Motor Imagery Signal Classification Using Constant-Q Transform for BCI Applications

Mustafa Alper BALIM Electrical and Electronics Engineering Bursa Technical University Bursa, Turkey mustafaalper.balim@btu.edu.tr Cemal HANİLÇİ Electrical and Electronics Engineering Bursa Technical University Bursa, Turkey cemal.hanilci@btu.edu.tr Nurettin ACIR Electronics Engineering National Defence University İstanbul, Turkey nacir@hho.edu.tr

Abstract—Electroencephalography (EEG) signals have been using for brain-computer interface applications for the last two decades. Motor imagery (MI) signals are one of the EEG signal types formed by imagining a limb's movement. Recently with the help of deep neural networks (DNN) for classifying MI signals using time-frequency (TF) features, considerable performance improvement has been reported. This paper proposes using a well-known TF representation technique called Constant-Q Transform (CQT) for the MI signal classification. Experiments conducted on BCI IV 2b dataset with DNN classifier using CQT spectrogram show that CQT outperforms traditional short-time Fourier transform (STFT) representation.

Index Terms—brain-computer interface, motor imagery, electroencephalography, constant-Q transform

I. INTRODUCTION

Brain-Computer Interface (BCI) applications continue to develop to become a powerful tool to make easier communication between humans and machines. It is also a promising phenomenon for individuals with neuromuscular disorders, making it possible to control machines without using neuromuscular pathways.

Electroencephalography (EEG) signals measured from the skull surface are commonly used for BCI applications [1]. Motor imagery (MI) signals are one of the EEG signal representations widely used in BCI studies. MI signals are considered as an event-related desynchronization (ERD) and event-related synchronization (ERS) pattern that occurs within the EEG signal, similar to the patterns observed under the presence of motor activity when imagining to move a limb (e.g. hands, feet and tongue) [2]. An EEG signal consists of five sub-frequency bands; gamma (>30 Hz), beta (12-30 Hz), alpha (8-12 Hz), theta (4-7 Hz) and delta (<4 Hz) waves. ERD and ERS patterns are generally observed in alpha and beta bands [3], [4].

Classification of MI signals is a typical pattern recognition task consisting of two phases: *feature extraction* and *classification*. Since EEG signals are weak, non-stationary, aperiodic and time-varying complex signals, a special attention is required while processing them for feature extraction [5]. Commonly used feature extraction methods for MI signal classification are time-domain features such as empirical mode decomposition (EMD) [6] and root-mean squares (RMS) [7], frequency-domain features extracted using fast Fourier transform (FFT) [8] and power spectral density (PSD) of the signal [9], spatial filtering features such as common spatial patterns (CSP) [10] and TF domain features obtained from wavelet transform (WT) [11] and STFT [12]–[14]. For the classification, in turn, principal component analysis (PCA) [5], support vector machines (SVM) [9], linear discriminant analysis (LDA) [9] and K-nearest neighbors (kNN) [6] methods are widely used. Besides traditional classifiers, deep learning frameworks such as convolutional neural networks (CNN) [12], [14], [15], deep belief networks (DBN) [5], [8] and long short-term memory (LSTM) [16] have become popular and promising results were reported [3]–[5], [17], [18]. Classifying MI signals using TF features with CNN architecture is the most common method [17].

Several studies used features extracted using STFT for MI signal classification. In [15], a CNN combined with stacked auto-encoders (SAE) was proposed using the STFT features extracted from two seconds long EEG signals. In that work, authors extracted the alpha (8-12 Hz) and beta (12-30 Hz) rhythms from the STFT representation and then the features of each EEG channel was concatenated to obtain a single input image for each signal. The combined input image was then used with CNN and 77.6% average accuracy was obtained for all subjects in BCI IV 2b dataset. In [16], STFT was used to extract features after filtering two seconds long EEG signals with a bandpass filter (5-30 Hz). The features were extracted from the desired frequency bands (8-30 Hz) using the STFT, resulting in a feature matrix of 7×15 dimension for each channel. Using the concatenated feature matrices as the input of the CNN and LSTM classifiers were compared, and it was shown that CNN is superior to LSTM. In [13], 14×14 features were extracted for each channel from the signal of duration two seconds signal by STFT. Then evaluated with several deep learning frameworks and it was shown that CapsNet based framework gives the highest classification accuracy of 78.44% on BCI IV 2b dataset. In [11], both STFT and continuous wavelet transform (CWT) were used for feature extraction, and CNN was used for classification. It was concluded that both STFT and CWT equally work well on classifying MI signals while CWT slightly (+1.17%) performs better than STFT.

