



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΕΥΦΥΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΖΩΓΡΑΦΟΥ 157 73, ΑΘΗΝΑ

Ανάλυση EEG και EOG σημάτων με ευφυείς τεχνικές για τηλεκίνηση RC αμαξιδίου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΑΡΤΕΜΙΔΟΣ ΖΩΓΡΑΦΟΥ

Επιβλέπων: Ανδρέας Γεώργιος Σταφυλοπάτης

Συνεπιβλέπων: Γεώργιος Σιόλας

Αθήνα, Μάρτιος 2018



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΕΥΦΥΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΖΩΓΡΑΦΟΥ 157 73, ΑΘΗΝΑ

Ανάλυση EEG και EOG σημάτων με ευφυείς τεχνικές για τηλεκίνηση RC αμαξιδίου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΑΡΤΕΜΙΔΟΣ ΖΩΓΡΑΦΟΥ

Επιβλέπων: Αντρέας Σταφυλοπάτης

Συνεπιβλέπων: Γεώργιος Σιόλας

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 27 Μαρτίου 2018.

(Υπογραφή)

.....

Ανδρέας Σταφυλοπάτης

Καθηγητής

Ε.Μ.Π.

(Υπογραφή)

.....

Γιώργος Στάμου

Αναπληρωτής καθηγητής

Ε.Μ.Π.

(Υπογραφή)

.....

Κωνσταντίνα Νικήτα

Καθηγήτρια

Ε.Μ.Π.

(Υπογραφή)

.....

Άρτεμις Ζωγράφου

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π

Copyright © Ζωγράφου Άρτεμις, 2018

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, η αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να αναφέρονται προς το συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν το συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Εργαστήριο Ευφυσών Συστημάτων

Ζωγράφου 157 73, Αθήνα

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον κύριο Σταφυλοπάτη που μου επέτρεψε να εκπονήσω την διπλωματική μου εργασία στο εργαστήριο Ευφυών Συστημάτων (Intelligent Systems Lab - ISL) καθώς και για την άψογη συνεργασία του. Ιδιαίτερως εκφράζω τις ευχαριστίες μου στον κύριο Σιόλα, οι συμβουλές του οποίου ήταν καταλυτικής σημασίας για την εκπόνηση της εργασίας αυτής. Ήταν πάντοτε διαθέσιμος και ιδιαίτερα βοηθητικός σε όλη την διάρκεια της διπλωματικής μου εργασίας και τον ευχαριστώ πολύ. Επίσης, θα ήθελα να ευχαριστήσω την κυρία Νικήτα και τον κύριο Στάμου, που μου έκαναν την τιμή να είναι μέλη της τριμελούς επιτροπής της διπλωματικής μου εργασίας.

Ένα μεγάλο ευχαριστώ θέλω να πω στην οικογένεια μου και ιδιαίτερα στους γονείς μου, Άγι και Όλγα και στην αδελφή μου Δάφνη για την στήριξη και υπομονή τους όλα αυτά τα χρόνια, σε ότι στόχους και αν έθετα. Τέλος, μιας και τελειώνει ένα πολύ όμορφο ταξίδι δεν θα μπορούσα να παραλείψω τους φίλους και τις φίλες μου, με τους οποίους περάσαμε όλα αυτά τα χρόνια και δημιουργήσαμε πολλές όμορφες αναμνήσεις μαζί.

στην γιαγιά μου Λαμπρινή και στον παππού μου Μιχάλη

Περίληψη

Στη διπλωματική εργασία αυτή μελετήθηκαν τα EEG και EOG σήματα του ανθρώπινου εγκεφάλου με στόχο την δημιουργία μίας διεπαφής ανθρώπου υπολογιστή (BCI: Brain Computer Interface). Συγκεκριμένα, ερευνήθηκαν τα σήματα εκείνα που σχετίζονται με τις κινήσεις των ματιών. Η διεπαφή που δημιουργήθηκε αποτελεί μία εφαρμογή πραγματικού χρόνου με την οποία ο χρήστης έχει την δυνατότητα βασικού τηλεχειρισμού ενός μικρού αμαξιού μόνο με τις κινήσεις των ματιών του. Το σύστημα που δημιουργήθηκε αποτελείται από το σύστημα απόφασης και την σύνδεση του με το αμαξάκι. Όσον αφορά το σύστημα απόφασης αποτελείται συνολικά από 4 στάδια: την καταγραφή των σημάτων, την προεπεξεργασία τους, την εξαγωγή των κατάλληλων χαρακτηριστικών από αυτά και την ταξινόμηση τους. Ο διαχωρισμός των σημάτων γίνεται σε 4 κλάσεις, ανάλογα με την κίνηση του ματιού που έχει εκτελέσει ο χρήστης, ανοιχτά/κλειστά μάτια και οριζόντια κίνηση των ματιών δεξιά/ αριστερή. Για τον σκοπό αυτό υλοποιήθηκε μία αλυσίδα από 3 SVM, καθένα εκπαιδευμένο για την ταξινόμηση διαφορετικής κίνησης. Στην συνέχεια, οι 4 αυτές κλάσεις αντιστοιχούνται μέσω κατάλληλης διεπαφής σε 4 απλές κινήσεις του μικρού αμαξιού. Κατά συνέπεια φαίνεται ότι, το προαναφερόμενο σύστημα έχει την δυνατότητα, ανάλογα με τα εισερχόμενα δεδομένα, να αποφασίζει ποια κίνηση έχει εκτελεστεί από τον χρήστη και να την μεταφράζει στην αντίστοιχη κίνηση του αμαξιού.

