35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Investigation of the trade-off between time window length, classifier update rate and classification accuracy for restorative brain-computer interfaces

Sam Darvishi, Michael C. Ridding, Derek Abbott *Fellow IEEE*, Mathias Baumert, *Member IEEE*.

Abstract— **Recently, the application of restorative brain-computer interfaces (BCIs) has received significant interest in many BCI labs. However, there are a number of challenges, that need to be tackled to achieve efficient performance of such systems. For instance, any restorative BCI needs an optimum trade-off between time window length, classification accuracy and classifier update rate. In this study, we have investigated possible solutions to these problems by using a dataset provided by the University of Graz, Austria. We have used a continuous wavelet transform and the Student t-test for feature extraction and a support vector machine (SVM) for classification. We find that improved results, for restorative BCIs for rehabilitation, may be achieved by using a 750 milliseconds time window with an average classification accuracy of 67% that updates every 32 milliseconds.**

I. INTRODUCTION

Based on the findings that imagination of motor functions can facilitate stroke rehabilitation [1], BCIs have been used by many BCI groups as a tool to assist stroke patients with mental practice [2-6]. Specific kinds of BCI systems, which are called *restorative BCIs*, have been proposed to reorganize the impaired neural networks in stroke patients through motor imagery provided in BCI sessions [7]. In contrast to invasive BCI, which involves implanting electrodes on the surface or within the brain, restorative BCI systems that are used for stroke rehabilitation use data collected non invasively. The signal of interest could be from the electroencephalogram (EEG), magnetoencephalogram (MEG), functional magnetic resonance imaging (fMRI), or near infrared spectroscopy (NIRS). Note that, EEG-based BCI, however, is the most commonly used modality for stroke rehabilitation and thus was selected in this study. Mechanisms of recovery (largely neuroplastic) are not fully understood. However, Hebbian based learning and metaplastic effects are thought important [8]. Hence, presuming the importance of coincidence between motor modulation and sensory feedback, providing real-time or near real-time feedback during training is expected to have a constructive role in this regard.

S. Darvishi, M. Baumert, and D. Abbott are with the Centre for Biomedical Engineering, School of Electrical and Electronic Engineering, The University of Adelaide, SA 5000 (phone:+618-8313-4115; fax:+618-8313-4360; e-mail: sam.darvishi@ adelaide.edu.au, mathias.baumert@ adelaide.edu.au, dabbott@eleceng.adelaide.edu.au)
M.C. Ridding is with the Robinson Institute, School of Pediatrics and Reproductive Health, The University of Adelaide, SA 5005 (michael.ridding@adelaide.edu.au)

In a recent study [5] a motor imagery based BCI system was used to actuate a robotic system for upper limb rehabilitation of stroke patients. In that study classifier result updated based on subject motor imagery or motor execution power spectrum density (PSD) of mu [8-13 Hz] and beta [18-26 Hz] bands during the first 500 milliseconds after cue onset and then every 300 ms compared to their resting PSD. Moreover, Buch *et al.* (2007) [4] and Shindo *et al.* (2011) [6] in very similar designs, used motor imagery based BCI and visual feedback for training aimed at recovery of finger movements. Complementary proprioceptive feedback was also provided after successful modulation of the requested motor imagery at the end of the trial. Besides a number of study design similarities between Buch *et al.* and Shindo *et al.,* they also had two main differences. One difference was in the signal acquisition, where Buch *et al.* used MEG while Shindo *et al.* used EEG. The other difference in their study designs was the classifier update rate, which was 300 ms for the former while it was defined to be 30 ms in the latter.

Even though in a previous study [9] the optimal delay time for feedback provision was proposed to be in the 200-300 ms range, no specific reason for suitability of that delay was provided. Considering the variable results reported in these studies, here we propose to investigate potential effects that different timing may have on the achieved results.

In addition, application of BCI for stroke rehabilitation typically starts with a calibration session to extract the optimum channels and frequency bands for feature classification followed by training the patients based on the results of the calibration session. However, as has been mentioned in [6], stroke patients experience various difficulties, including spasticity in their muscles and lack of proper sleep. Thus, minimizing the number of sessions may improve stroke patients' adherence to BCI based therapy.

Thus we also examined whether it is viable to improve classification accuracy of a motor imagery BCI system for a typical healthy subject by adding further healthy subject training data to train the classifier (subject- independent classifier) compared to the method that uses only training data of the same subject for testing its classification accuracy (subject-dependent classifier).

