Research Article

Improving EEG-Based Motor Imagery Classification for Real-Time Applications Using the QSA Method

Patricia Batres-Mendoza,1 Mario A. Ibarra-Manzano,2,3 Erick I. Guerra-Hernandez,1 Dora L. Almanza-Ojeda,2,3 Carlos R. Montoro-Sanjose,3,4 Rene J. Romero-Troncoso,3,5 and Horacio Rostro-Gonzalez1,3,6

1Laboratorio de Sistemas Bioinspirados, Departamento de Ingeniería Electrónica, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago Km. 3.5 + 1.8 Km., 36885 Salamanca, GTO, Mexico
2Laboratorio de Procesamiento Digital de Señales, Departamento de Ingeniería Electrónica, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago Km. 3.5 + 1.8 Km., 36885 Salamanca, GTO, Mexico
3Cuerpo Académico de Telemática, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago Km. 3.5 + 1.8 Km., 36885 Salamanca, GTO, Mexico
4Departamento de Arte y Empresa, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago Km. 3.5 + 1.8 Km., 36885 Salamanca, GTO, Mexico
5Departamento de Ingeniería Electrónica, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago Km. 3.5 + 1.8 Km., 36885 Salamanca, GTO, Mexico
6Neuroscientific System Theory, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany

Correspondence should be addressed to Horacio Rostro-Gonzalez; hrostrog@ugto.mx

Received 14 July 2017; Revised 16 October 2017; Accepted 9 November 2017; Published 3 December 2017

Academic Editor: Pedro Antonio Gutierrez

Copyright © 2017 Patricia Batres-Mendoza et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present an improvement to the quaternion-based signal analysis (QSA) technique to extract electroencephalography (EEG) signal features with a view to developing real-time applications, particularly in motor imagery (MI) cognitive processes. The proposed methodology (iQSA, improved QSA) extracts features such as the average, variance, homogeneity, and contrast of EEG signals related to motor imagery in a more efficient manner (i.e., by reducing the number of samples needed to classify the signal and improving the classification percentage) compared to the original QSA technique. Specifically, we can sample the signal in variable time periods (from 0.5 s to 3 s, in half-a-second intervals) to determine the relationship between the number of samples and their effectiveness in classifying signals. In addition, to strengthen the classification process a number of boosting-technique-based decision trees were implemented. The results show an 82.30% accuracy rate for 0.5 s samples and 73.16% for 3 s samples. This is a significant improvement compared to the original QSA technique that offered results from 33.31% to 40.82% without sampling window and from 33.44% to 41.07% with sampling window, respectively. We can thus conclude that iQSA is better suited to develop real-time applications.

1. Introduction

In the last few years, interest in inferring information from the human brain stemming from cognitive thoughts by means of electroencephalography (EEG) has expanded to various disciplines such as neuroscience, robotics, computational science, physics, and mathematics. Research in these areas tends to revolve around the development of new communication and control technologies based on brain-computer interface (BCI) devices to support people with severe neuromuscular conditions in ways that can enable them to express their wishes or use devices as neuroprosthetics [1], wheelchairs [2, 3], control a cursor on a computer screen [4], or even a robot [5, 6]. Wolpaw et al. [7] argue that BCIs “give their users communication and control channels that do not depend on the brain’s normal output
Motor imagery (MI) is a conscious process defined as a mental simulation of a particular movement [9]. The motor imagery is endowed with the same functional relationship to the imagined or represented movement and the same causal role in the generation of the movement in question [10]. In other words, MI is related to the intention and preparation of movements, whereby the subject imagines carrying out a particular action without making any real movements. This has led to studies using motor imagery to decipher processes that precede the execution of an action. For instance, Bai et al. [11] claim that mental practice using motor imagery of limb movement may facilitate motor recovery in persons who have experienced cerebrovascular injuries. In addition, McFarland et al. [12] conducted a comparative analysis of EEG topographies associated with actual hand movements and imagined hand movements, concluding that motor imagery plays an important role in EEG-based communication and suggesting that mu and beta rhythms might provide independent control signals.

