K-NN Classification of Brain Dominance

Khairul Amrizal Abu Nawas, Mahfuzah Mustafa, Rosdiyana Samad, Dwi Pebrianti, Nor Rul Hasma Abdullah
Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang, Malaysia

**ABSTRACT**

The brain dominance is referred to right brain and left brain. The brain dominance can be observed with an Electroencephalogram (EEG) signal to identify different types of electrical pattern in the brain and will form the foundation of one’s personality. The objective of this project is to analyze brain dominance by using Wavelet analysis. The Wavelet analysis is done in 2-D Gabor Wavelet and the result of 2-D Gabor Wavelet is validated with an establish brain dominance questionnaire. Twenty-one samples from University Malaysia Pahang (UMP) student are required to answer the establish brain dominance questionnaire has been collected in this experiment. Then, brainwave signal will record using Emotiv device. The threshold value is used to remove the artifact and noise from data collected to acquire a smoother signal. Next, the Band-pass filter is applied to the signal to extract the sub-band frequency components from Delta, Theta, Alpha, and Beta. After that, it will extract the energy of the signal from image feature extraction process. Next the features were classified by using K-Nearest Neighbor (K-NN) in two ratios which 70:30 and 80:20 that are training set and testing set (training: testing). The ratio of 70:30 gave the highest percentage of 83% accuracy while a ratio of 80:20 gave 100% accuracy. The result shows that 2-D Gabor Wavelet was able to classify brain dominance with accuracy 83% to 100%.

1. **INTRODUCTION**

This project is carried to provide more precise method for brain dominance classification by using Wavelet analysis. The idea is to provide new analysis method to determine the brainwaves signal of brain dominance for the personality of person in terms of four sub-band frequencies which Delta, Theta, Alpha and Beta. In this project, there are four types of basic bands is used to determine the dominance of brainwaves. Each of the bands are different in frequencies range where Delta is 0 Hz – 4 Hz, Theta is 4 Hz – 8 Hz, Alpha is 8 Hz – 12 Hz and Beta is more than 12 Hz. These four bands would result a pattern that used to determine the brain dominance by signal processing.

EEG signal will be analyzing by using Two-Dimensional Wavelet technique. Our initial experiment is done using Power Spectral Density (PSD) and Energy Spectral Density(PSD) [1], [2]. This paper is to explore more technique to analysis EEG signal in brain dominance application.

Best extracted feature will be proceeded to classification by using K-Nearest Neighbour (K-NN) classification. This classification will be grouping and classify the data from Wavelet analysis with establish questionnaire to obtain the accuracy percentage [3]. At the end of analysis, the accuracy from classification will shows the brain dominance for each sample.

**Corresponding Author:**
Mahfuzah Mustafa,
Faculty of Electrical & Electronics Engineering,
Universiti Malaysia Pahang,
Pekan Campus, 26600 Pekan, Pahang, Malaysia.
Email: mahfuzah@ump.edu.my

**Article Info**

**Article history:**
Received Mar 9, 2018  
Revised Jun 12, 2018  
Accepted Jun 19, 2018

**Keyword:**
Brain dominance  
K-NN  
Wavelet
In classical, brain dominance can be determined by establish questionnaire. Brainwaves have several rhythm frequencies that can be measured by device called as EEG. When to analyze the signal, there are a lot of technique back then such Fast Fourier Transform (FFT), Power Spectral Density (PSD) and others. However, those methods still not enough for an EEG signals as the features extraction. This is because FFT and PSD determine the pattern of brainwave by highest peak value of the signal and transportation signals are non-stationary. Moreover, times and frequency information cannot be seen at the same time. It is difference with Wavelet analysis where it took the whole signal to extract the energy value of the signal. Therefore, Wavelet analysis is preferred. This is also Wavelet can perform multi resolution time-frequency analysis.

Research has indicated that most people are dominant in one brain hemisphere. Determining for dominant brain hemisphere can play a role in determining how the people learn best and how his relate to others. But, any inventory is just a guide. People are most successful when they integrate and develop both sides of the brain. From the Herrmann brain dominance questionnaire that it can establish which side the brain for some person [4]. It is also useful to consider the face validity of the concept of brain dominance measurement to experts in scientific fields associated with the brain. Here it is safe to say that the very idea of left brain-right brain dominance remains somewhat controversial.