II. METHODS

A. Data set

We utilized a widely used data set (Dataset 2b of BCI 2008 competition) from the University of Graz, which is accessible via http://www.bbci.de/competition/iv/. This

Table. 1: Dataset Specifications

Data set Name	BCI Competition	Number of classes	Channels used	Filter Bandwidth	Sampling rate	Number of subjects	Training trials (clean)	Test trials (clean)
2b	2008	2 (Left/ Right hand)	2 (C3, C4)	0.5–100 Hz	250 Hz	9	1182	2239

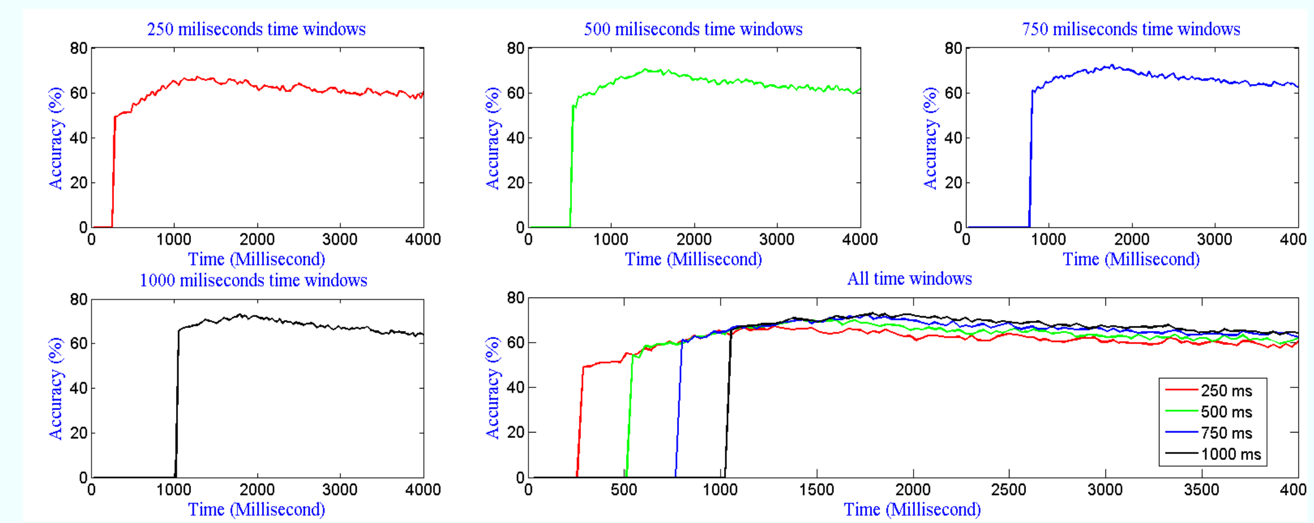

Fig1. Classification accuracy results for different time window lengths

dataset comprises EEG data from 9 healthy subjects, who were asked to modulate motor imagery based on provided visual cues. Each subject had 120 calibration trials with no feedback for each hand's movement imagery, 80 training trials for each hand (with feedback) and 320 trials for testing. For every subject, a number of trials (around 20%) were contaminated with EOG artifacts that were rejected and analysis was only conducted on clean data. Data has been recorded with 3 electrodes (C3, Cz and C4). During training and test sessions, continuous visual feedback was provided after 2 seconds. Table 1 summarizes the dataset features.

B. Feature extraction

It has been shown in [10] that desynchronization followed by synchronization of event related potentials (ERD/ ERS) in both the mu and beta bands occurs during motor imagery. Furthermore, such ERD/ERS only occurs in short and non-stationary periods. Thus, we decided to use a continuous wavelet transform based on the Morlet wavelet for feature extraction, as it has been shown to be more powerful than its discrete version in extraction of subtle EEG features [11]. Moreover, it has been demonstrated in [10] that 500 ms after cue onset in beta band, ERD in the contralateral hemisphere is coincident with ERS in the ipsilateral hemisphere. Thus, to be able to extract such representative features, that in turn leads to higher classification accuracy, we based our feature reduction methodology on finding the most discriminant features using the Student t-test. The full description of our feature extraction methodology is provided in our previous work [12]. Since the exploited data set contained 3421 trials, to make its classification more computationally efficient, we decided to only extract 6 features per channel for each [8-13 Hz] (mu) and [18-26 Hz] (beta) frequency band that in total provided 24 features for each trial.