Similarly, there have been studies focusing on MI support and BCI systems, proposing algorithms for feature extraction and classification. For instance, Pfurtscheller et al. [13] studied the reactivity of mu rhythms associated with the imagination of hand, foot, and tongue movements with 60 EEG electrodes in nine able-bodied subjects (with a 66.16% performance rate). In turn, Aghaei et al. [14] argue for the use of the separable common spatiotemporal patterns (SCSSP) method to extract discriminant spatiotemporal EEG features and a Laplacian filter of data set V of BCI competition III with the following mental imagery tasks: left-hand movement, right-hand movement, and generation of words beginning with a random letter involving 3 subjects. As for Gao et al. [15], they conducted EEG signal analysis during left-hand movements, right-hand movements, and resting with 10 subjects using the Kolmogorov complexity to extract the features and an AdaBoost multiclass classifier, achieving a 79.5% accuracy rate. In a separate study, Schlögl et al. [16] conducted a comparative study involving four classifiers to determine the global separability of data in relation to four different MI tasks with 5 subjects, modelling the EEG signal by means of an adaptive autoregressive (AAR) process whose parameters were extracted through Kalman filtering. Trad et al. [17] used empirical mode decomposition (EMD) and band power (BP) to extract EEG signals and classify MI in experiments involving 10 subjects aged 22–35 as they imagined left-hand and right-hand movements. Choi and Cichocki [18] implemented a MI-based algorithm to control a wheelchair using spatial filters to extract the features by means of a common spatial pattern (CSP) method and the linear SVMs to classify feature vectors. Three healthy men participated in the experiments where they had to imagine clenching the right hand, squeezing the left hand, and walking. In [19] results of tests conducted with a 74% accuracy rate to control a robot indoors on the basis of three mental states are presented. In particular, there are several studies focusing on the development of processing techniques, feature extraction, and classification to improve the BCI systems. In Tables 1 and 2, we present the most commonly used algorithms for these tasks from Lotte’ et al. works [20]. However, we only extracted information related to our work, that is, motor imagery activities (imagined movement of the left hand, imagined movement of the right hand, imagined movement of the foot, imagined movement of the tongue, relaxation, and mental calculation). These works use different algorithms in the classification stages, such as SVM, KNN, LDA, MLP, HMM, Gaussian classifier, and Bayes quadratic, including combinations of these, which result in most studies in acceptable performance rates. However, most experiments are performed on a small number of subjects, which returns a low number of trials per session. In the feature extraction stage, most techniques used to analyze EEG signals extract information within the frequency or time-frequency domain, which may lead to information loss when information is transformed. In addition, they require noise elimination filters and frequency band localization to identify the patterns of motor imagery.

On the other hand, in [21], the design of a motor imagery experiment was reported based on three mental processes: arrow moving to the left, arrow moving to the right, and waiting time, lasting 5 seconds per image. This study analyzed EEG signals with the implementation of the quaternion-based signal analysis (QSA). The quaternion-based signal analysis (QSA) method is a technique that uses EEG signals within the time domain because it is based on quaternion algebra. The use of quaternion algebra with the QSA technique makes it possible to describe signals within the time domain by means of rotations and orientations of 3D objects and represent multichannel EEG signals as a single entity, preventing data ambiguity and producing a more accurate representation doing fewer calculations than are needed with other techniques. The offline analysis was conducted in the feature extraction phase considering the total number of samples in each class, which produced an 84.92% accuracy rate. However, while the analysis of offline signals is convenient and efficient, offline analysis results may not generalize their performance to online applications. In the case of the QSA method, the online analysis obtained was just 33.31% considering window sizes of 0.5 seconds.