However, those have some another technique to determine brain dominance that is through by brainwave [5]. In other term, brainwave may call as neural oscillation. It is rhythmic or repetitive neural activity in the central nervous system which also known as brain. This kind of oscillation can be observed in the EEG. The EEG is the recording of electrical activity along the scalp. Thus, dominance brainwave is one of the brains activities that flows within the neurons of the brain that can be classify by EEG [6]. From the Emotiv EPOC device, that can use for collecting the data from brainwave. It will be pre-processed and extracted frequency band in some of range classification whether in Delta, Theta, Alpha and Beta.

Emotiv EPOC with 14 electrodes include with two references are most commonly as shown in Figure 1(a) and Figure 1(b) to detect EEG signal. Normally, the EEG is determined through AF3 and AF4 as shown in Figure 2. Apparently, these two channels also known as Fp 1 and Fp 2 that been proposed in most of projects and researches. EEG can be analyzed in time domain, frequency domain and even time-frequency domain. However, most of projects and researches carried previously, they were analyzing in either time domain or frequency domain. It is seldom happened to be time-frequency domain due to its limited technique that found currently.

Figure 1. (a) Emotiv EPOC device [7], (b) Emotiv EPOC channel [7]

EEG signal that get from Emotiv Epo device is in non-stationary signal. For non-stationary signal, there are need information in both the frequency and time domains at once [8]. Fourier Transform (FT) is normally uses for transformation signal technique. But, the times and frequency information cannot be seen at the same time. That only gives the frequency exist in the signal. Therefore, Fourier analysis isolates well in frequency, but not in time.

The Wavelet analysis, that support the finite signal because Wavelets are localized waves, although some Wavelets do not have compact support compare with (FT). In this case, the performance of Doubechies 4 (db4) decomposition is exact due to a great similarity between the analyzed signal and the Wavelet itself as shown in Figure 2 [9]. The declaration is importance to choose of Wavelet for the decomposition of the signal and denoising technique. Generally, choosing longer Wavelets leads to a better frequency resolution but a worse time resolution [9].

2-D wavelet analysis is the method to detect the edge of information signal based on the feature of image texture [10]. Frequency and orientation representations of Gabor filter are like those of the human
visual system, and they have been found to be particularly appropriate for texture representation and discrimination. Gabor function provides the spectral energy density concentrated around a given position and frequency in a certain direction. Moreover, the popular Gabor parameters, 5 scales X 8 orientations is the best choice in many studies and that will apply to this Gabor function [11]. Therefore, Gabor Wavelet is preferable as it can extract most of the features from a signal such as amplitude and energy [11].

K-Nearest Neighbors (K-NN) is the method that use for classifying the brain dominance and to determine which frequency band of the EEG Power Spectrum whether that is Delta, Theta, Alpha, and Beta. The classifier works by comparing a new sample (testing data) with the baseline data (training data) [12]. The classifier will determine with k neighbourhood in the training data and assign class which appear more frequently in the neighbourhood of k [13], [14].

![Wavelet decomposition tree](image)

Figure 2. Wavelet decomposition tree used for the real EEG data set and used to reconstruct the brain waves [9]

K-NN was widely used by several researches and projects for the classification purpose. This is due to K-NN is very easy to understand and simple method that would return a good accuracy. For example, training point of sample, Y = [1, 5] that will be used to classified sample point, X where X mean to be positive or negative. Let assume that sample point, X= [2, 6] and put into classifier with training point, Y = [1, 5]. Then, the classifier will classify the value of X such as 2 to the nearest value of Y and put them into the group. Therefore, where value of X is equal to 2, it is nearest to value of Y equal to 1. So, value of X equal to 2 will classify into that group.

2. RESEARCH METHOD

The experiment stage is divided into data collection, pre-processing, 2-D Wavelet and K-NN classification. The research method will be discussed in the next section.

2.1. Data collection

Participants were volunteering undergraduate students from Universiti Malaysia Pahang. Total participants were 21 among 6 males and 15 females. The EEG signals was downloading offline from Emotiv Test Bench after the participant answer brain dominance questionnaire and they were given rest 2 to 5 minute before EEG data acquisition. The brainwave signal is recorded using 16 channels including two references in EEG channel. The period protocol of the experiment is shown in Figure 3. The total time required for each participant is approximately 15 to 20 minutes. Initially, participant rested for 2 minutes while EEG experimental procedures were carried out. Consequently, EEG signal were recorded continuously for 5 minutes during eyes closed.