Λέξεις κλειδιά

Ηλεκτροεγκεφαλογράφημα, EEG, EOG σήματα, Μηχανική Μάθηση, Ταξινόμηση, Διεπαφή ανθρώπου - υπολογιστή

Abstract

In this diploma thesis, we study the EEG and EOG signals of the human brain in order to create a Brain - Computer Interface (BCI). In particular, the signals that were examined are those related to eye movement. The proposed interface constitutes a real-time application, which enables the user to control a demo car only via eye movements. The whole system consists of 2 main parts: the decision making subsystem and the connection between the latter and the demo car. The decision making system is made up of 4 stages: the recording of the signals, the preprocessing of the signals, the feature extraction and finally the classification. More specifically, the classification is conducted in 4 classes, corresponding to a particular eye movement: open eyes/close eyes, right and left movement of the eyes. For this purpose a chain of 3 SVMs is implemented, each one trained to recognize a specific movement. Subsequently, the 4 aforementioned classes are matched to 4 simple movements of the demo car. Thus it is clear that the aforementioned system is capable of deciding which eye movement has been made, based on the input data, and translate it to a corresponding action of the car.

Keywords

EEG, EEG signals, EOG signals, Signal Preprocessing, Machine Learning, Brain Computer - Interface (BCI), Classification, application demo car

Περιεχόμενα

Ευχαριστίες	9
Περίληψη	11
Abstract	13
1 Κεφάλαιο : Εισαγωγή.....	19
1.1 Κίνητρα για την ανάπτυξη διεπαφών ανθρώπου – υπολογιστή.....	19
1.2 Αντικείμενο και συνεισφορά της εργασίας.....	19
1.3 Διάρθρωση του κειμένου.....	20
2 Κεφάλαιο: Μία ανασκόπηση των βιοσημάτων	21
2.1 Βιοσήματα (Biosignals).....	21
2.2 EEG	21
2.3 Σύστημα 10-20.....	21
2.4 Ηλεκτροεγκεφαλικά σήματα (EEG signals).....	23
2.4.1 SSVEP (Steady State Visual Evoked Potential)	25
2.4.2 P300.....	27
2.4.3 MI (Motor Imagery).....	29
2.5 EOG signals	31
2.6 Hybrid BCI system	34
3 Κεφάλαιο: Θεωρητικό υπόβαθρο	35
3.1 Προεπεξεργασία – Αποθоруβοποίηση (Preprocessing).....	35
3.1.1 Βαθυπερατό φίλτρο (Low pass filter).....	35
3.1.2 Ζωνοπερατό φίλτρο (Bandpass filter)	35
3.1.3 Φίλτρο διαμέσου τιμής (Median filter).....	36
3.1.4 Wiener filter	36
3.2 Εξαγωγή χαρακτηριστικών (feature extraction).....	36
3.2.1 Δειγματοληψία	36
3.2.2 DFT (Discrete Fourier Transform).....	37
3.2.3 Power Spectral Density (PSD)	38
3.3 Μηχανική Μάθηση	39
3.3.1 Επιβλεπόμενη μάθηση (Supervised learning)	39
3.3.2 Αλγόριθμος KNN (K- Nearest neighbors).....	39

3.3.3	Τεχνητά Νευρωνικά Δίκτυα (Neural Networks).....	39
3.3.4	SVM (Support Vector Machines)	42
3.3.5	SVM ως μηχανή πυρήνα	44
3.4	Επαλήθευση με την μέθοδο Cross – Validation	45
4	Κεφάλαιο: Περιγραφή συστήματος.....	47
4.1	Πειραματικό μέρος	47
4.1.1	Καταγραφή δεδομένων (Recording)	47
4.1.2	Experimental setup.....	49
4.2	Προεπεξεργασία και εξαγωγή χαρακτηριστικών.....	52
4.2.1	Κανάλια O1-O2	52
4.2.2	Κανάλια F7-F8	57
4.3	Δημιουργία δεδομένων εκπαίδευσης (training dataset).....	61
4.4	Artefacts (Σφάλματα)	62
4.5	Μηχανική Μάθηση	63
4.5.1	Αλγόριθμος KNN.....	63
4.5.2	SVM.....	64
4.5.3	Neural Networks.....	65
5	Κεφάλαιο: Συνολικό σύστημα απόφασης.....	67
5.1	Γενική περιγραφή συστήματος.....	67
5.1.1	1° SVM.....	68
5.1.2	2° SVM.....	69
5.1.3	3° SVM.....	71
5.2	Εύρεση βέλτιστου SVM	73
5.3	Σύστημα πραγματικού χρόνου (Online system)	73
5.3.1	Βασικός περιορισμός.....	74
5.4	Πειραματικές δοκιμές	76
6	Κεφάλαιο: Υλοποίηση σε RC car.....	78
6.1	Σχεδιασμός demo car	78
6.1.1	Περιγραφή λειτουργίας.....	79
6.2	Αντιστοιχία κινήσεων ματιών - demo car	80
6.3	Πειραματικές δοκιμές	80
7	Κεφάλαιο : Συμπεράσματα και μελλοντική εργασία	81