Thus, this paper presents an improvement to the QSA method that we shall call improved quaternion-based signal analysis (iQSA) for use in the feature extraction and classification phases, whose contribution consists in providing a technique for use in real-time applications, focusing on analyzing EEG signals online reducing the sample sizes needed to a tenth of the ones required by QSA, resulting in a faster response and fewer delays to improve execution times in real-time actions. The experiment involved using an Emotiv-Epoc device to acquire the brain signals from the motor and visual regions of the cerebral cortex. Similarly, during the training and validation phases, the EEG signal
Table 1: Accuracy of classifier in motor imagery based BCI: multiclass. The classes are (T1) left imagined hand movements, (T2) right imagined hand movements, (T3) imagined foot movements, (T4) imagined tongue movements, (T5) relax (baseline), (T6) word generation starting with a letter specific, and (T7) mental calculation.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Activity</th>
<th>Subjects</th>
<th>Trials</th>
<th>Filter</th>
<th>Feature</th>
<th>Classifier</th>
<th>Accuracy</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCI competition III</td>
<td>T1, T2, T3, T4</td>
<td>5</td>
<td>240 to 360</td>
<td>Yes</td>
<td>AAR</td>
<td>Linear SVM, KNN, LDA, Mahalanobis distance</td>
<td>63%</td>
<td>[16]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T6</td>
<td>3</td>
<td>Yes</td>
<td></td>
<td></td>
<td>Neuro-fuzzy algorithm S-dFasArt</td>
<td>41.74%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KNN, LDA, Mahalanobis distance</td>
<td>54.46%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1, T2, T3, T4</td>
<td>3</td>
<td>240</td>
<td>Yes</td>
<td>CSP</td>
<td>CSP + SVM, PCA + ICA + SVM, CBN + SVM</td>
<td>53.50%</td>
<td>[23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FBCSP-NBPW, FBCSP-Lin, CSP + SVM + KNN + LDA</td>
<td>89.04%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T1, T2, T6</td>
<td>3</td>
<td>12</td>
<td>Yes</td>
<td>SCSSP/LDA</td>
<td>SCSSP-Lin, SCSSP-NBPW, SCSP-Lin</td>
<td>79%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CSP/ERD-ERS, SVM</td>
<td>53.21 ± 6.94</td>
<td></td>
</tr>
<tr>
<td>BCI competition IV</td>
<td>T1, T2, T3, T4</td>
<td>9</td>
<td>288</td>
<td>Yes</td>
<td>CSP</td>
<td>CSP + SVM, LDA + SVM, CSP + SVM</td>
<td>59.77 ± 9.97</td>
<td>[24]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T3, T4</td>
<td>9</td>
<td>288</td>
<td>Yes</td>
<td>SCSSP/LDA</td>
<td>CBN + SVM, PCA + LDA, CSP + SVM + KNN + LDA</td>
<td>52.42 ± 6.94</td>
<td></td>
</tr>
<tr>
<td>Other data sets</td>
<td>T1, T2, T3, T4</td>
<td>8</td>
<td>288</td>
<td>Yes</td>
<td>ICA, CSP</td>
<td>SVM (RBF Kernel)</td>
<td>91%</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T3</td>
<td>3</td>
<td>480</td>
<td>Yes</td>
<td>SampEn</td>
<td>FDA</td>
<td>50%</td>
<td>[13]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T2</td>
<td>33</td>
<td>30</td>
<td>Yes</td>
<td>AAR</td>
<td>SMR, two-dimensional linear classifier</td>
<td>85%</td>
<td>[26]</td>
</tr>
<tr>
<td>Epoc</td>
<td>T1, T2, T3</td>
<td>5</td>
<td>800</td>
<td>Yes</td>
<td>ICA, CWT</td>
<td>SVM</td>
<td>66.16%</td>
<td>[29]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T4</td>
<td>1</td>
<td>75</td>
<td>Yes</td>
<td></td>
<td>Simple logistic, Meta, MLP, Neural networks (PSO)</td>
<td>80.40%</td>
<td>[30]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T3, T4, T7</td>
<td>3</td>
<td>40</td>
<td>Yes</td>
<td>BP</td>
<td>HMM</td>
<td>9%</td>
<td>[31]</td>
</tr>
<tr>
<td></td>
<td>T1, T2, T3</td>
<td>3</td>
<td>30</td>
<td>Yes</td>
<td>BP</td>
<td>LDA</td>
<td>77.50%</td>
<td>[32]</td>
</tr>
</tbody>
</table>
Table 2: Accuracy of classifier in motor imagery based BCI: two classes. The classes are (T1) left imagined hand movements and (T2) right imagined hand movements.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Activity</th>
<th>Subjects</th>
<th>Trials</th>
<th>Filter</th>
<th>Feature</th>
<th>Classifier</th>
<th>Accuracy</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCI competition III</td>
<td>T1, T2</td>
<td>5</td>
<td>280</td>
<td>Yes</td>
<td></td>
<td>LS-SVM (RBF kernel)</td>
<td>95.72%</td>
<td>[34]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>4</td>
<td>140</td>
<td>Yes</td>
<td></td>
<td>HMM</td>
<td>77.50% +</td>
<td>[35]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>4</td>
<td>200</td>
<td>Yes</td>
<td>CSP</td>
<td>SVM (Gaussian kernel)</td>
<td>77.50%</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>7</td>
<td>100</td>
<td>Yes</td>
<td>CSP/EMD, PCA</td>
<td>KNN</td>
<td>85.8%</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>4</td>
<td>100</td>
<td>Yes</td>
<td></td>
<td>SVM</td>
<td>81%</td>
<td>[29]</td>
</tr>
<tr>
<td>Other data sets</td>
<td>T1, T2</td>
<td>4</td>
<td>90</td>
<td>Yes</td>
<td>CSP</td>
<td>SVM (Gaussian kernel)</td>
<td>74.10%</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>1</td>
<td>140</td>
<td>Yes</td>
<td>CSP/ERS</td>
<td>BSSFO-SVM</td>
<td>97.57%</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>8</td>
<td>200</td>
<td>Yes</td>
<td>CSP</td>
<td>Based on the decoding principle</td>
<td>64%</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>109</td>
<td>90</td>
<td>Yes</td>
<td>CSP/ERD-ERS, SVM</td>
<td>SUTCCSP</td>
<td>90%</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>4</td>
<td>480</td>
<td>Yes</td>
<td>CSP/EMD, FFT</td>
<td>LDA, SVM, BPNN</td>
<td>84%+</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>3</td>
<td>90</td>
<td>Yes</td>
<td>BSS, CSP</td>
<td>SVM</td>
<td>92%</td>
<td>[40]</td>
</tr>
<tr>
<td>Epoc</td>
<td>T1, T2</td>
<td>5</td>
<td>120</td>
<td>Yes</td>
<td>CSP/EMD, MIDKRA</td>
<td>PSD, Hjort, CWT, DWT</td>
<td>97.79%</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>8</td>
<td>100</td>
<td>Yes</td>
<td>ERS, ERD</td>
<td>LDA</td>
<td>70.37%</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>2</td>
<td>140</td>
<td>Yes</td>
<td>CSP</td>
<td>Naive Bayes</td>
<td>79%</td>
<td>[43]</td>
</tr>
<tr>
<td></td>
<td>T1, T2</td>
<td>15</td>
<td>40</td>
<td>Yes</td>
<td>Wavelet, PSD, EMD</td>
<td>KNN</td>
<td>91.80%</td>
<td>[44]</td>
</tr>
</tbody>
</table>

database was strengthened by adding a greater number of subjects and by combining decision trees in the classification phase based on use of the boosting technique.

