist of two active and one passive tasks. The subjects are instructed to imagine swallowing a cup of water for *motor imagery of swallow (MI)*; and to physically move the tongue freely within the oral cavity as many times as possible for *tongue movement (TN)*. Nevertheless, the subjects should not perform any of the above tasks during idle state and should not close their eyes.

Each session consists of three runs with short breaks between any two runs. Each run consists of 20 trials for each action, yielding a total of 60 trials per session for each action. The timing scheme is shown in Fig. 1. In the experiments, the subjects were sitting in a comfortable armchair in front of a computer screen. A resting state shown by a progress blue bar is firstly appeared on the black screen. A short acoustic tone was presented at 5s. After one second (t=6s), a cue in the form of an image lasting for 1s is shown to prompt the subject to perform the required actions, e.g., motor imagery of swallow, tongue movement and idle. Each action lasts for 6s, followed by a short break of 6s. Note that the sequence of the actions is randomized.

The EMG recording is done using two pairs of electrodes taped beneath the skin of the submental and infrahyoid muscle groups for maximal signal detection, as shown in Fig. 2. The choice of these locations to measure the electromyographic activity is based on the facts that submental muscles are connected to hyoid bone, which is important in hyoid laryngeal excursion during swallowing [15]. The electrodes are connected to Neuroscan NuAmps EEG amplifier. In total, 38

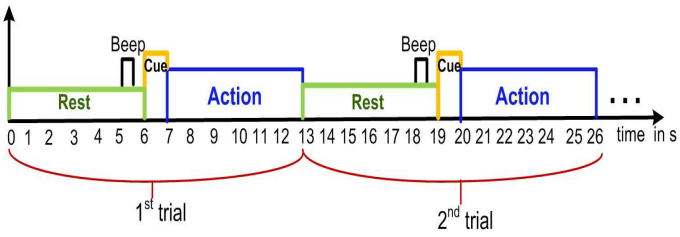


Fig. 1. The timing scheme of the experimental paradigm. The "action" refers to "motor imagery of swallow", "tongue movement" and "idle".

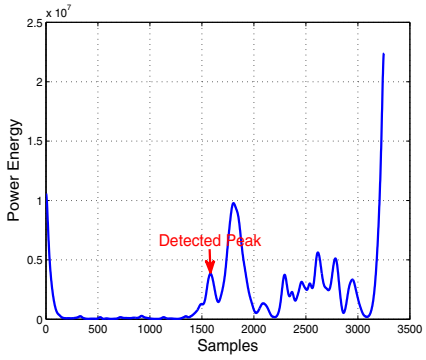
Fig. 2. Locations of two pairs of electrodes for EMG recording.

channels are used. Four channels, i.e., HEOL, HEOR, VEOU and VEOL, are used for EMG recording. The other thirty-two channels, i.e., Fp1, Fp2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz, O2, PO1 and PO2, are used for EEG recordings. While the average of signals in channels A1 and A2 (i.e., $\frac{A_1+A_2}{2}$) is used as the reference. The EEG and EMG are bandpass filtered between 0.5Hz and 100Hz and a notch filter of 50Hz is enabled. The sampling rate is set to 250Hz. Note that in the following experiments, the data acquired for 4 channels of EMG and 32 channels of EEG are used. 10×10 fold cross-validation is employed to analyze the performance and Support Vector Machine (SVM) with linear kernel function is chosen as the classifier.

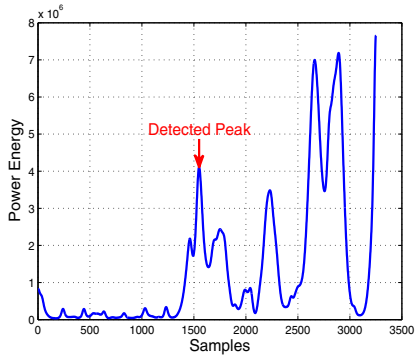
B. Dynamic Initiation Detection and Performance of Using Different Window Sizes

The dynamic initiation identified using the EMG signal of tongue movements is used to extract the effective time segments for the classification of motor imagery of swallow versus idle. This is based on the assumption that the response time of each person to the actions is similar. Two examples of the located onset of the initiations of tongue movements are shown in Fig. 3, which confirm the effectiveness of the proposed peak localization technique.