7.1	Σύνοψη	81
7.2	Μελλοντική εργασία	82
	Βιβλιογραφία	84

Κατάλογος σχημάτων

Εικόνα 1.1:	10-20 Placement system	22
Εικόνα 1.2:	Electrodes placement	22
Εικόνα 1.3:	Brain regions and functions	23
Εικόνα 1.4:	Εγκεφαλικοί ρυθμοί	24
Εικόνα 1.5:	BCI system using SSVEP	26
Εικόνα 1.6:	ERP	27
Εικόνα 1.7:	P300 Speller	28
Εικόνα 1.8:	Περιοχές εγκεφάλου που διεγείρονται από διάφορες κινήσεις μυών	29
Εικόνα 1.9:	Γραφικό περιβάλλον για την εκπαίδευση BCI συστήματος	30
Εικόνα 1.10:	Γραφική απεικόνιση ματιού	31
Εικόνα 1.11:	Μύες κίνησης ματιού	31
Εικόνα 1.12:	Κανάλια μέτρησης EEG σημάτων που σχετίζονται με τα μάτια	32
Εικόνα 1.13:	Τοποθέτηση EOG ηλεκτροδίων	33
Εικόνα 1.14:	Γωνία κίνησης ματιών	33
Εικόνα 2.1:	Απόκριση βαθυπερατού φίλτρου	35
Εικόνα 2.2:	Απόκριση ζωνοπερατού φίλτρου στο πεδίο της συχνότητας	35
Εικόνα 2.3:	Σχηματική περιγραφή FT	37
Εικόνα 2.4:	Διαδικασία δειγματοληψίας α) αρχικό σήμα β) ικανοποίηση της συνθήκης γ) aliasing	38
Εικόνα 2.5:	Μονοστρωματικός νευρώνας	40
Εικόνα 2.6:	Multi Layer Perceptron	41
Εικόνα 2.7:	Σχηματική αναπαράσταση του σκεπτικού ενός βέλτιστου υπερεπιπέδου	43
Εικόνα 2.8:	Μετασχηματισμός δεδομένων	45
Εικόνα 2.9:	Πυρήνες διανυσμάτων υποστήριξης	45
Εικόνα 2.10:	Μέθοδος cross validation	46
Εικόνα 3.1:	Στάδια υλοποίησης της εφαρμογής	47
Εικόνα 3.2:	EMOTIV EPOC	48
Εικόνα 3.3:	Ηλεκτρόδια EPOC	48
Εικόνα 3.4:	Τοποθέτηση και προσαρμογή συσκευής	49
Εικόνα 3.5:	Αλληλουχία κινήσεων GUI	50
Εικόνα 3.6:	Αρχική οθόνη διεπαφής	51
Εικόνα 3.7:	GUI: Κίνηση προς τα δεξιά	51
Εικόνα 3.8:	GUI: Κίνηση προς τα αριστερά	52
Εικόνα 3.9:	Αρχικά σήματα α) πεδίο του χρόνου	53
Εικόνα 3.10:	Αρχικά σήματα β) πεδίο της συχνότητας	54

Εικόνα 3.11: Φιλτραρισμένα σήματα α) πεδίο του χρόνου	54
Εικόνα 3.12: Φιλτραρισμένα σήματα β) πεδίο της συχνότητας	55
Εικόνα 3.13: Αρχικά σήματα α) πεδίο του χρόνου	55
Εικόνα 3.14: Αρχικά σήματα β) πεδίο της συχνότητας	56
Εικόνα 3.15: Φιλτραρισμένα σήματα O1, O2 α) πεδίο του χρόνου	56
Εικόνα 3.16: Φιλτραρισμένα σήματα O1, O2 β) πεδίο της συχνότητας	57
Εικόνα 3.17: Αρχικά σήματα των F7, F8 καναλιών	58
Εικόνα 3.18: Αποτέλεσμα εφαρμογής ζωνοπερατού φίλτρου στα κανάλια F7, F8	58
Εικόνα 3.19: Αποτέλεσμα εφαρμογής median φίλτρου στα κανάλια F7, F8	59
Εικόνα 3.20: Αποτέλεσμα εφαρμογής wiener φίλτρου στα κανάλια F7, F8	60
Εικόνα 3.21: Αποτέλεσμα εφαρμογής συνδυασμού βαθυπερατού και wiener φίλτρου στα κανάλια F7, F8	60
Εικόνα 3.22: Διαφορά των καναλιών F7, F8	61
Εικόνα 3.23: Πίνακας δεδομένων εκπαίδευσης	61
Εικόνα 4.1: Αρχιτεκτονική συστήματος απόφασης	67
Εικόνα 4.2: Χαρακτηριστικά δεδομένων εκπαίδευσης 1ου SVM	68
Εικόνα 4.3: Αρχικό σήμα 512 δειγμάτων	69
Εικόνα 4.4: Διαδικασία εύρεσης παραθύρου στο αρχικό σήμα	70
Εικόνα 4.5: Χαρακτηριστικά των δεδομένων εκπαίδευσης του 2ου SVM	71
Εικόνα 4.6: Αρχικά σήματα F7-F8 καναλιών (δεξιά-αριστερή κίνηση)	72
Εικόνα 4.7: Τελικά σήματα των F7-F8 (128 δείγματα)	72
Εικόνα 4.8: Αρχιτεκτονική online συστήματος	74
Εικόνα 5.1: Σύνδεση εξαρτημάτων	78
Εικόνα 5.2: Τιμές των pins για την κίνηση του αμαξιού	79

Κατάλογος πινάκων

Πίνακας 1: Παράμετροι SVM	64
Πίνακας 2: Βέλτιστο SVM	65
Πίνακας 3: Παράμετροι Νευρωνικού Δικτύου	65
Πίνακας 4: Βέλτιστο Νευρωνικό Δίκτυο	66
Πίνακας 5: Αποτελέσματα πειραματισμού διαφορετικών παραθύρων	75
Πίνακας 6: Αποτελέσματα δοκιμών των 4 κινήσεων	76
Πίνακας 7: Πίνακας συνδυαστικών αποτελεσμάτων	77
Πίνακας 8: Αντιστοιχία κινήσεων ματιών-demo car	80