2. Materials and Methods

2.1. Quaternions. Quaternions were proposed in 1843 by Hamilton [22], as a set of four constituents (a real and three imaginary components) as follows: \( q = w + ix + jy + kz \), where \( w, x, y, z \in \mathbb{R} \) and \( i, j, k \) are symbols of three imaginary quantities known as imaginary units. These units follow these rules:

\[
i^2 = j^2 = k^2 = ijk = -1
ij = k,
jk = i,
ki = j
ji = -k,
kj = -i,
kj = -j.
\]

A quaternion can be described as

\[
q = (s + a), \quad a = (x, y, z),
\]

where \( s \) and \( a \) are known as the quaternion's scalar and vector, respectively. When \( s = 0 \), \( q \) is known as a pure quaternion.

Based on the expanded Euler's formula, the rotation for quaternion around the axis \( n = [nx, ny, nz] \) by angle theta is defined as follows (see [21, 45] for further details).

A rotation of angle \( \theta \) around a unit vector \( a = [a_x, a_y, a_z] \) is defined as follows:

\[
q = \cos \left( \frac{\theta}{2} \right) + (a_x \cdot i + a_y \cdot j + a_z \cdot k) \sin \left( \frac{\theta}{2} \right).
\]

Furthermore, the operation to be performed on a vector \( r \) to produce a rotated vector \( r' \) is

\[
r' = q r q^{-1} = \left( \cos \left( \frac{\theta}{2} \right) + a \sin \left( \frac{\theta}{2} \right) \right) \cdot r \left( \cos \left( \frac{\theta}{2} \right) - a \sin \left( \frac{\theta}{2} \right) \right).
\]

Equation (4) is a useful representation that makes the rotation of a vector easier. We can see that \( r \) is the original vector, \( r' \) is the rotated quaternion, and \( q \) is the quaternion that defines the rotation.

2.2. iQSA Algorithm. The iQSA is a method that improves the performance and precision of the QSA method, which is a technique to analyze EEG signals for extracting features based
The iQSA approach consists of three modules: quaternion, classification, and learning, which are described as follows.

(1) Quaternion module: in this module a features matrix (M) is defined from the description of the quaternion q and a vector R, where each of them corresponds to an array of 4 and 3 EEG channels, respectively. This module is divided into three steps, which are described as follows:

(a) Sampling window: here, we define the sample size (ns) to be analyzed and a displacement of the window in the signal (t, disp). That is, the iQSA method performs the sampling by means of superposing, producing a greater number of samples to reinforce the learning stage of the algorithm.

(b) Calculate rotation and module: then, a rotated vector \( q_{\text{rot}} \) is calculated using the quaternion q and vector R, where q is a m-by-4 matrix containing m quaternions and R is an m-by-3 matrix containing m quaternions displaced on the basis of a \( dt \) value. Later, the modulus is applied to the quaternion \( q_{\text{mod}} \) resulting in the vector \( q_{\text{mod}} \).

(c) Building an array of features: finally, the array \( q_{\text{mod}} \) is used to form a matrix with \( M_{i,j} \) features, where \( i \) corresponds to the analyzed segment and \( j \) is one of the 4 features to be analyzed using the equations included in Table 3, that is, mean (\( \mu \)), variance (\( \sigma^2 \)), contrast (con), and homogeneity (H).

Equation (5) shows matrix M with its features vector. In this matrix rows correspond to samples and columns to features.

\[
M = \begin{bmatrix}
\mu_1 & \sigma_1^2 & \text{con}_1 & H_1 \\
\mu_2 & \sigma_2^2 & \text{con}_2 & H_2 \\
\vdots & \vdots & \vdots & \vdots \\
\mu_i & \sigma_i^2 & \text{con}_i & H_i
\end{bmatrix} .
\]

(2) Classification module: the aim of this module is to create a combination of models to predict the value of a class according to its characteristics; to do this we use the boosting method adapted to the QSA model. To be more specific, we combine ten decision trees and a weight is obtained for each of them, which will be used during the learning to obtain the prediction by a majority rule. Here, we take 70% of samples from \( M \), of which the 80% are used for training and the rest for validation as follows:

(a) Training: the subset of training data (80% samples) is assessed, using the models of decision trees to obtain a matrix with accurately classified samples (G) and a matrix with inaccurately classified samples (B). The learning of each tree is done by manipulating the training data set and partitioning the initial set in several subsets according to the classification results obtained; that is, a new subset is formed, generated, and created with the accurately classified samples, twice that number of inaccurately classified samples and the worst-classified samples from the original training data set. Later, when the trees of classification are created, these are used with the test data set to determine a weight in function of the accuracy of their predictions.

(b) Validation: in this block we obtain a matrix with the reliability percentages given by the decision trees. The subset of validation data (20% samples) is assessed, using the decision trees to obtain an array with reliability values given by the processing of decision trees and the classes identified in the training process.

(3) Learning module: in this module, the subset of test data of \( M \) (30% samples) is assessed to obtain reliability values matrix (R) and the prediction by majority rule (C). That is, the learning process will be conducted assessing features matrix of test using the trees that have been generated in classification process. The final predictor comes from a weighted majority rule of the predictors from various decision trees. Equation (6) shows the recognition and error rates obtained in each of the prediction models. In this matrix RT corresponds to accuracy rate, ET corresponds to error rate, and \( \alpha \) corresponds to reliability value of each decision tree.

\[
R = \begin{bmatrix}
RT_1 & ET_1 & \alpha_1 \\
RT_2 & ET_2 & \alpha_2 \\
\vdots & \vdots & \vdots \\
RT_i & ET_i & \alpha_i
\end{bmatrix} .
\]

In Algorithm 1, we present the pseudocode for the main elements of the iQSA algorithm towards real-time applications.
2.2.1. iQSA versus QSA Methods. In Figure 1, we show through block diagrams the main differences between the iQSA and QSA methods. It can be observed that QSA method considers quaternion and classification modules. On the other hand, iQSA method considers an improved classification module and one more for learning.