Taking the promising results obtained using STFT TF

representation into account, in this paper, we propose to use the CQT to obtain the TF representation for EEG signal classification. The advantage of the CQT over discrete Fourier transform (DFT) is the fact that while DFT provides a fixed frequency resolution, CQT ensures higher frequency resolution in low-frequency bands by employing geometrically spaced frequency bins. Thus, intuitively, extracting TF features using a variable frequency resolution would be more informative for MI signal classification. STFT and CQT spectrograms are used with a CNN-based classifier and we compare their performances on BCI IV 2b dataset.

II. TIME-FREQUENCY REPRESENTATION

Spectral brain activity varies over time when different MI tasks are carried out. Although EEG signals are one-dimensional arrays, they can be represented as twodimensional images by utilizing WT, STFT and CQT. Such representation makes it possible to observe energy variation over time during MI tasks [19]. In this section, we briefly describe the well-known STFT and the proposed CQT methods.

A. Short-time Fourier Transform

Since EEG is a non-stationary signal, its spectral content changes over time. Hence, applying DFT to EEG signal over a long window does not reveal the spectral changes over time. In order to avoid this, DFT is applied over short segments of the signal. STFT is a popular technique based on DFT which helps to analyze the energy variation of a signal over time. To compute the STFT of a discrete-time signal x[n], the signal is first divided into short overlapping segments. Each segment is multiplied by a window function (generally a data tapering window is used). This results in a set of windowed frames, and STFT is defined as the DFT of the windowed segments:

$$X(m, k) = \sum_{m=-\infty}^{+\infty} x[m]w[n-m]e^{-j\frac{2\pi}{N}km}$$
(1)

where w[n] is the window function consisting of a total of N samples and the sequence x[m]w[n-m] represents the short segment of the EEG signal x[n]. The STFT of the signal X(m, k) is therefore a function of two variables where m is the time variable (generally represents the frame index) and k is the discrete-frequency index. Hence STFT is a TF representation of the signal, and its graphical visualization is known as the spectrogram of the signal.

From the above definition, STFT can be considered as a filterbank where the window function w[n] can be thought as the impulse response of the filter. So that the signal is first passed through the filter with linear phase factor $(w[n]e^{-j\frac{2\pi}{N}k}, k = 0, 1, ..., N - 1)$ and then the output of the each filter is demodulated by $e^{-j\frac{2\pi}{N}k}$. The Q factor is a metric which measures the selectivity of each filter $(w[n]e^{-j\frac{2\pi}{N}k})$ and it is defined as:

$$Q = \frac{f_c}{\delta f} \tag{2}$$

where f_c and δf are the center frequency and bandwidth of each filter, respectively. Since the bandwidth of each filter

 $(w[n]e^{-j\frac{2\pi}{N}k})$ is constant in the STFT, the Q factor is therefore an increasing function of the frequency.

B. Constant Q Transform

Although STFT is the most popular TF representation, one of the main drawbacks of the STFT is that it increases Qfactor with frequency. In order to cope with that, another TF analysis known as the CQT was proposed in [20]. It was originally developed to process music signals since it provides perceptually better representation according to the human perception system. Since speech and EEG signals have very similar characteristics (non-stationary, aperiodic and varying in time), CQT is potentially a good candidate for EEG signal analysis.