1 Κεφάλαιο : Εισαγωγή

1.1 Κίνητρα για την ανάπτυξη διεπαφών ανθρώπου – υπολογιστή

Είναι γεγονός ότι ο εγκέφαλος είναι ένα μυστήριο που απασχολεί επιστήμονες πολλών ειδικοτήτων και η αποκωδικοποίηση του αποτελεί πρόκληση. Σήμερα ο αυξανόμενος αριθμός ερευνητικών μελετών σε συστήματα διεπαφής εγκεφάλου – υπολογιστή (BCI: Brain Computer Interface), καταδεικνύει ιδιαίτερα ενδιαφέρουσες δυνατότητες για αυτόν τον νέο ερευνητικό τομέα. Το BCI διαμορφώνει νέα σύνορα στην επιστήμη και την τεχνολογία διότι συνδυάζει δεξιότητες από πολλούς τομείς όπως η νευροεπιστήμη, η μηχανική, επιστήμη των υπολογιστών, η ψυχολογία και η κλινική αποκατάσταση. Πολλές έρευνες έχουν συνταχθεί, οι οποίες προσπαθούν να αποκρυπτογραφήσουν την λειτουργία του εγκεφάλου, μέσω μελέτης της εγκεφαλικής δραστηριότητας. Η εγκεφαλική δραστηριότητα καταγράφεται με πολλές μεθόδους, όπως για παράδειγμα το ηλεκτροεγκεφαλογράφημα (EEG) και το μαγνητοεγκεφαλογράφημα (MEG). Πιο διαδεδομένη στις εφαρμογές διεπαφής εγκεφάλου – υπολογιστή είναι η πρώτη μέθοδος, λόγω του εξοπλισμού που χρειάζεται για την καταγραφή, ο οποίος είναι πιο ελαφρύς και φορητός. Ένα σύστημα διασύνδεσης εγκεφάλου - υπολογιστή (BCI) παρέχει έναν τρόπο αξιοποίησης των σημάτων του εγκεφάλου, για την δημιουργία ενός διαφορετικού τρόπου επικοινωνίας του ανθρώπου με το περιβάλλον του. Τέτοια συστήματα είναι ιδιαίτερα χρήσιμα και ελπιδοφόρα για ανθρώπους που δεν μπορούν να αλληλοεπιδράσουν με το περιβάλλον τους με τον συνηθισμένο τρόπο, όπως για παράδειγμα άτομα με αναπηρία ή μεγαλύτερης ηλικίας. Παράλληλα, η έρευνα στα BCI συστήματα έχει χρησιμοποιηθεί με επιτυχία και για υγιείς ανθρώπους, ως πρόσθετο κανάλι εισόδου δεδομένων σε διάφορες εφαρμογές όπως για παράδειγμα στον έλεγχο παιχνιδιών, στις εφαρμογές επαυξημένης πραγματικότητας, για τον έλεγχο ρομπότ και σε πολλές άλλες.

1.2 Αντικείμενο και συνεισφορά της εργασίας

Αυτή η διπλωματική εργασία μελετά την αξιοποίηση των EEG και EOG σημάτων με στόχο την δημιουργία μίας διεπαφής εγκεφάλου – υπολογιστή. Πιο συγκεκριμένα ερευνήθηκαν τα σήματα εκείνα που συσχετίζονται με την κίνηση των ματιών. Για τον σκοπό αυτό χρησιμοποιήθηκαν EEG και EOG σήματα. Χρησιμοποιώντας τον συνδυασμό των 2 αυτών κατηγοριών, υλοποιήθηκε μία hybrid εφαρμογή πραγματικού χρόνου, με την οποία είναι δυνατός ο βασικός έλεγχος ενός μικρού demo car, μόνο με τις κινήσεις των ματιών. Οι κινήσεις των ματιών που μελετήθηκαν ήταν 4, ανοιχτά/κλειστά μάτια και αριστερή/δεξιά κίνησή τους. Πιο αναλυτικά το συνολικό σύστημα αποτελείται από δύο επιμέρους μέρη: α) από το σύστημα απόφασης και β) την σύνδεση του με το μικρό αμαξίδιο. Όσων αφορά το σύστημα απόφασης, αποτελείται από 4 βασικά στάδια:

- Καταγραφή σημάτων: χρησιμοποιήθηκε η συσκευή EPOC της εταιρείας EMOTIV.
- Προεπεξεργασία σημάτων: εφαρμόστηκαν διάφοροι τρόποι φιλτραρίσματος με σκοπό την αποθρομβοποίηση τους.

- Εξαγωγή χαρακτηριστικών: δοκιμάστηκαν διάφορες μέθοδοι για την εξαγωγή και έπειτα την επιλογή των κατάλληλων χαρακτηριστικών της κάθε κίνησης.
- Ταξινόμηση: κατασκευάστηκε μία αλυσίδα 3 κατάλληλα εκπαιδευμένων SVM, με σκοπό των διαχωρισμό των προαναφερθέντων 4 κινήσεων. Αναλυτικότερα, το πρώτο SVM λαμβάνει την απόφαση για το αν τα δεδομένα εισόδου ανήκουν στην κατηγορία των ανοιχτών ή κλειστών ματιών. Εάν αποφασιστεί ότι ανήκουν στην πρώτη κλάση, τότε ο έλεγχος μεταβαίνει στο δεύτερο SVM, το οποίο αποφασίζει για το αν υπάρχει κίνηση ή όχι των ματιών. Ακολουθώντας την ίδια λογική εάν το 2ο SVM βγάλει την απόφαση ότι υπάρχει κίνηση, ο έλεγχος μεταβαίνει στο 3ο και τελευταίο SVM της αλυσίδας το οποίο ταξινομεί την κίνηση σε δεξιά και αριστερή. Τέλος, σημειώνεται ότι τα δεδομένα με τα οποία εκπαιδεύτηκαν τα SVM, καταγράφηκαν ομοιόμορφα μέσω ενός πρωτοκόλλου που δημιουργήθηκε για τον σκοπό αυτό.