In Table 4, we present some of the main improvements made in the iQSA algorithm.

<table>
<thead>
<tr>
<th>Modules</th>
<th>QSA</th>
<th>iQSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling window</td>
<td>X</td>
<td>•</td>
</tr>
<tr>
<td>Quaternion module</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Classification module</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Boosting technique</td>
<td>X</td>
<td>•</td>
</tr>
<tr>
<td>Learning module</td>
<td>X</td>
<td>•</td>
</tr>
<tr>
<td>Superposition technique</td>
<td>X</td>
<td>•</td>
</tr>
<tr>
<td>Weight normalization</td>
<td>X</td>
<td>•</td>
</tr>
<tr>
<td>Majority voting</td>
<td>X</td>
<td>•</td>
</tr>
<tr>
<td>Processing type</td>
<td>Batch</td>
<td>Real-time</td>
</tr>
</tbody>
</table>

2.3. BCI System. A brain-computer interface is a system of communication based on neural activity generated by the brain. A BCI measures the activity of an EEG signal by processing it and extracting the relevant features to interact in the environment as required by the user. An example of this device is the Emotiv-Epoc device (Figure 3(a)), a noninvasive mobile BCI device with a gyrosensor and 14 EEG channels (electrodes) and two reference channels (CMS/DRL) with a 128 Hz sample frequency. The distribution of sensors in the headset is based on the international 10–20-electrode placement system with two sensors as reference for proper placement on the head with channels labeled as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 (Figure 3(b)).

An advantage of the Emotiv-Epoc device is its ability to handle missing values, very common and problematic when dealing with biomedical data.

2.4. Decision Trees and Boosting. Decision trees (DT) [46–48] are a widely used and easy-to-implement technique that offers high speed and accuracy rates. DT are used to analyze data for prediction purposes. In short, they work by setting conditions or rules organized in a hierarchical structure where the final decision can be determined following conditions established from the root to its leaves.

Algorithm 1: iQSA method

**Require**: \( \{	ext{quat} = \{q_1, q_2, q_3, q_4\}\} \), \( dt, ns, t_{disp}, \text{task} \)

1. \( y(t) \leftarrow \) segments of signals
2. for each \( ns \) in \( y(t) \) do
   a. \( q(ns) \leftarrow \text{quat}(ns) \)
   b. \( r(ns) \leftarrow \text{quat}(ns - dt) \)
   c. \( q_{\text{mod}}(ns) \leftarrow \text{mod}(q(ns)) \)
   d. \( M_{i,j} \leftarrow f_j(q_{\text{mod}}(ns)) \) \( \{j = 1, \ldots, n\} \)
   e. \( c_i \leftarrow \{c = (1, 2, 3, \ldots, n) | y_j(t) \in c\} \)
   f. \( ns \leftarrow ns + t_{disp}; \)
3. end for
4. \( M_j \leftarrow \{M_{i,j} | \%k/\#i = \%d\} \)
5. \( \beta, \text{BTree} \leftarrow \text{Classify module}(M_{i,j}, c_i) \)
6. \( [R, \hat{C}, V_m] = \text{Learning module}(M_{i,j}, \beta, \gamma) \)
7. \( \%rt = \#C_i / \#C_k / \#C_j \)
8. \( \%rv = \#C_i / \#C_j / \#C_k \)
9. Function: Classify module \((M, c)\)
   a. \( M_{i,j} \leftarrow \{M_{i,j} | \%a/\#i = \%d\} \)
   b. \( M_{i,j} \leftarrow \{M_{i,j} | \%b/\#a = 1 - \%d\} \)
10. /*Training process*/
   11. for \( k = 1 \) to \( 10 \) do
       a. \( [G_{m,j}, B_{m,j}, \text{BTree}] \leftarrow \text{training}(M_{i,j}, c_i) \)
       b. \( M_{i,j} \leftarrow \{M_{i,j} | \%a/\#m_j = X\} \)
   12. end for
   13. /*Validation process*/
   14. for each \( \text{BTree} \) do
       a. \( \beta_i \leftarrow \text{validation}(\text{BTree}, M_{i,j}, c_i) \)
   15. end for
   16. return \( \beta, \hat{C}, \text{BTree} \)
17. Function: Learning module \((M, c, \beta, \text{BTree})\)
   a. for each \( \text{BTree} \) do
       a. \( [R, \hat{C}, V_m] \leftarrow \text{classify}(\text{BTree}, M_{i,j}, \beta, c_i) \)
   18. end for
   19. return \( R, \hat{C}, V_m \)
Recently, several alternative techniques have been presented to construct sets of classifiers whose decisions are combined in order to solve a task and to improve the results obtained by the base classifier. There exist two popular techniques to build sets: bagging [49] and boosting [50]. Both methods operate under a base-learning algorithm that is invoked several times using various training sets. In our case, we implemented a new boosting method adapted to QSA method, whereby we trained a number of weak classifiers iteratively so that each new classifier (weak learner) focuses on finding weak hypothesis (inaccurately classified). In other words, boosting calls the weak learner \( w \) times thus determining, at each iteration, a random subset of training samples by adding the accurately classified samples, twice that number of inaccurately classified samples and the worst-classified samples from the original training data set to form a new weak learner \( w \).