Comparison is made on how the classification performance will be affected by choosing different sizes of the sliding window to detect the initiation of tongue movement. Generally, choosing a small window size tends to select the small, local peaks. While choosing a large window size tends to select the larger global peaks. The classification accuracies of choosing different window sizes are presented in Fig. 4. The average accuracies are: 65.49%, 66.35%, 67.57%, 69.96%, 66.87% and 66.17% for window sizes of $1/4f_s$, $1/3f_s$, $1/2f_s$, f_s, $2f_s$ and $4f_s$, respectively, where f_s is the sampling rate of



(a)



(b)

Fig. 3. Examples of the located onset of the initiations of tongue movements (marked in "red" arrow).

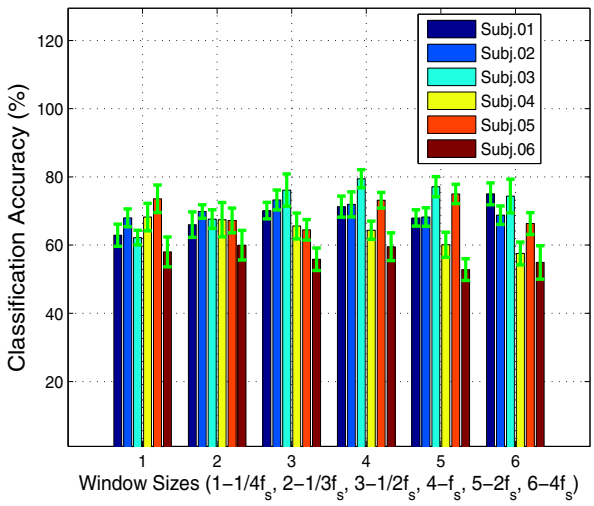


Fig. 4. Classification accuracy (motor imagery of swallow versus idle) achieved by choosing different window sizes to detect the dynamic initiation of tongue movement, where f_s is the sampling rate.

the signal. Note that the linear SVM classifier is used for the classification for all the window sizes and other parameters of SVM classifier are not optimized catering for different window sizes. The average accuracy and variance for the six subjects are: 68.88%±18.01, 70.01%±4.59, 72.82%±43.23, 63.88%±17.83, 69.96%±19.86, 56.85%±7.86, respectively. From these results, it is not difficult to see that all the subjects except subject 6 perform well above the chance level. It should be noted that subject 6 was a bit sleepy during the experiments, which may affect the performance. It can be further noticed that the accuracy does not vary much by choosing different window sizes, indicating the reliability of the method.

C. Performance Comparison of Using and Not Using Dynamic Initiation, and Classification Using EMG Signals

Experiments are conducted to show the advantages of proposed classification scheme to detect the motor imagery of swallow using dynamic initiation-based time segments (P-DI), which is compared with our scheme that uses the fixed cue-based time segments (P-FC). For this purpose, various time segments with reference to the onset of the action are selected for the fixed cue-based scheme, with the comparison results shown in Table II. The results confirm that our proposed dynamic initiation-based scheme is more effective in detecting the motor imagery of swallow, compared with that of the fixed cue-based scheme. A paired sample t-test at significance level of 0.07 is conducted to test the null hypothesis that the accuracies for the proposed scheme using dynamic initiation (P-DI) and the average accuracies of using fixed cue (AP-FC) come from populations with equal means. The results show that the hypothesis is rejected with p values of 0.0466, indicating the significant difference exists in the accuracies by using and not using dynamic initiation to extract the time segments for the classification. It should be noted that the fixed time segments are used in our current approach. Optimizing the classification accuracy by choosing different time windows for each subject, i.e., selection of subject-specific time windows, is also possible. The training model obtained using the optimized time segments can eventually be used to classify the test data for each subject.