The motivation behind the CQT is to obtain a TF feature with a constant Q factor along the entire frequency axis by spacing center frequencies of the filters geometrically. Thus, it can obtain a higher frequency resolution at lower frequencies and a higher temporal resolution at higher frequencies [21]. Given a discrete-time signal x[n], the CQT is defined as:

$$X_{CQ}(k,n) = \sum_{i=0}^{N_k-1} x[i]\alpha_k^* \left[i - n + \frac{N_k}{2}\right]$$
(3)

where k is the frequency bin, N_k is the window length of the kth bin and $\alpha_k^*[n]$ is the complex conjugate of the complex basis function for the bin k which is defined by:

$$\alpha_k[n] = \frac{1}{N_k} w \left[\frac{n}{N_k} \right] \exp\left[-j2\pi n \frac{f_k}{f_s} \right]$$
(4)

Here f_k is the center frequency of the kth bin, f_s is the sampling frequency and w[n] is the window function. The center frequency values are computed by:

$$f_k = f_1 2^{\frac{\kappa - 1}{B}} \tag{5}$$

where f_1 is the center frequency of the first bin and B corresponds to the number of bins per octave. For more details on the computation of CQT, the readers are referred to [20].

From the above definitions, it is observed that the analysis window length (N_k) and center frequency values (f_k) vary with the frequency values in contrast to traditional STFT. Thus, CQT yields a constant Q value for all frequency values. Fig. 1 shows TF representations of an example EEG signal obtained using STFT and CQT techniques. From the figure, although the largest energy variation around 10 Hz can be observed from both representations, STFT spectrogram rapidly attenuates above approximately 20 Hz. This is probably because of the fact that the spectral resolution of the STFT is considerably high to detect small variations. Although the frequency resolution of the STFT can be improved by using a larger window, it clearly reduces the time resolution which introduces a trade-off. In contrast to STFT, CQT spectrogram clearly shows the energy variations at low-frequency values.

One would argue that observing STFT and CQT spectrograms of a single EEG signal can not be generalized well to show the effectiveness of the CQT over STFT. To this end



Fig. 1. Comparison of STFT and CQT spectrograms



Fig. 2. Long-term spectra obtained from the STFT and CQT

we compute the long-term average STFT and CQT spectra using 740 EEG signals from a subject for C3 channel. More specifically STFT and CQT spectrograms are averaged over all frames and all signals by:

$$LTAS(k) = \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{N} \log |X_t(n,k)|$$
(6)

where $X_t(n, k)$ correspond to STFT or CQT spectrograms of the tth signal, respectively and T is the total number of EEG signals used to compute average spectra. Fig. 2 shows the long-term spectra obtained from the STFT and CQT for class 1 (left hand) and class 2 (right hand). From the figure, it can clearly be seen that STFT provides a smooth spectra. However, a more detailed spectral variation can be observed by CQT. Hence, intuitively CQT will provide more detailed information about the EEG signals.

III. EXPERIMENTAL SETUP

A. Dataset and Pre-processing

The dataset 2b from BCI IV competition [22] is used in the experiments. It consists of EEG signals collected from nine right-handed subjects. For each subject, the database consists of 3 channels (C3, Cz, C4) EEG signals sampled at 250 Hz which are collected in five sessions. An 50 Hz notch filter and

a band-pass filter with 0.5-100 Hz pass band were applied to the signals. Total number of EEG signal for each subject varies between 680 and 760. Since the MI signals are predominantly alpha and beta waves, a 5-50 Hz band-pass filter has been applied before feature extraction.

B. Feature Extraction

In this study, both STFT and CQT spectrograms used as features. It was shown in [4], [12], [15] that observed energy in the motor cortex area in alpha (8-12 Hz) and beta (12-30 Hz) bands vary while performing MI task. Therefore, TF representations are widely used for EEG signal classification. As mentioned in the preceding section, STFT spectrogram does not properly reveals the detailed information because of the fixed frequency resolution. Hence, in this paper, we propose to use CQT spectrograms as the TF representation.