As a result, inaccurately classified samples of the previous iteration are given an \( (\alpha) \) weight in the next iteration, forcing the classifying algorithm to focus on data that are harder to classify in order to correct classification errors of the previous iteration. Finally, the reliability percentage of all the classifiers is added and a hypothesis is obtained by a majority vote, whose prediction tends to be the most accurate.

2.5. Channel Selection for iQSA. From the 14 electrodes that the Emotiv-Époc device provides, we decided to perform an analysis on the electrodes located in three different regions of the cerebral cortex looking for those with better performance to form the quaternion: (1) electrodes located on the motor cortex, which generates neural impulses that control movements, (2) those on the posterior parietal cortex, where visual information is transformed into motor instructions [51], and (3) the ones on the prefrontal cortex, which appear as a marker of the anticipations that the body must make to adapt to what is going to happen immediately after [52].

In this regard, in Table 5 we present the performance for each of the sets of channels that were selected and further analyzed by the iQSA algorithm.
Comparing the data sets shown in Figure 4, it is observed that all have a similar behavior in each sample size. For a 64-sample analysis, data sets 2 and 4 obtained 82.30% and 82.33%, respectively. The channels for data set 2 are related to the frontal and motor areas of the brain and the channels for data set 4 are related to the parietal and motor areas. From these results and due to the fact that we mainly focus on motor control tasks, set 2 (F3, F4, FC5, and FC6) has been chosen.

3. Experiment

As seen in earlier sections, one of the objectives of this paper is to analyze EEG signals towards real-time applications. Therefore, it is necessary to reduce the number of samples of EEG signals that have been acquired by subjecting them to a process based on the iQSA method to decode motor imagery activities, while keeping or improving
Table 5: Accuracy rate to several data sets of channel blocks.

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>384</td>
</tr>
<tr>
<td>FC5, FC6, P7, P8</td>
<td>74.16</td>
</tr>
<tr>
<td>F3, F4, FC5, FC6</td>
<td>73.16</td>
</tr>
<tr>
<td>F3, F4, F7, F8</td>
<td>73.48</td>
</tr>
<tr>
<td>F3, F4, P7, P8</td>
<td>73.47</td>
</tr>
<tr>
<td>AF3, AF4, FC5, FC6</td>
<td>74.17</td>
</tr>
</tbody>
</table>

Figure 4: Comparison of data sets with different sample sizes. Data set 1: FC5, FC6, P7, P8; data set 2: F3, F4, FC5, FC6; data set 3, F3, F4, F7, F8; data set 4: F3, F4, P7, P8; data set 5: AF3, AF4, FC5, FC6.

Figure 5: Methodology for activities in motor-cognitive experiments.

accuracy results achieved with this technique. In this way, the experiment was designed to register EEG signals from several individuals and to identify three motor-imagery-related mental states (think left motion, think right motion, and waiting time).

3.1. Description of the Experiment. The experiment was conducted in three sessions: Session 1 involved motor imagery with tagged visual support ($V + L$), that is to say, an arrow with the word LEFT or RIGHT written on it. Session 2 involved motor imagery with visual ($V$) support only. In Session 3 the subject only received a tactile stimulus ($T$) to evoke motor imagery while keeping his or her eyes closed. In each session, the aim was to identify three mental states (think left motion, think right motion, and waiting time). The visual $V$ and VL stimuli were provided using a GUI interface developed in Python 2.7 to show the movement of a red arrow for 5 seconds for each of the motor brain actions and a fixed cross at the center of the GUI to indicate a rest period for 3 seconds (the third brain action). The tactile stimulus was provided by touching the left/right shoulder of the individual to indicate the brain action to be performed. Each session lasted for 5 minutes and was done on separate days.

To start with, the participant was asked to sit on a comfortable chair in front of a computer screen. As the participant was given instructions regarding the test, an Emotiv-Epoc headset was placed on his or her head, making sure that each of the Emotiv device electrodes was making proper contact with the scalp (Figure 5). Once the participant was ready, he or she was asked not to make sudden body movements that could interfere with the signal acquisition results during the experiment.

The training paradigm consisted of a sequential repetition of cue-based trials (Figure 5). Each trial started with an empty blank screen; during the time $t = 0$ to $t = 3$ s a cross was displayed to indicate to the user that the experiment had started and that it was time to relax. Then at second 3 ($t = 3$ s) an arrow appearing for 5 s pointed either to the left or to the right. Each position indicated by the arrow instructed the subject to imagine left or right movement, respectively. The next trial started at $t = 8$ s with a cross. This process was repeated for 5 minutes, displaying the arrows 32 times and the cross 33 times within each run. Thus, the data set recorded for the three runs consisted of 96 trials.