Experiments are further conducted to confirm that the classification results are based on the motor imagery of swallow EEG signal and not due to other movement artifacts. For this purpose, we classify the motor imagery of swallow versus idle using only the EMG data. A time segment of $[-0.5s\ 0.5s]$ from onset of the action is used. The same wavelet features (discussed in Section II. B) are extracted for the classification and the obtained accuracies and variances are: 51.48%±4.14, 46.69%±3.51, 51.32%±2.76, 48.57%±3.28, 57.35%±2.93 and 57.58%±3.77 for the six subjects, respectively. This further demonstrates that most of the subjects performed at chance level, indicating no EMG artifacts exist, whereas only subjects 5 and 6 show slight EMG artifacts.

The 95% confidence estimate of the accuracy on the respective action at chance level is approximately 0.41 to 0.58 using binomial inverse cumulative distribution function. This indicates those subjects whose accuracy falls between 41% and 58% can be deemed as performing at chance level. The average accuracy of using different window sizes to detect the dynamic initiation (presented in Section III. B) demonstrates that five out of six subjects performed above chance level,

TABLE II
CLASSIFICATION ACCURACIES OF MOTOR IMAGERY OF SWALLOW FOR THE PROPOSED METHODS USING/NOT USING DYNAMIC INITIATION.

Methods	Time Segments(s)	Sub. 1 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 2 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 3 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 4 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 5 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 6 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Ave. Acc. ($\mathcal{A}_{as} \pm \mathcal{V}_r$)
Proposed fixed cue (P-FC)	[-1 0]	65.94±1.69	61.55±3.38	64.99±4.14	67.54±1.99	62.67±3.70	59.53±2.43	63.70±2.89
	[0 1]	73.07±3.06	46.53±3.42	71.84±2.12	66.13±4.08	60.45±2.74	55.28±2.60	62.22±3.00
	[0.5 1.5]	66.53±3.92	52.62±2.61	67.98±2.35	66.65±2.91	56.56±2.84	55.65±3.97	61.00±3.10
	[1 2]	72.34±2.60	57.25±2.64	65.80±5.09	69.40±1.29	67.25±2.80	49.46±4.83	63.58±3.21
	[1.5 2.5]	72.02±4.01	53.02±4.35	58.64±3.41	69.77±3.63	50.23±3.54	51.63±2.57	59.22±3.59
Average of P-FC (AP-FC)		69.98±3.06	54.19±3.28	65.85±3.42	67.90±2.78	59.43±3.12	54.31±3.28	61.94±3.16
Proposed dyn. Initiation (P-DI)	dynamic	71.27±3.08	71.95±3.66	79.47±2.64	64.36±1.91	73.16±2.29	59.54±4.06	**69.96±2.94**

$\mathcal{A}_c$: accuracy (%); $\mathcal{V}_r$: variance; $\mathcal{A}_{as}$: average accuracy across subjects (%). Note that time segments shown in the table are relative to the onset of action time. The best average accuracies achieved among all subjects to detect the motor imagery of swallow is our proposed dynamic initiation (DI)-based scheme, shown in black in the last column.

for the classification of the motor imagery of swallow using our proposed dynamic initiation and wavelet feature-based approach.

D. Performance Comparisons Between Our Proposed and Other Methods

To show the efficacy of our proposed dynamic initiation (DI) and wavelet features-based classification scheme, it is compared with several existing methods: 1) common spatial pattern (CSP)-based method [8][10]. In CSP, the EEG signals are decomposed into the spatial patterns such that the discriminability between the two classes is maximized. It is achieved by jointly diagonalizing the covariance matrices of the EEG signals of each class, which is one of the most popular feature extraction techniques for EEG signal classification. 2) filter bank CSP (FBCSP)-based method [9][11]. In FBCSP, the discriminant frequency bands (DFBs) are chosen by computing the mutual information between the CSP features and that of the class labels, which has shown good performance on the competition data set. 3) sliding window discriminative CSP (SWDCSP)-based method [12]. It selects the DFBs based on the CSP features from the overlapping frequency bands using the clustering technique such as affinity propagation. Time segments of 0.5s prior and after the detected dynamic initiation and the onset of the visual cue are extracted for the proposed dynamic initiation-based, and other methods, respectively, the results are shown in Table III. The average accuracy of CSP, FBCSP, SWDCSP and our proposed method using dynamic initiation are: 58.83%, 63.39%, 57.82% and 69.96% for the classification of motor imagery of swallow versus background idle state. On average, the accuracy achieved by our proposed dynamic initiation-based approach is 9.95% higher than that achieved by the three existing methods, indicating significant superior performance of our proposed scheme.