C. Classification

Even though we showed in our previous work [12] that neuro-fuzzy classifiers outperform SVM in accuracy, linear SVM provides faster results for online applications. Thus we used a linear SVM classifier, which was demonstrated to be one of the most powerful techniques for EEG classification [13] using functions provided by Matlab software. To compare the accuracy of classification for different window lengths, we defined the window length of training and test data to be 250, 500, 750, and 1000 ms and then shifted the time windows in steps of 32 ms. We used a 32 ms time shift, which was the closest value to the 30 ms classifier update rate (used in Shindo *et al.* study), considering that the sampling frequency of the EEG data was 250 Hz. We used the entire subjects training data for classifier training and then tested its classification accuracy for all test data regardless of their correspondent subjects. We first used calibration data as well as training data for classifier training, hypothesizing that the larger the training dataset, the better the training of the classifier and consequently, the higher its classification accuracy would be. Thus, we only used the first four-second part of each 4.5-second-long trial, because we wanted to make them similar to the calibration data trials, which were only 4 seconds long. However, when we trained the classifier based on the mentioned hypothesis, it turned out that its average accuracy with 10-fold cross validation was very close to chance (53%). In search for the reason for these poor classification results, we noticed that calibration data, due to lack of feedback provision, might not be quite similar to training data in which subjects were provided with continuous feedback after 2 seconds.
In other words, while increasing the training dataset size may lead to improving classification accuracy, it only occurs when we provide homogenous data for classifier training.

Table. 2: Classification statistics for different window lengths

Window length (ms)	First accuracy (%)	Rise time (ms)	Maximum accuracy (%)	Last accuracy (%)	Mean accuracy	Standard deviation
250	**49**	**1000**	**67**	**60**	**61.09**	**3.8**
500	**54**	**900**	**71**	**62**	**64.06**	**3.3**
750	**61**	**1000**	**72.3**	**62**	**66.45**	**2.7**
1000	**65**	**750**	**72.9**	**64**	**68.16**	**2.5**

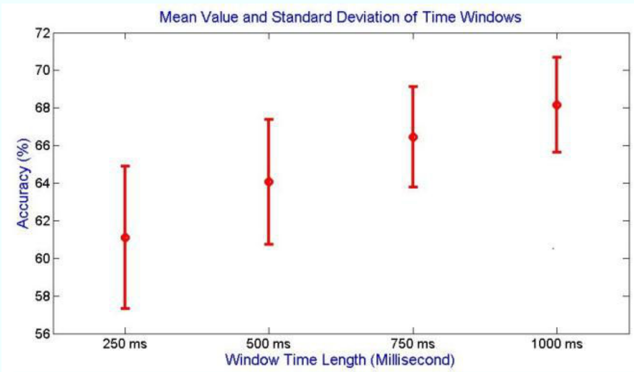

Figure 2. Mean and standard deviation for different time window

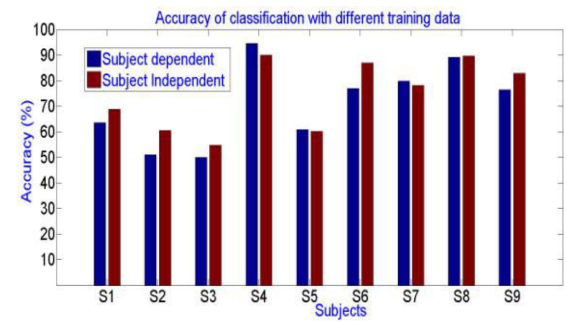

Figure 3. Comparison of subject-dependent and subject-independent training methods for subjects S1 - S9

Thus, we decided to use only training data and excluded calibration data for classifier training based on the aforementioned findings. By doing so, we limited the training data to only the most similar ones to test data, to improve the classifier accuracy.

Regarding comparison between subject-dependent and subject-independent classifier training, for the former we used only each subject's own training data (around 120 trials) for classifier training and tested its accuracy with the same subject's test data. However, in the latter we used the entire subjects training data (1182 trials) and then tested its accuracy for each subject test data.