3.2. iQSA Method Implementation. After data acquisition, the next stage consists in extracting EEG signal features in order to find the required classes, that is, think left motion, think right motion, and waiting time. To start with, the iQSA method was implemented to represent the four EEG signals within a quaternion and to carry out the feature extraction related to the stimuli presented. To that effect, data set 2 was prepared to assess their performance when time (and the number of samples) was reduced. Under these conditions, 3 seconds (384 samples), 2.5 seconds (320 samples), 2 seconds (256 samples), 1.5 seconds (192 samples), 1 second (128
samples), and 0.5 seconds (64 samples) were considered. Values 0, 1, and 2 were used to refer to the three mental activities: waiting time (0), think left motion (1), and think right motion (2).

So, a segments matrix was obtained to detect sudden changes between classes for each data set. The segments matrix consisted of samples from 32 trials from left and right classes and samples from 33 trials for the waiting time class. In addition, the quaternion was created with the block signals proposed (F3, F4, FC5, and FC6), considering F3 as the scalar component and F4, FC5, and FC6 as the imaginary components (Figure 6). From the samples to be analyzed, several segments were created to generate the quaternion \( q \) and vector \( R \) with a displacement \( dt = 4 \) and thus obtain the rotation \( q_{\text{rot}} \) and modulus \( q_{\text{mod}} \).

Once the module was obtained, the mean (\( \mu \)), contrast (\( \text{con} \)), homogeneity (\( H \)), and variance (\( \sigma^2 \)) features were calculated to generate the matrix \( M \) and the vector \( c \) with the required classes. Later, we returned to the current segment, and a displacement of 64 samples for the signal was effected to obtain the next segment.

Later, \( M \) was used in the processing stage using 70% of data from training and 30% from test; here each class has the same sample size and was randomly chosen. In the classification module, we generate the matrix \( M_1 \) with the 80% from training data of \( M \) and \( M_2 \) with the remaining 20% of the training data, which will be used for training and validation, where 10 training trees were created to force the algorithm to focus on the inaccurately classified data. Here, for each iteration we generate new subsets of samples with the double of inaccurately classified samples and fewer accurately classified samples, along with a reliability rate for each tree. Later, with matrix \( M_2 \) the data were validated to obtain the percentage of reliability of each tree. Finally, with the remaining data of \( M \), a prediction by majority rule is voted and a decision is reached based on the reliability rate of each classification tree.

4. Results

As said earlier, one of the aims of this study was to reduce the assessment time (sample size) without loss of the accuracy rate. Therefore, a comparison of the iQSA and QSA techniques was performed to show its behavior when the number of samples decreases.

4.1. Comparison between iQSA and QSA Methods. To evaluate and compare the performance of the iQSA online versus QSA offline algorithms, the same sets of data from the 39 participants were used, considering the different sample sizes (384, 320, 256, 192, 128, and 64 samples) as input for the algorithm.
Table 6: Table of accuracy rate for iQSA and QSA method with different number samples.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>QSA (1)</th>
<th>QSA (2)</th>
<th>iQSA (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>384</td>
<td>40.82%</td>
<td>41.07%</td>
<td>73.16%</td>
</tr>
<tr>
<td>320</td>
<td>40.74%</td>
<td>40.29%</td>
<td>74.87%</td>
</tr>
<tr>
<td>256</td>
<td>43.56%</td>
<td>43.96%</td>
<td>80.23%</td>
</tr>
<tr>
<td>192</td>
<td>42.41%</td>
<td>42.71%</td>
<td>81.39%</td>
</tr>
<tr>
<td>128</td>
<td>37.45%</td>
<td>38.28%</td>
<td>81.65%</td>
</tr>
<tr>
<td>64</td>
<td>33.31%</td>
<td>33.44%</td>
<td>82.30%</td>
</tr>
</tbody>
</table>

In spite of the good performance rates reported with the QSA algorithm for offline data analysis, the classification results were not as good when the number of samples was reduced up to 64 samples as shown in Figure 7(a). Graphically, as can be seen, the error rate increases gradually when the sample size is reduced for analysis, reaching up to 66.56% error for a reading of 64 samples. In this way, it was necessary to make several adjustments to the algorithm, in such a way as to support an online analysis with small samples sizes, without losing information and sacrificing the good performance rates provided by the offline QSA algorithm.

Thus, the data of the 39 subjects were evaluated, considering three cases: (1) QSA method without window samples, (2) QSA method with window samples, and (3) iQSA proposed method. Comparing the results of Table 6, the precision percentages for the iQSA method range from 73.16% for 384 samples to 82.30% for 64 samples, unlike the original QSA method whose percentage decreases to 33.31% for 64 samples in case 1 and 33.44% in case 2 although the sampling window has been implemented.

Figure 8 shows the behavior of the iQSA technique in blue and QSA technique in red, where the mean, maximum, and minimum of each technique are shown. Comparing both techniques, it is observed that the data analyzed with the original QSA method drastically loses precision as the number of samples selected decreases. This was not the case for data sets using the iQSA technique whose precision did not vary too much when the number of samples for classification was reduced, improving even its performance.

4.2. iQSA Results. Figure 9 shows the behavior of 39 participants under both techniques (iQSA and QSA), considering all the sample sizes. The values obtained using the iQSA method are better than values obtained with QSA method whose values are below 50%. Numerical results show that the iQSA method provides a higher accuracy over the original QSA method for tests with a 64-sample size.