Brainwave signal will record by using Emotiv device and it has 16 channels, but in this analysis required only two out of 16 channels. That is Fp 1 and Fp 2 channel which channels contains the most features that help in determination of brain dominance because that are in near of left and right frontal brain. The recorded brainwave signal will be saved in edf format.
2.2. Pre-processing

Pre-processing was carried out to remove the unnecessary signal or called as artifact from EEG signal. In pre-processing, there are two phases that is threshold and band pass filter. The signals were further processed using a specific algorithm of filtration to reject artifact and to allow only the information signal from the brainwave to pass through the correlation analysis.

Noise from a signal is resulting when the participant does an any minor movement, eye blinking, when his breath and sweating. The analysis would need only EEG signal. Therefore, other than EEG signal, they are assumed as a noise and artifact towards the signal. To remove all the unnecessary artefacts or noise from a signal, that will be use a threshold to outbound from a certain range of value. As the EEG was measured in 16 channels.

The range of EEG value is within -100μV to 100μV in amplitude. Thus, the signal in channel Fp 1 and Fp 2 was threshold within the range so that the signal contains correct EEG signal.

In this project, it required four sub-band frequencies for analysis. They are Delta, Theta, Alpha and Beta. Each of the brain represents different state of consciousness where the band wave will change according to the activity they are doing. In the MATLAB software, the Butterworth low pass and band pass filter to identify the signal whether of frequency sub-band. The EEG signal was segmented into independent four sub-bands according to their own frequency range as stated below with the same technique where the range of frequency for the filter was changed according the Table 1. It is applied to all channels as well in Fp 1 and Fp 2.

| 10-15 minutes | 2 minutes | 5 minutes |
|---------------|-----------|-----------|
| Psychoanalysis Test (answer the 15 item brain dominance questionnaires) | Rest during electrodes placement | EEG data acquisition |

Table 1. Brainwave Sub-band Pattern Description [15]

| Brainwave | Frequency (Hz) | Amplitude (μV) | Description |
|-----------|----------------|----------------|-------------|
| Beta      | >13            | lowest         | Awake       |
| Alpha     | 8 - 13         | low            | Relaxed     |
| Theta     | 4 – 8          | high           | Tired       |
| Delta     | 0.5 - 4        | highest        | Deep Sleep  |

2.3. 2-D wavelet

The time-domain analysis in biomedical signals such as EEG does not seen the time-frequency at the same time [16]. Spectral or frequency-domain analysis would be very helpful in this case. However, spectral analysis does not show us at what times the frequency changes occur. The Wavelet technique will solve the problem of resolution by using Multi-resolution analysis. The Wavelet is tools that take the data, function or operator into different frequency components, and then studies each component, and then studies each component with a resolution matched to its scale.

This project uses 2-D Gabor Wavelet to extract the expression feature from sub-regions of expression images. According to analyze and compare the expression features extracted from different images texture with different frequency band and different sample, that have obtained satisfactory experimental results that demonstrates the efficiency of our algorithm. Gabor transfer function (tfrgbabor) is used to get coefficient contains an information relative to the time-frequency content of the signal around the time-frequency location. Generally, some degree of oversampling is considered (t₀ x v₀ < 1), which introduces redundancy in the coefficients, to 'smooth' the biorthonormal window h, for the sake of numerical stability. 2-D Gabor Wavelet that extract the image in R, G, B colour. In this project, our feature extraction has similar size of image texture that is 560 X 420 pixels. That will remove the scale, axes, and title of image by using auto crop declaration in MATLAB. The removing was doing because that only analyze and segmented the information extraction. Before to get mean amplitude and mean energy from the image, that image will change into gray-scale because the Gabor function used to extract the feature from the Gabor image and that only extract from the gray-scale image. This extraction is segmented the sub-regions from an expression image and that having greatly containing expression.