Σχετικά με την σύνδεση του συστήματος απόφασης που αναφέρθηκε προηγουμένως και του demo car, υλοποιήθηκε μία διεπαφή, έτσι ώστε να γίνει εφικτή η αντιστοίχιση των κινήσεων των ματιών με εκείνων του αμαξιδίου, σε πραγματικό χρόνο, δηλαδή όταν λαμβάνεται η απόφαση από την αλυσίδα των SVM. Συμπερασματικά, βλέπουμε ότι η εφαρμογή αυτή επιτρέπει στον χρήστη να ελέγχει μόνο μέσω της κίνησης των ματιών του ένα μικρό αμαξίδιο. Κάτι τέτοιο είναι ιδιαίτερο χρήσιμο για άτομα με αναπηρία, όπως για παράδειγμα άτομα που πάσχουν από Πλάγια Μυατροφική Σκλήρυνση (ALS).

1.3 Διάρθρωση του κειμένου

Η δομή του κειμένου της διπλωματικής εργασίας είναι η εξής:

Στο Κεφάλαιο 2 παρουσιάζεται μία επισκόπηση του ερευνητικού πεδίου των βιοσημάτων και ο τρόπος με τον οποίο χρησιμοποιούνται σε εφαρμογές διεπαφής εγκεφάλου - υπολογιστή.

Στο Κεφάλαιο 3 παρουσιάζεται το θεωρητικό υπόβαθρο που απαιτείται για την εκπόνηση της συγκεκριμένης διπλωματικής εργασίας, σχετικά με τους τρόπους προεπεξεργασίας των σημάτων αλλά και οι απαραίτητες γνώσεις από τον τομέα της Μηχανικής Μάθησης.

Στο Κεφάλαιο 4 περιλαμβάνεται μία εκτενής περιγραφή της διαδικασίας που ακολουθείται για την δημιουργία του συστήματος, μαζί με τις πρώτες πειραματικές δοκιμές.

Στο Κεφάλαιο 5 παρουσιάζεται το συνολικό σύστημα που υλοποιήθηκε μαζί με τα συμπεράσματα των πειραματικών δοκιμών.

Στο Κεφάλαιο 6, περιγράφεται η κατασκευή και η σύνδεση του αμαξιδίου (demo car) με το προαναφερθέν σύστημα απόφασης.

Στο Κεφάλαιο 7 συνοψίζονται τα συμπεράσματα της εργασίας και τίθενται κατευθυντήριες γραμμές για την μελλοντική περαιτέρω ανάπτυξη της εφαρμογής που δημιουργήθηκε.

2 Κεφάλαιο: Μία ανασκόπηση των βιοσημάτων

2.1 Βιοσήματα (Biosignals)

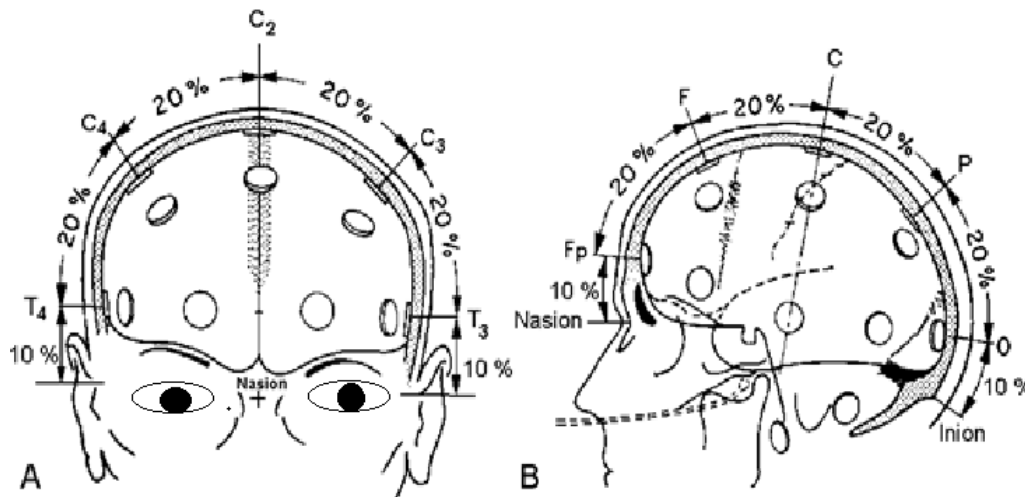
Ως biosignal ορίζεται οποιοδήποτε σήμα, το οποίο περιγράφει κάποιο φυσιολογικό φαινόμενο του οργανισμού και μπορεί να παρατηρηθεί και να μετρηθεί οποιαδήποτε στιγμή. Ο όρος αυτός αναφέρεται συνήθως σε βιοηλεκτρικά σήματα (bioelectrical signals). Τα bioelectrical signals σχετίζονται με την αλλαγή του ρεύματος η οποία παράγεται από την συνολική διαφορά δυναμικού σε ένα συγκεκριμένο ιστό, οργανικό ή κυτταρικό σύστημα, όπως για παράδειγμα το νευρικό σύστημα. Έτσι, μεταξύ των πολλών αυτών σημάτων, τα πιο γνωστά είναι τα ακόλουθα: EEG (electroencephalography), EOG (electrooculography), EMG (electromyography) και ECG (Electrocardiography). Στην συνέχεια θα αναλύσουμε τα δύο πρώτα είδη σημάτων, καθώς με αυτά ασχολείται και η παρούσα διπλωματική.

2.2 EEG

Η ηλεκτροεγκεφαλογραφία (EEG: electroencephalography) είναι η μέθοδος της καταγραφής της ηλεκτρικής δραστηριότητας του εγκεφάλου, όπως αυτή σχηματίζεται από την επαλληλία των ηλεκτρικών ρευμάτων που παράγονται από τους νευρώνες του εγκεφάλου και πολλά ακόμη νευρικά κύτταρα του Κεντρικού Νευρικού Συστήματος (Central Nervous System - CNS). Το ηλεκτροεγκεφαλογράφημα προκύπτει μετρώντας τις διακυμάνσεις της τάσης που οφείλονται στο ιοντικό ρεύμα που διέρχεται από τους νευρώνες του εγκεφάλου.