To statistically validate our arguments, a null hypothesis test is conducted, which is defined as: the difference in accuracy of motor imagery of swallow versus idle for our proposed method using dynamic initiation localization, and that of other methods (e.g., CSP, FBCSP, SWDCSP), is a random sample from a normal distribution with mean zero. A paired sample t-test conducted with a significance level of 0.05 shows that the null hypothesis is rejected for the methods of CSP and SWDCSP (p=0.0366 and 0.0019). Nevertheless, the null hypothesis cannot be rejected for FBCSP (p=0.0816) at the significance level of 0.05. Hence, the accuracies of proposed approach for motor imagery of swallow using dynamic initiation is significantly different from that of CSP and SWDCSP, but not so significant different from that of FBCSP. Same arguments can be made for the tongue movements, in which the p values are 0.0007, 0.1338 and 0.0029 for the methods of CSP, FBCSP and SWDCSP, respectively.

IV. CONCLUSIONS

In this paper, we investigate the novel problem of detection of motor imagery of swallow from EEG signals for dysphagia rehabilitation of stroke patients. Specifically, we propose a novel scheme to locate the dynamic initiation of tongue movement using EMG signals, which has subsequently been employed to extract the effective time segments for the classification of motor imagery of swallow versus background idle state. This has led to an increase in accuracy of 8.02% compared with that obtained using the time segments extracted based on the fixed cue. Further, we propose a novel feature extraction scheme in the dual-tree complex wavelet transform domain. The energy and phase of the wavelet coefficients in different levels and directions, and coarse representation of the signal are used as the features. Comparison results show that the accuracy obtained using our proposed dynamic initiation and wavelet feature-based scheme is on average 9.95% higher than that obtained using the three existing approaches. Classification results using the EMG data confirm that our results are not due to movements artifacts. Statistical tests with 95% confidence show that five out of six subjects performed above chance level. It is noted that the current results are based on six healthy subjects. Our future work will focus on conducting clinical trials using proposed method for the rehabilitation of patients with dysphagia.

REFERENCES

[1] J. J. Daly, J. R. Wolpaw, "Brain-Computer Interfaces in Neurological Rehabilitation," *Lancet Neurology*, vol. 7, no. 11, pp. 1032-1043, 2008.

TABLE III
COMPARISON OF CLASSIFICATION ACCURACIES FOR MOTOR IMAGERY OF SWALLOW AND TONGUE MOVEMENTS USING DIFFERENT METHODS.

Methods	Task	Sub. 1 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 2 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 3 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 4 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 5 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Sub. 6 ($\mathcal{A}_c \pm \mathcal{V}_r$)	Ave. Acc. ($\mathcal{A}_{as} \pm \mathcal{V}_r$)
CSP	MI	56.17±2.22	52.09±2.29	56.92±3.62	65.83±2.64	65.08±3.26	56.92±2.77	58.83±2.79
	TN	61.05±4.36	71.00±2.90	68.17±3.30	73.33±3.37	77.42±2.28	58.83±3.34	68.38±3.26
FBCSP	MI	63.00±3.27	62.00±2.89	75.83±3.45	69.42±2.64	55.92±3.15	54.17±3.81	63.39±3.20
	TN	61.83±3.35	88.67±1.68	82.17±3.63	80.83±1.11	94.50±1.37	65.75±4.49	78.96±2.60
SWD-CSP	MI	56.85±3.72	60.19±2.01	66.07±2.02	59.70±3.96	53.82±5.19	50.30±4.03	57.82±3.49
	TN	58.16±3.18	79.91±2.47	75.22±1.60	70.24±3.10	84.06±1.92	54.62±2.35	70.37±2.44
Proposed (DI)	MI	71.27±3.08	71.95±3.66	79.47±2.64	64.36±1.91	73.16±2.29	59.54±4.06	**69.96±2.94**
	TN	85.09±2.62	90.48±2.46	91.58±1.78	87.45±1.80	90.54±1.48	69.21±3.56	**85.72±2.28**

The best average accuracies achieved among all subjects and different methods for motor imagery of swallow and tongue movement are our proposed dynamic initiation (DI)-based scheme, shown in black of last column.