III. Results

Based on the described methodology, we achieved the results summarized in Fig. 1. The accuracy level for time windows 250 and 500 ms started with classification accuracy very close to chance, however, for wider time windows (750, 1000 ms), the classifiers start with accuracies higher than 60%. Moreover, for all time windows, after around a second the classification reaches its maximum and that maximum level has a direct relationship with the length of the time window. Fig. 2 presents a comparison of the mean values and standard deviations for different time windows that show the direct correlation between time window length and corresponding mean accuracy values and an inverse correlation between time window length and corresponding accuracy values' standard deviation. Moreover, considering the highest accuracy of around 75% that was achieved by exploiting the whole 4-second-long time window for training, the mean accuracy achievable by 750 ms and 1000 ms time windows seems to be reasonable.

Regarding the feasibility of a subject-independent classifier, we tested both subject-dependent and subject-independent methods and results of the experiments are shown in Fig 3.

There it can be seen that training the classifier with the entire subjects training data (subject-independent method) improves the accuracy of classifier for 6 subjects (S1, S2, S3, S6, S8, S9) while degrades it for the other three (S4, S5, S7). However, it is worth mentioning that the subject dependent data provides very close to chance level accuracies for subjects S2, S3, while subject independent classifier provides higher classification accuracies for them. On the other hand, subject dependent classifier creates better results for subject S4 that has the highest accuracy level for both training methods.

IV. Discussion

The main finding of our study is that whenever a part or whole of the time windows are located in the first 500 ms of trials, accuracy is lower compared to the rest of the times as demonstrated in Fig. 1. This finding is in conformity with the literature [14-16], showing that the EEG data recorded during the first 250-500 ms of motor imagery trials do not include discriminant information. In terms of neurophysiology this delay in modulation of motor imagery may reflect the few hundred milliseconds of time that subjects need to recognize the visual stimulus followed by motor planning.

Moreover, the trend of all time windows shows that after reaching a maximum value that varies between 750 to 1000 ms for different time windows, the level of accuracy starts degrading gradually until the end of trial. This behavior might reflect the decrease in subject attention to the task as the trial goes on. In addition, it might be concluded that based on the accuracies provided by 250 ms time windows, the most discriminating time window is the 750-1750 ms period after the cue onset as this time period provides the highest accuracy level.

Considering the optimum choice for time window length while taking into account the trade-off between its length, and its transient and stable accuracy levels, it seems that the 750 ms long time window could be considered as the optimum choice since it starts with 61% accuracy and ends with 62% while its mean value (66.45%) is only 1.71% less than the average accuracy level of the 1000 ms window. In addition, the 750 ms time window allows the exploitation of the above mentioned golden window of time (750– 1750 ms) in which the most discriminating features can be extracted.

Comparing the rehabilitation level of stroke patients of the studies by Buch *et al.* and Shindo *et al.*, it must be considered that patients in the Buch *et al.* group had a more severe impairment in their affected hand's finger movement (no residual finger movement) compared to the patients of the Shindo *et al.* study, who had some residual motor functions. Moreover, it needs to be considered that recovery in mild to moderate stroke patients is more likely than in more severely paralyzed stroke patients [17]. Nonetheless, the role of providing ten times faster feedback in the Shindo *et al.* study, which used a 30 ms classification update rate compared to Buch *et al*. study, which used a 300 ms classification update rate, cannot be ruled out and it could have played an important role in the higher recovery level in the Shindo *et al.* study.

In addition, based on [18] a key factor for choosing the time shift for any BCI system that needs to be considered is that its value has to be greater than the sum of data transfer time, signal processing time and application delay. Otherwise, the provided feedback cannot be considered to be real-time, which consequently degrades performance of the BCI system. It has been explicitly mentioned in [18] that assuming a 30 ms buffer length of the EEG amplifier, it is far greater than the required time for data transfer, signal processing and application delay in total, when recording data with 4 channels. This time (30 ms) is very close to the 32 ms that we used in this study as window shifting width. These authors used a typical home PC configuration with a 16 channel gUSB amp EEG amplifier and open source BCI2000 as their software platform. Since their experiment design did not include any specialized hardware, considering the accessibility of the hardware and software to any BCI laboratory our suggested update rate seems to be feasible and practical.

It has been demonstrated in this study that using subject independent training method, it is possible to improve the classification accuracy level of the subjects with low accuracy levels. However, it would be only possible in cost of decreasing the accuracy level of subjects who could have achieved higher accuracies with subject dependent method. In addition, caution must be taken when applying this strategy to stroke patients. It may work only for the groups of stroke patients who have similar lesions and similar motor imagery capabilities.

V. Conclusion

In this study we have shown that 750 ms long time windows with classification accuracy update rate of 32 ms and average classification accuracy of 67% results in a viable design for restorative BCI applications. This design would provide patients with fast feedback that is likely to be critical for Hebbian learning and leads to greater functional recovery of stroke patients [8].