2.3 Σύστημα 10-20

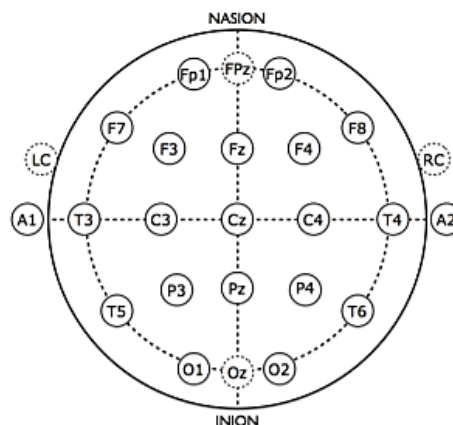
Η μέθοδος καταγραφής των ηλεκτροεγκεφαλογραφημάτων είναι συνήθως μη επεμβατική, και λαμβάνει χώρα στην εξωτερική δερματική επιφάνεια του ανθρώπινου κρανίου, με τα ηλεκτρόδια τοποθετημένα κατά μήκος του. Πιο συγκεκριμένα, τοποθετούνται στην δερματική επιφάνεια του κεφαλιού του χρήστη τα ηλεκτρόδια σε συγκεκριμένες θέσεις, μαζί με κάποια ειδική αλοιφή ή κάποιο αλατούχο διάλυμα, το οποίο βοηθάει στην αύξηση της αγωγιμότητας μεταξύ του δέρματος του κρανίου και του ηλεκτροδίου. Το τελευταίο έχει ως αποτέλεσμα την καλύτερη επαφή και επομένως μετάδοση των EEG σημάτων μέσω αυτών. Τα ονόματα και οι θέσεις των ηλεκτροδίων δεν είναι αυθαίρετα, αλλά καθορίζονται από το Διεθνές Σύστημα Τοποθέτησης 10-20 (International 10-20 Placement System) για τις περισσότερες κλινικές και ερευνητικές εφαρμογές, έτσι ώστε να διασφαλίζεται η κοινή ονομασία των ηλεκτροδίων καθώς και η δυνατότητα αναπαραγωγής των μελετών μεταξύ διαφορετικών εργαστηρίων, χωρίς να υπάρχουν προβλήματα συνέχειας.



Εικόνα 2.1: 10-20 Placement system

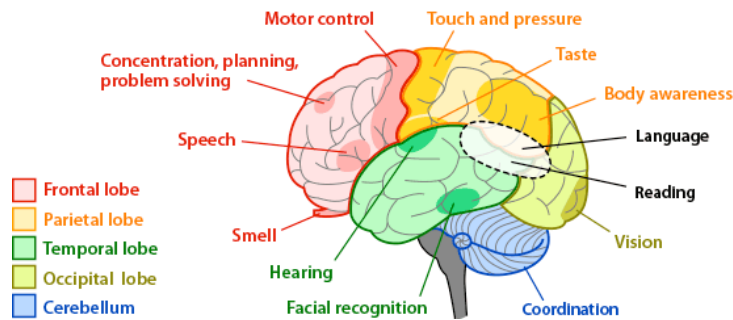
Πιο αναλυτικά το συγκεκριμένο σύστημα βασίζεται στη σχέση μεταξύ της θέσης ενός ηλεκτροδίου και της υποκείμενης περιοχής του εγκεφαλικού φλοιού. Κάθε τοποθεσία του ηλεκτροδίου έχει ένα γράμμα για να αναγνωρίζεται σε ποιο λοβό βρίσκεται και έναν αριθμό για να προσδιορίσει τη θέση του στο ημισφαίριο. Ειδικότερα :

- Το '10' και το '20' στο όνομα του συστήματος αναφέρονται στο γεγονός ότι η απόσταση μεταξύ γειτονικών ηλεκτροδίων αντιστοιχεί είτε στο 10% είτε στο 20% της συνολικής απόστασης του κρανίου από το μπροστινό έως και το πίσω αλλά και από το δεξί έως και το αριστερό μέρος (front - back and right - left distance).
- Δύο ανατομικά ορόσημα χρησιμοποιούνται για την τοποθέτηση των ηλεκτροδίων EEG. Πρώτον, το ριζορρίνιο (nasion) που είναι η περιοχή μεταξύ των ματιών, ακριβώς πάνω από τη γέφυρα της μύτης. Δεύτερον, το ίνιο (inion), το οποίο είναι το χαμηλότερο σημείο στο πίσω μέρος του κεφαλιού και συνήθως υποδεικνύεται από ένα προεξέχον "βουναλάκι" του κρανίου.



Εικόνα 2.2: Electrodes placement

- Τα γράμματα F, T, C, P, O αντιπροσωπεύουν τους μετωπικούς (Frontal), κροταφικούς (Temporal), κεντρικούς (Central), βρεγματικούς (Parietal) και ινιακούς (Occipital) λοβούς αντίστοιχα. Σημειώνεται ότι δεν υπάρχει κεντρικός λοβός, το γράμμα "C" χρησιμοποιείται μόνο για σκοπούς αναγνώρισης. Παρακάτω φαίνεται μία αναπαράσταση των περιοχών του εγκεφάλου και των αντίστοιχων λειτουργιών του.