To compute the STFT spectrogram, each signal is first divided into short frames consisting of 32 samples with 8 sample frame shift (24 samples of consecutive frames overlap). Then each frame was windowed using a Hamming window. 256 point DFT of each windowed frame was computed to obtain the spectrum of each frame. Due to the symmetry property of the DFT, the first 129 samples were retained. This results in a spectrogram image of size 129×122 . In the previous studies, it was shown that 6-36 Hz frequency band contains the most discriminative information for EEG signal classification [4], [12], [15]. Similar observation was found in our preliminary experiments. Therefore; only 6-36 Hz frequency band was selected from the spectrogram images rather than using the full 0-125 Hz band. This yields a feature matrix of size 30×122 . Since each EEG signal of each subject in BCI IV 2b dataset composed from three channels (C3, Cz, C4), each channel was proposed independently and resulting features were combined. Hence for each trial in the database, a 90×122 dimensional STFT features were extracted. To compute the CQT spectrogram, we use short windows consisting of 8 samples and the number of frequency bins and the number of frequency bins per octave are selected as 30 and 12, respectively. A CQT spectrogram of dimension 30×126 was obtained using these parameters for each channel. The standardization, concatenation and scaling steps were applied similar to STFT, and a 90×126 COT feature matrix was obtained. The window lengths used to extract STFT and CQT spectra were determined according to preliminary experiments in order to obtain the TF images of the similar size for both methods. One could argue that using three dimensional TF spectra for each EEG signal as the input would perform better than combining the EEG channels. However, we found that combining the channels yields better performance than threedimensional representation.

C. Classifier

EEG signal classification was performed using the CNN model shown in Fig. 3. In order to make a fair performance comparison of the STFT and CQT spectrograms, the same



Fig. 3. The proposed method and the details of the CNN-based classifier used in the experiments.

network is used for both features. The parameters at the network (number of filters, size of the filters, activation function etc.) were optimized by preliminary experiments. As shown in Fig. 3, the proposed network consists of two convolutional layers where the first convolutional block contains 30 filters of dimension 15x1 followed by a max-pooling layer. ReLU non-linearity was used in the convolutional layers. The output layer consisting of two nodes which performs the classification task while the first output node corresponds to the first class (the class that represents the subject imagines moving his/her left hand), the second output unit represents the second class (subject imagines moving his/her right hand). Two different classification scenarios were considered in the experiments:

- Intra-subject classification: A separate model was trained for each subject.
- Inter-subject classification: A single model was trained for all subjects.

In the previous studies, generally intra-subject approach was used and it was shown to outperform inter-subject method, as expected. Although inter-subject approach would be more preferable since a single generic model is trained for all subjects, it is inferior to the intra-subject method because of the fact that EEG signal specifications are highly dependent on the subjects.

Stochastic gradient descent algorithm was used to train the network with an initial learning rate of 0.01 and it was reduced to 0.005 after 60 epochs. Maximum number of epochs was selected as 400 and if validation loss was not reduced for 25 successive epochs, training was terminated. Categorical crossentropy was used as the loss function.

Classification experiments were carried out using 10-fold cross-validation approach where EEG signals of each subject is divided into ten disjoint subsets and nine subsets were used to train the system and the remaining one subset was used for testing and this is repeated for ten times. With this, all of the signals of each subject are used to test the system. In each fold, 20% of the training subset is used as validation set to optimize the network parameters during training. For the intersubject approach, in each fold, training subset of all subjects were combined and a single network was trained.

D. Performance Metrics

In the experiments, the average classification accuracy averaged over all ten-folds is used as the primary performance criterion. Besides the average accuracy, the standard deviation of the classification accuracies was calculated and reported in the experiments.

IV. RESULTS

The average classification accuracy values and the standard deviations obtained with intra-subject classification approach using CQT and STFT features are summarized in Table I. From the table, CQT features systematically outperform STFT for the majority of the subjects. For example, while STFT spectrogram yields a 71.2% accuracy for the subject 1, CQT representation gives 79.7% accuracy which corresponds to a 8.5% relative improvement over STFT. This shows CQT is significantly superior to the STFT. The last row of the Table I shows the accuracy and standard deviation values averaged over all subject and again CQT gives better values. One interesting observation from the results given in Table I is that although reasonable accuracy values are obtained for the vast majority of the subjects, considerably lower accuracy values are observed for the subjects 2 and 3 for both feature representations. Similar observations were reported independently in the previous studies [5], [23]. This suggests that subjects 2 and 3 are unsuccessful in imagining movement. Hence these two subjects make the average accuracy value biased towards them.