Acknowledgment

We would like to thank the Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology for providing the EEG Data.

REFERENCES:

[1] D. M. Nilsen, G. Gillen, and A. M. Gordon, Use of mental practice to improve upper-limb recovery after stroke: a systematic review, *Am J Occup Ther,* vol. 64, pp. 695-708, 2010.

[2] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, A clinical study of motor imagery-based brain-computer interface for upper limb robotic rehabilitation, *Conf Proc IEEE Eng Med Biol Soc,* pp. 5981-5984, 2009.

[3] K. K. Ang. (2011, *A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface.*

[4] E. Buch, C. Weber, L. G. Cohen, C. Braun, M. A. Dimyan, T. Ard, J. Mellinger, A. Caria, S. Soekadar, A. Fourkas, and N. Birbaumer, Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke, *Stroke,* vol. 39, pp. 910-917, 2008.

[5] M. Gomez-Rodriguez, M. Grosse-Wentrup, J. Hill, A. Gharabaghi, B. Scholkopf, and J. Peters, Towards brain-robot interfaces in stroke rehabilitation, *IEEE Int Conf Rehabil Robot,* 2011.

[6] K. Shindo, K. Kawashima, J. Ushiba, N. Ota, M. Ito, T. Ota, A. Kimura, and M. Liu, Effects of neurofeedback training with an electroencephalogram-based brain-computer interface for hand paralysis in patients with chronic stroke: a preliminary case series study, *J Rehabil Med,* vol. 43, pp. 951-957, 2011.

[7] S. Soekadar, N. Birbaumer, and L. Cohen, Brain–Computer Interfaces in the Rehabilitation of Stroke and Neurotrauma, in *Systems Neuroscience and Rehabilitation*, K. Kansaku and L. Cohen, Eds., ed: Springer Japan, 2011, pp. 3-18.

[8] T. H. Murphy and D. Corbett, Plasticity during stroke recovery: from synapse to behaviour, *Nat Rev Neurosci,* vol. 10, pp. 861-872, 2009.

[9] M. Grosse-Wentrup, D. Mattia, and K. Oweiss, Using brain-computer interfaces to induce neural plasticity and restore function, *J Neural Eng,* vol. 8, 2011.

[10] G. Pfurtscheller and F. H. Lopes da Silva, Event-related EEG/MEG synchronization and desynchronization: basic principles, *Clin Neurophysiol,* vol. 110, pp. 1842-1857, 1999.

[11] V. Bostanov and B. Kotchoubey, The t-CWT: a new ERP detection and quantification method based on the continuous wavelet transform and Student's t-statistics, *Clin Neurophysiol,* vol. 117, pp. 2627-2644, 2006.

[12] S. Darvishi and A. Al-Ani, Brain-computer interface analysis using continuous wavelet transform and adaptive neuro-fuzzy classifier, *Conf Proc IEEE Eng Med Biol Soc,* pp. 3220-3223, 2007.

[13] S. Darvishi, Moradi, MH., Baumert, M., Abbott, D., Classification of motor imagery EEG signals using adaptive neuro-fuzzy inference system trained by particle swarm optimization, EMBC 2012, San Diego, 2012.

[14] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss, Beamforming in noninvasive brain-computer interfaces, *IEEE Trans Biomed Eng,* vol. 56, pp. 1209-1219, 2009.

[15] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlogl, B. Obermaier, and M. Pregenzer, Current trends in Graz Brain-Computer Interface (BCI) research, *IEEE Trans Rehabil Eng,* vol. 8, pp. 216-219, 2000.

[16] G. Pfurtscheller, C. Neuper, H. Ramoser, and J. Muller-Gerking, Visually guided motor imagery activates sensorimotor areas in humans, *Neurosci Lett,* vol. 269, pp. 153-156, 1999.

[17] L. Oujamaa, I. Relave, J. Froger, D. Mottet, and J. Y. Pelissier, Rehabilitation of arm function after stroke. Literature review, *Ann Phys Rehabil Med,* vol. 52, pp. 269-293, 2009.

[18] J. A. Wilson, C. Guger, and G. Schalk, BCI hardware and software, in *Brain-computer interfaces : principles and practice*, J. R. Wolpaw and E. W. Wolpaw, Eds., ed Oxford; New York: Oxford University Press, 2012, pp. 165-188.