Εικόνα 2.3: Brain regions and functions

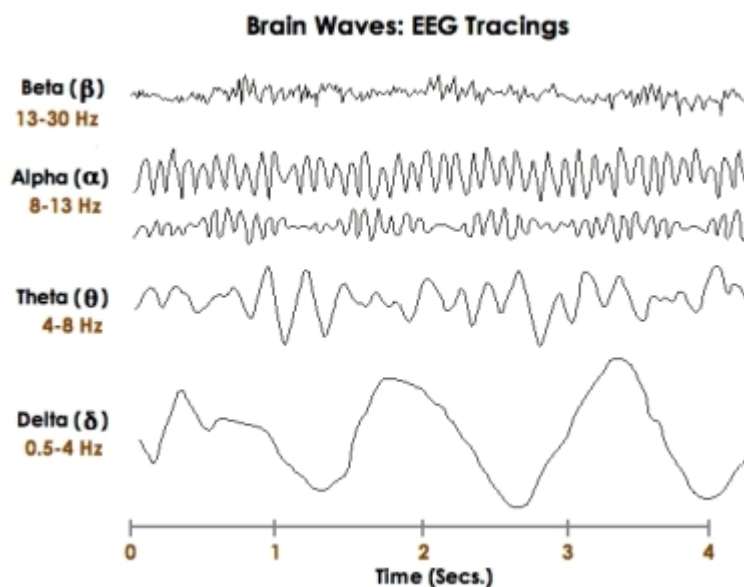
- Οι άρτιοι αριθμοί (2, 4, 6, 8) αναφέρονται σε θέσεις ηλεκτροδίων στο δεξιό ημισφαίριο, ενώ οι περιττοί αριθμοί (1, 3, 5, 7) αναφέρονται σε θέσεις του αριστερού ημισφαιρίου. Το γράμμα 'z' αναγράφεται στους αισθητήρες εκείνους που τοποθετούνται στην μεσαία γραμμή. Εκτός από αυτές τις ονομασίες, οι κωδικοί των χαρακτήρων A, P_g και F_p αναφέρονται στους λοβούς του αυτιού, στις ρινοφαρυγγικές και στις εμπρόσθιες πολικές περιοχές αντίστοιχα.

Όπως βλέπουμε στο σύστημα 10-20, τοποθετούνται 21 ηλεκτρόδια. Αν θέλαμε μία πιο λεπτομερή καταγραφή των EEG σημάτων, επιπλέον ηλεκτρόδια τοποθετούνται με βάση την 10% διαίρεση. Πιο συγκεκριμένα τα επιπρόσθετα ηλεκτρόδια εφαρμόζονται ακριβώς ανάμεσα από τα υπάρχοντα του συστήματος 10-20.

2.4 Ηλεκτροεγκεφαλικά σήματα (EEG signals)

Αποτέλεσμα της καταγραφής με την μέθοδο της ηλεκτροεγκεφαλογραφίας είναι τα EEG σήματα. Αναλυτικότερα, τα ηλεκτροεγκεφαλικά σήματα είναι στοχαστικά, θορυβώδη και αρκετά ασθενή. Γι' αυτό τον λόγο χαρακτηρίζονται από μικρό σηματοθορυβικό λόγο (SNR) και χαμηλό πλάτος, το οποίο κυμαίνεται περίπου από 1μV έως 100μV. Το φάσμα των ηλεκτροεγκεφαλικών σημάτων εντοπίζεται στο εύρος 0 έως 100 Hz, ενώ διακρίνονται συγκεκριμένες κυματομορφές βάσει των συχνοτήτων των αρμονικών από τις οποίες αποτελούνται, δηλαδή βάσει του φασματικού περιεχομένου τους. Οι κυματομορφές αυτές είναι γνωστές ως **εγκεφαλικοί ρυθμοί** (brain rhythms) και έχουν ταξινομηθεί με βάση κάποια κοινά χαρακτηριστικά, όπως το εύρος συχνοτήτων και την χωρική τους κατανομή. Με αυτό τον τρόπο, χωρίζονται σε 4 θεμελιώδεις κατηγορίες: **δέλτα** (0.5-4 Hz), **θήτα** (4-8 Hz), **άλφα** (8-13Hz), **βήτα** (13-30 Hz). Πέρα από τις παραπάνω βασικές ζώνες, φαίνεται πως το φάσμα του ηλεκτροεγκεφαλογραφήματος αποκαλύπτει και μια ακόμη κατηγορία κυμάτων, την **γάμμα** (30-70 Hz), την οποία πολλοί

ερευνητές κατατάσσουν στην ευρύτερη κατηγορία των βήτα. Τέλος, σημειώνεται ότι στην ίδια ζώνη συχνοτήτων με τα άλφα κύματα έχουμε και τα **κύματα μ** (mu ή sensorimotor rhythm), τα οποία θα μας απασχολήσουν στην συνέχεια. Πιο συγκεκριμένα, όσον αφορά τα δέλτα κύματα, εμφανίζονται κυρίως στον μετωπιαίο λοβό και σχετίζονται με στάδια βαθύ ύπνου. Σχετικά με το εύρος συχνοτήτων θήτα, εντοπίζεται σε καταστάσεις υπνηλίας. Οι κυματομορφές της ζώνης άλφα, εκφράζουν κυρίως την δραστηριότητα του ινιακού λοβού του εγκεφάλου. Εμφανίζονται σε κατάσταση χαλάρωσης και έξαρσης, καθώς επίσης μεταβάλλονται και με το άνοιγμα και κλείσιμο των ματιών. Σε αντίθεση με τα κύματα άλφα, τα οποία επαναλαμβάνονται με παρόμοια συχνότητα στον οπτικό φλοιό, το κύμα mu εντοπίζεται πάνω από τον κινητικό φλοιό. Όσον αφορά τη ζώνη συχνοτήτων βήτα, παρατηρείται κυρίως στον πρόσθιο λοβό του κρανίου, και συνδέεται τόσο με την κινητική συμπεριφορά, αλλά και με την κατάσταση συγκέντρωσης ή εγρήγορσης του εγκεφάλου. Επιπλέον, έχει πρόσφατα ανακαλυφθεί ότι τα κύματα γάμμα σχετίζονται με την συνείδηση και την αντίληψη, κάτι ιδιαίτερα ενδιαφέρον.