In the inter-subject classification experiments approach CQT again systematically outperforms the STFT irrespective of the subjects (except subject 1). Interestingly, inter-subject approach slightly outperforms the intra-subject classification on average when CQT features are used. This shows the effectiveness of the CQT on motor imagery signal classification. However, the average standard deviation value is considerably lower than intra-subject approach.

Finally, we compare the results we obtain in this study with the previous works that used the same BCI IV 2b database in Table II. Although some of the studies used a different experimental setup such as using only the first three sessions rather than all five sessions [8], [14], [15] or a different

	Intra-subject				Inter-subject			
	CQT		STFT		CQT		STFT	
Sub.	Acc.	S.D.	Acc.	<i>S.D.</i>	Acc.	<i>S.D.</i>	Acc.	<i>S.D.</i>
1	79.7	5.0	71.2	4.6	69.9	6.0	70.8	4.9
2	55.6	8.7	55.6	5.9	62.9	1.3	60.9	0.8
3	52.7	5.3	56.6	6.1	60.5	1.4	57.6	0.5
4	95.1	1.7	94.4	2.4	92.1	0.9	90.1	0.6
5	90.3	2.7	85.2	4.7	82.1	1.0	80.9	1.1
6	80.6	5.1	74.6	4.9	82.9	1.0	77.1	0.7
7	77.6	4.5	72.5	2.7	79.6	1.3	71.9	1.0
8	80.5	5.2	80.1	5.0	80.5	1.1	77.9	1.0
9	79.4	5.1	79.0	4.5	85.7	0.6	81.3	1.0
Avg.	76.8	4.8	74.4	4.5	77.3	1.6	74.3	1.3

 TABLE I

 CLASSIFICATION RESULTS USING STFT AND CQT FEATURES

classifier, we compare the overall accuracy values in order to get an idea about whether the performance of the CQT can be further improved. From the table, we observe that although the classifiers and the experimental setups are different, CQT outperforms the results reported in some studies. This indicates that the performance of the proposed CQT method can be further improved by using a different model such as RBM and using only first three sessions as in [8].

 TABLE II

 Comparison of Recent Studies Evaluated on BCI IV 2b

Study	Feature	Classifier	Avg. Acc.
[8]	FFT	RBM	84.0
[10]	FB-CSP	GRU-RNN	82.75
[12]	STFT	CNN	73.81
[13]	STFT	Caps.Net	73.81
[14]	STFT	CNN-VAE	78,2
[15]	STFT	CNN-SAE	77.6
This study	STFT	CNN	74.4
This study	CQT	CNN	76.8

V. CONCLUSIONS

In this study, we proposed to use CQT as a TF feature extraction method instead of STFT for motor imagery signal classification. The motivation behind using CQT features instead of STFT features was lower frequency components (alpha and beta waves) of EEG signals are more discriminative than higher frequency bands, and CQT provides higher frequency resolution in low-frequency bands [2]. Experiments conducted on BCI IV 2b dataset using CNN classifier showed that CQT representation systematically outperforms the traditional STFT features. It was found that although inter-subject approach using CQT features slightly increases the average classification accuracy in comparison to intra-subject method, the average standard deviation value is considerably lower than the intra-subject approach. This reveals the generalization capability of the CQT features. We showed that CQT achieves comparable performance with the previous studies. Future studies using more powerful classifiers such as RBM and RNN can further improve the performance of the CQT features.