[2] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin and H. Zhang, "A Large Clinical Study on the Ability of Stroke Patients in Using EEG-Based Motor Imagery Brain-Computer Interface," *Clinical EEG and Neuroscience*, pp. 253-258, Oct. 2011.

[3] N. Sharma, V. M. Pomeroy and J.-C. Baron, "Motor Imagery: A Backdoor to the Motor System After Stroke?" *Stroke,* vol. 37, pp. 1941-1952.

[4] M. Naeem, C. Brunner, R. Leeb, B. Graiman and G. Pfurtscheller, "Separability of Four-class Motor Imagery Data Using Independent Components Analysis," *J. Neural Eng.*, vol. 3, pp. 208-216, 2006.

[5] Martin J. McKeowna, Dana C. Torpeyd, Wendy C. Gehmf, "Non-invasive Monitoring of Functionally Distinct Muscle Activations During Swallowing," *Clinical Neurophysiology,* vol. 113, pp. 354-366, 2002.

[6] M. Kiger, C. S. Brown and L. Watkins, "Dysphagia Manage-ment: An Analysis of Patient Outcomes Using VitalStim^{TM} Therapy Compared to Traditional Swallow Therapy," Dysphagia, vol. 21, no. 4, pp. 243-253, 2006.

[7] C. Fraser, M. Power, S. Hamdy, J. Rothwell, D. Hobday, I. Hollander, P. Tyrell, A. Hobson, S. Williams and D. Thompson, "Driving Plasticity in Human Adult Motor Cortex Is Associated with Improved Motor Function after Brain Injury," *Neuron*, vol. 34, no. 5, pp. 831-840, May 2002.

[8] H. Ramoser, J. Muller-Gerking and G. Pfurtscheller, "Optimal Spatial Filtering of Single Trial EEG during Imagined Hand Movements," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 4, pp. 441-446, 2000.

[9] K. K. Ang, Z. Y. Chin, H. Zhang and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface," *IEEE World Congress on Comp. Intel. & Int. Joint Conf. on Neural Networks (IJCNN)*, pp. 2390-2397, June 2008, Hong Kong.

[10] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Muller, "Optimizing Spatial Filters for Robust EEG Single-Trial Analysis," *IEEE Signal Proc. Magazine,* vol. 25, no. 1, pp. 41-56, 2008.

[11] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan,"Mutual Information-based Selection of Optimal Spatial-Temporal Patterns for Single-Trial EEG-based BCIs," *Pattern Recognition*, vol. 45, no. 6, pp. 2137-2144, June 2012.

[12] G. Sun, J. Hu and G. Wu, "A Novel Frequency Band Selection Method for Common Spatial Pattern in Motor Imagery Based Brain Computer Interface," *IEEE World Congress on Comp. Intel. (WCCI 2010)*, pp. 335-340, July 2010, Spain.

[13] Catherine Tallon-Baudry, O. Bertrand, F. Peronnet, and J. Pernier, "Induced γ-Band Activity during the Delay of a Visual Short-Term Memory Task in Humans," *The Journal of Neuroscience*, vol. 18, no. 11, pp. 4244-4254, June 1998.

[14] Ivan W. Selesnick, Richard G. Baraniuk, and Nick G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," *IEEE Signal Process. Magazine*, pp. 123-151, Nov. 2005.

[15] K. M. Wheeler, T. Chiara and C. M. Sapienza, "Surface Electromyographic Activity of the Submental Muscles during Swallow and Expiratory Pressure Threshold Training Tasks," *Dysphagia,* vol. 22, pp. 108-116, 2007.