**K-NN Classification of Brain Dominance (Khairul Amrizal Abu Nawas)**
2-D Gabor Wavelet will give the value of mean amplitude and squared energy after the image turned into gray-scale. From the value, that will obtain the frequency band which is the brain dominance of the person. That will detect the edge of information into the image. Phase symmetry (phasesym) function will use to compute and calculate Gabor features from the gray-scale image to get the value of mean squared energy and mean amplitude in automatically by using MATLAB. However, apart from the Gabor image processing has parameters to be set that is scale and orientation.

Moreover, the popular Gabor parameters is 5 scales x 8 orientations, have been assume to be the best choice in many studies without careful discussion and examination on their performance X. The normalization is used to change the energy value into several ranges that is ‘0’ to ‘1’. Histogram chart in Microsoft Excel was used to obtain and analyse the patterns of brain dominance with frequency band (Delta, Theta, Alpha, and Beta) after normalized the energy value X.

2.4. K-NN classification

After feature extraction from 2-D Gabor Wavelet, the features were classified by using K-Nearest Neighbor (K-NN). The accuracy from classification will shows the brain dominance for each sample. It shows the accuracy of the analysis been through. So, K-NN was used to classify them. In K-NN, there are three things need to understand in advanced. There are testing set, training set and group set. Testing set is the data to classify with approach of training set. Training set are the data that used to compare with the sample set to classify the testing set in certain that been set in group set whether left dominance and right dominance of brain. Group set can only be defined in number instead of character. For example, right brain dominance set number as ‘2’, left brain dominance set number as ‘1’ and testing will set as number ‘0’. That can ease to obtain the accuracy of data from 2-D Gabor Wavelet means that 30 percent out of 20 samples become testing set and 70 percent of 20 samples become training set. It same goes for 80:20, that is 20 percent of 20 samples become testing set and 80 percent of 20 samples become training set. Hence, the data was classified based on these two ratios and it returned a percentage of accuracy of classification.

Lastly, the testing set was classified with k from 1 to 15 set value of training set. For example, 70:30 means that my sample set is 6 out of 20 samples while training set is 14 out of 20 samples. Therefore, that will be classified the sample set from k = 1 to k = 15 for purpose to obtain which k returns me highest percentage of accuracy.

3. RESULTS AND ANALYSIS

The EEG signals collected to obtain four frequency band (Delta, Theta, Alpha, Beta) from two channels. There were 21 sample from UMP student and all EEG signal has been recorded using Emotive after each of student answering the questionnaire. There are two channels to use for completing the observation and analysed, that is Fp1 and Fp2. The MATLAB software has been used to process the data.

Table 2 shows the brain dominance questionnaire results. The observation for left and right brain dominance using this questionnaire also similar with case study before where when the person strong in mathematics and science there are more to left-brain dominance in engineering student.

| No. of samples | Brain Dominance |
|----------------|-----------------|
| 7              | R               |
| 13             | L               |
| 1              | Bilateral       |

Two-Dimensional Wavelet is the other technique where to determine sub-band frequency of brain dominance and show the result for Fp1 and Fp2 2 channel in Gabor Wavelet. The 2D Gabor functions proposed are local spatial band pass filters that achieve the theoretical limit for conjoint resolution of information in the 2D spatial and 2D Fourier domains. The Figure 4 to Figure 7 show the feature and information of the image for each sub-band frequency where the channel of Fp1.

After that, the features are extracted through the images. But, the images need to modify in advance before extraction. This because that the function applied only able to extract the features from a gray-scale form. And even the images must be excluding the axes and others after the axes. Therefore, crop image process was done for that purpose and then the images were turn into gray-scale.
K-NN Classification of Brain Dominance (Khairul Amrizal Abu Nawas)

Based on Figure 8 and Figure 9, the average data after normalization energy can be utilized the left-brain and right-brain dominance by refers the both channel and the highest histogram for each sub-band frequency. The Fp 1 channel as the left-brain dominance is indicate Beta-band whereas Fp 2 channel as the right-brain dominance is indicate Theta-band. For bilateral cases, it cannot indicate the sub-band frequency because there is only one sample.