REFERENCES

- T. Nguyen, I. Hettiarachchi, A. Khatami, L. Gordon-Brown, C. P. Lim, and S. Nahavandi, "Classification of Multi-Class BCI Data by Common Spatial Pattern and Fuzzy System," IEEE Access, vol. 6, pp. 27873–27884, 2018.
- [2] G. Pfurtscheller and C. Neuper, "Motor imagery and direct braincomputer communication," Proceedings of the IEEE, vol. 89, no. 7, pp. 1123–1134, 2001.
- [3] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain–computer interface paradigms," Journal of Neural Eng., vol. 16, no. 1, p. 011001, 2019.
- [4] N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges," Sensors, vol. 19, no. 6, p. 1423, 2019.
- [5] L. Cheng, D. Li, G. Yu, Z. Zhang, X. Li, and S. Yu, "A Motor Imagery EEG Feature Extraction Method Based on Energy Principal Component Analysis and Deep Belief Networks," IEEE Access, vol. 8, pp. 21453–21472, 2020.
- [6] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," Biomed. Signal Process. and Control, vol. 31, pp. 398–406, Jan. 2017.
- [7] M. Hamedi, S. Salleh, A. M. Noor and I. Mohammad-Rezazadeh, "Neural network-based three-class motor imagery classification using time-domain features for BCI applications," 2014 IEEE Region 10 Symposium, Kuala Lumpur, Malaysia, 2014, pp. 204-207.
 [8] N. Lu, T. Li, X. Ren and H. Miao, "A Deep Learning Scheme for Motor
- [8] N. Lu, T. Li, X. Ren and H. Miao, "A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines," IEEE Trans. on Neural Sys. and Reh. Eng., vol. 25, no. 6, pp. 566-576, 2017.
- [9] M. Zheng, B. Yang, and Y. Xie, "EEG classification across sessions and across subjects through transfer learning in motor imagery-based brainmachine interface system," Med. Biol. Eng. Comput., vol. 58, no. 7, pp. 1515–1528, Jul. 2020.
- [10] T. Luo, C. Zhou, and F. Chao, "Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network," BMC Bioinformatics, vol. 19, no. 1, p. 344, 2018.
- [11] S. Chaudhary, S. Taran, V. Bajaj and A. Sengur, "Convolutional Neural Network Based Approach Towards Motor Imagery Tasks EEG Signals Classification," IEEE Sensors, vol. 19, no. 12, pp. 4494-4500, 15 June15, 2019.
- [12] G. Xu et al., "A Deep Transfer Convolutional Neural Network Framework for EEG Signal Classification," IEEE Access, vol. 7, pp. 112767–112776, 2019.
- [13] K.-W. Ha and J.-W. Jeong, "Motor Imagery EEG Classification Using Capsule Networks," Sensors, vol. 19, no. 13, p. 2854, 2019.
- [14] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "EEG Classification of Motor Imagery Using a Novel Deep Learning Framework," Sensors, vol. 19, no. 3, p. 551, 2019.
- [15] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," Journal of Neural Eng., vol. 14, no. 1, p. 016003, 2016.
- [16] Z. Wang, L. Cao, Z. Zhang, X. Gong, Y. Sun, and H. Wang, "Short time Fourier transformation and deep neural networks for motor imagery brain computer interface recognition," Concurrency Comp. Pract. and Exp., vol. 30, no. 23, p. e4413, 2018.
- [17] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," J. of Neural Eng., vol. 16, no. 3, p. 031001, 2019.
- [18] S. Aggarwal and N. Chugh, "Signal processing techniques for motor imagery brain computer interface: A review," Array, vol. 1–2, p. 100003, 2019.
- [19] D. Wang, D. Miao, and G. Blohm, "Multi-Class Motor Imagery EEG Decoding for Brain-Computer Interfaces," Frontiers in Neuroscience, vol. 6, p. 151, 2012.
- [20] J. Youngberg and S. Boll, "Constant-Q signal analysis and synthesis," ICASSP '78. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Tulsa, OK, USA, 1978, pp. 375-378.
- [21] M. Todisco, H. Delgado, and N. Evans, "A New Feature for Automatic Speaker Verification Anti-Spoofing: Constant Q Cepstral Coefficients," Odyssey pp. 283–290, 2016.
- [22] R. Leeb, C. Brunner, G. -Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI Competition 2008 – Graz data set B," BCI Competition IV, 2008.
- [23] M. Tangermann et al., "Review of the BCI Competition IV," Frontiers in Neuroscience, vol. 6, p. 55, 2012.