Given the above results, a data analysis was conducted using 64 samples to identify the motor imagery actions. The performances obtained with the set of classifiers were compared using various assessment metrics, such as recognition rate (RT) and error rate (ET), and sensitivity (S).
and specificity (Sp). The sensitivity metric shows that the classifier can recognize samples from the relevant class, and the specificity is also known as the real negative rate because it measures whether the classifier can recognize samples that do not belong to the relevant class.

\[
RT = \frac{\# \{c | c = \hat{c} \}}{\# \{c \}} \tag{7}
\]

\[
ET = \frac{\# \{c | c \neq \hat{c} \}}{\# \{c \}} \tag{8}
\]

\[
S_d = \frac{\# \{c | c = d, c = \hat{c} \}}{\# \{c | c = d \}} \tag{9}
\]

\[
Sp_d = \frac{\# \{c | c \neq d, c = \hat{c} \}}{\# \{c | c \neq d \}} \tag{10}
\]

As Table 7 shows, the highest accuracy percentage was 84.50% and the lowest 68.75%. In addition, the sensitivity average for class 0 \((S_0)\) associated with the waiting time mental state was 82.53%, and the sensitivity average for class 1 \((S_1)\) and class 2 \((S_2)\) associated with think left motion and think right motion was 81.07% and 81.65%, respectively, which indicates that the classifier had no problem classifying classes. In turn, the specificity rate for class 0 \((Sp_0)\) shows that 81.36% of the samples classified as negative were actually negative, while class 1 \((Sp_1)\) performed at 82.09% and class 2 \((Sp_2)\) at 81.80%.

To evaluate the performance of our approach, we have carried out a comparison with other methods such as FDCSP [53], MEMD-SI-BCI [54], and SR-FBCSP [55] using data set 2a from BCI IV [56] and the results are shown in Table 8. Our method shows a slight improvement (1.09%) compared to the SR-FBCSP method, which presents the best results of the three.

To analyze our results, we performed a significant statistical test making use of the STAC (Statistical Tests for Algorithms Comparison) web platform [57]. Here, we chose the Friedman test with a significance level of 0.10 to get a ranking of the algorithms and check if the differences between them are statistically significant.

Table 9 shows the Friedman test ranking results obtained with the \(p\) value approach. From such table, we can observe that our proposal gets the lowest ranking; that is, iQSA has the best results in accuracy among all the algorithms.

In order to compare whether the difference between iQSA and the other methods is significant, a Li post hoc procedure was performed (Table 10). The differences are statistically significant because the \(p\) values are below 0.10.

In addition, the classification shown in Table 8 is achieved by the iQSA in real time and compared to the other methods a prefitering process is not required.

In Table 11, we show the time required for each of the tasks performed by the iQSA algorithm: (1) features extraction, (2) classification, and (3) learning. The EEG signal analysis in processes 1 and 2 was responsible for obtaining the quaternion from the EEG signals, learning trees, and training. Process 3 evaluates and classifies the signal based on the generated decision trees, as should be done in real-time analysis.

According to the experiment, for offline classification (processes 1 and 2), the processing time required was 0.3089 seconds. However, the time required to recognize the motor imagery activity generated by the subject was of 0.0095 seconds, considering that the learning trees were already generated, and even when the processing phase was done in real time, it would take 0.3184 seconds to recognize a pattern after the first reading. With this, our proposed method can be potentially used in several applications such as controlling a robot, manipulating a wheelchair, or controlling home appliances, to name a few.

As said earlier, the experiment consisted of 3 sessions with each participant (visual tagged, visual, and tactile). The average accuracy obtained from the 3 experiments is between 76% and 78%. Figure 10 shows results from the
visual stimulation session, which produced the lowest rate at 76.63%, followed by tagged visual session which reached 76.98% and the tactile session 77.28%.

In Figure 10, it is shown that the mental activity generated with the aid of a tactile stimulus is slightly more accurate than visual and visual tagged stimuli where recognition is more imprecise and slow with visual stimulus.

To summarize, the results of tests conducted with 39 participants using this new method to classify motor imagery brain signals, 20 times with each participant, considering 70%–30% of the data have been presented and creating several subsets of 80–20% for the classification process. The average performance accuracy rate was 81.75% when using 10 decision trees in combination with the boosting technique at 0.5-second sampling rates. The results show that this methodology for monitoring, representing, and classifying EEG signals can be used for the purposes of having individuals control external devices in real time.
and half a second. This new method was compared to the original QSA technique, whose accuracy rates ranged from 40.82% to 33.31% without sampling window and from 41.07% to 33.34% with sampling window. We can thus conclude that iQSA is a promising technique with potential to be used in motor imagery recognition tasks in real-time applications.

**Conflicts of Interest**

The authors declare that there are no conflicts of interest regarding the publication of this paper.

**Acknowledgments**

This research has been partially supported by the CONACyT project “Neurociencia Computacional: de la teoría al desarrollo de sistemas neuromórficos” (no. 1961). Horacio Rostro-Gonzalez acknowledges the University of Guanajuato for the support provided through a sabbatical year.

**References**


Submit your manuscripts at
https://www.hindawi.com