There is possibility of 2-D Wavelet results is more precise because that it analyse energy of signal in scale ‘s’, orientation ‘θ’ and physical location (x, y). To prove this analysis, that will refer the brain dominance criteria. The brain dominance has mention where right-brain people are more creative and intuitive. The meditation analysis, Delta-band signal are associated with the deepest levels of relaxation and restorative, healing sleep. That is not similar for right-brain dominance criteria. The Theta-band frequency range is involved in daydreaming and make ‘highly suggestible’ because they are in a deeply relaxed, semi-hypnotic state. That is helping to improve our intuition, creativity and makes person more natural. These criteria also suitable and related with right-brain dominance.
K-NN is the technique that can use to analyse the result from 2-D Wavelet and to prove this analysis based on the highest accuracy percentage. K-NN was done in two of ratios and the value of k. There are 70:30 and 80:20 for ratios while the maximum k for both ratios is 15. Figure 10 and Figure 11 show the result into two classifications of ratio.

From the classification, the result obtained based on data correction which means that in certain value of k, it gave certain number of correct classification for the testing set. For example, in Figure 10, when k = 4, the result shows that 5 out of 6 testing set were correct in the classification.

![Figure 10: Accuracy percentage for ratio 70:30. The k is varied from 1 to 15. From the figure, the highest accuracy is obtained from k = 4, with accuracy 83%.](image)

![Figure 11: Accuracy percentage for ratio 80:20. The k is varied from 1 to 15. From the figure, the highest accuracy is obtained from k = 2, with accuracy 100%](image)

The result from Figure 11 show the highest percentage accuracy when k = 2 it gave 100% of accuracy. That is the best of data analysis. The means of this result also shows that if the amount of data set is higher, then the accuracy of data may go higher than this result. This is because when the ratio is changed 80:20, it results 100% where increase around 17% compared to ratio of 70:30. As a conclusion, the accuracy will be increase when the amount of training set goes higher.

4. CONCLUSION

The project able to classify brain dominance based on EEG signal referred to brain dominance questionnaire. The 2-D Gabor Wavelet could analyze brain dominance with accuracy 83% to 100% that obtained from K-NN classification.

For the recommendation and future development, this suggestion can upgrade to make this analysis is more interactive and valuable. The EEG data can be improved when recording the brainwave signal into the online system. From the brainwave signal that can show there brain dominance result into smart phone by using GSM system. The parent or lecturer can download the application and try to implement for children who are cannot read or write to change their study technique based on brain dominance classification.

The Wavelet technique used in this project in order to analyze the EEG signal gave a successful result when that can classify the brain dominance of a person. That it because the Wavelet can analyze the brain signal based on scale and orientation. However, hidden information of brain dominance in terms of brainwave can be explored in future research. At the last, the result can be improved when to use Neurosky Mindset device because the device is more accurate to capture the brain signal compared with Emotiv device.

ACKNOWLEDGEMENTS

This research was supported by a Grant at the Universiti Malaysia Pahang (RDU170365).

REFERENCES

[1] N. S. H. B. A. Hamid. 2015, “Brain Dominance Using Brainwave Signal”, Bachelor, Thesis, University Malaysia Pahang, Pahang.
K-NN Classification of Brain Dominance (Khairul Amrizal Abu Nawas)
Dwi Pebrianti is a senior lecturer in Faculty of Electrical & Electronics Engineering, University Malaysia Pahang, UMP, Malaysia since 2013. She received Bachelor of Engineering in electronics engineering from Universitas Indonesia, Indonesia, in 2001 and joined an oil gas company, Caltex Pacific Indonesia in the same year. She received Master of Engineering from the Department of Engineering Synthesis, The University of Tokyo, Japan in 2006. Her Philosophy Doctor was received in 2011 from Chiba University in the field of Artificial System Science. Since then, she was serving Chiba University as a Postdoctoral fellow in the same field. Her main works are including image processing for robot navigation, control theory for robot navigation, automation, control system, robotics, wearable computer, artificial intelligence, nonlinear system and control, Unmanned Aerial Vehicle, Motion & dynamics control, embedded system design, Motion tracking system, Swarm robot and Optimization technique.

Nor Rul Hasma Abdullah obtained a Bachelor of Electrical Engineering (Hons) from Universiti Teknologi Malaysia, MEng in Electrical Engineering from UiTTHO and PhD in Electrical Engineering from Universiti Teknologi MARA. Her research interest includes power system stability, optimization techniques, distributed generation, swarm optimization and meta-heuristic techniques. To date, she is currently a senior lecturer at Universiti Malaysia Pahang, Malaysia.