Εικόνα 2.4: Εγκεφαλικοί ρυθμοί

Πέρα από τους εγκεφαλικούς ρυθμούς που έχουν μελετηθεί, έχει παρατηρηθεί ότι εγκεφαλικές διαδικασίες προκαλούνται και ως απόκριση σε συγκεκριμένα ερεθίσματα (stimuli), τα οποία μπορεί να είναι αισθητικά (sensory), κινητικά (motor) και διανοητικά (cognitive). Για παράδειγμα, στην άλφα μπάντα με το κλείσιμο των ματιών παρουσιάζεται αύξηση της φασματικής ισχύος, ενώ με το άνοιγμα τους παρουσιάζεται μείωση. Ορίζουμε έτσι ως **ERS (Event Related Synchronization)** φαινόμενα τα οποία συσχετίζονται με την αύξηση της ισχύος του EEG σήματος, και ως **ERD (Event Related Desynchronization)** τα γεγονότα εκείνα που σχετίζονται με την μείωση της ισχύος του. Τα ERS/ERD είναι σε συνάρτηση με το ερέθισμα κλειδωμένα ως προς τον χρόνο (time-locked), αλλά όχι και ως προς την φάση (phase-locked), και εμφανίζονται επίσης σε πολύ συγκεκριμένη περιοχή του φάσματος (band specific). Παράλληλα, μία άλλη κατηγορία των EEG σημάτων είναι τα **Βιωματικά Δυναμικά (Event Related Potentials - ERP)**, τα οποία χαρακτηρίζουν την απόκριση του εγκεφάλου στην εμφάνιση ή εξαφάνιση ενός ερεθίσματος. Τα

τελευταία χωρίζονται σε **Προκλητά (Evoked Potentials)** και **Εκπεμπόμενα (Emitted Potentials)**. Τα δυναμικά της πρώτης κατηγορίας χρησιμοποιούνται κατά κύριο λόγο στις διεπαφές ανθρώπου-υπολογιστή μέσω εγκεφάλου (BCI: Brain-Computer Interfaces), με κυριότερα τα **Οπτικά προκλητά δυναμικά (Visual Evoked Potentials - VEP)** και τα **Σωματοαισθητικά προκλητά δυναμικά (Somatosensory Evoked Potentials - SEP)**.

Παρακάτω θα αναλύσουμε τα τρία βασικά είδη που ανήκουν στην κατηγορία των ERP και χρησιμοποιούνται συχνότερα στις εφαρμογές διεπαφών ανθρώπου-υπολογιστή μέσω εγκεφάλου.

2.4.1 SSVEP (Steady State Visual Evoked Potential)

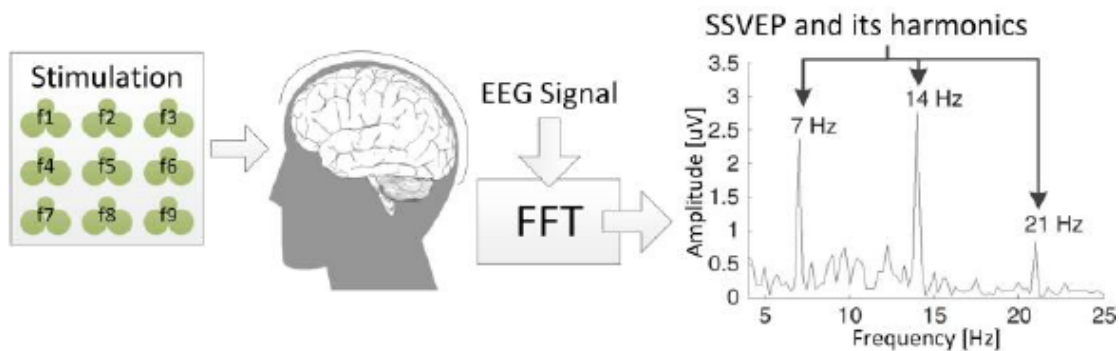
Τα SSVEP λαμβάνονται από την οπτική περιοχή του εγκεφάλου και αποτελούν την ανταπόκριση του πρωταρχικού οπτικού εγκεφαλικού λοβού σε κάποιο περιοδικό οπτικό ερέθισμα, σε συγκεκριμένες συχνότητες. Οι συχνότητες αυτές ποικίλουν από 1 έως 90 Hz, αλλά συνήθως χρησιμοποιείται το εύρος 4-60 Hz. Ο τρόπος δημιουργίας των SSVEP βασίζεται στο εξής: Όταν ο αμφιβληστροειδής διεγείρεται από ένα οπτικό ερέθισμα που κυμαίνεται στο παραπάνω φάσμα, ο εγκέφαλος παράγει ηλεκτρικά κύματα στην ίδια (ή πολλαπλάσια) συχνότητα του οπτικού ερεθίσματος. Η λήψη των δυναμικών αυτών γίνεται μέσω ηλεκτροδίων που τοποθετούνται στον ινιακό λοβό του κρανίου, επιλέγουμε συνήθως τα O1, O2 και Oz, διότι εμφανίζουν την μεγαλύτερη ισχύ στο συγκεκριμένο φάσμα.

Πειραματική διαδικασία

Παρακάτω θα περιγραφεί ένα απλουστευμένο σύστημα αλληλεπίδρασης εγκεφάλου υπολογιστή (BCI: Brain Computer Interface), το οποίο χρησιμοποιεί τα συγκεκριμένα σήματα. Αρχικά, ο χρήστης κεντράρει το βλέμμα του σε κάποιο στόχο, ο οποίος αποτελεί το οπτικό ερέθισμα. Ως στόχο μπορεί να έχουμε συστοιχίες από LED, πλέγματα ή σκακιέρες σε οθόνη υπολογιστή, που αναβοσβήνουν με κάποια συγκεκριμένη συχνότητα. Ανάλογα με την διαμόρφωση του stimulus, έχουμε τις εξής 3 κατηγορίες:

- f-VEP: Έχουμε συχνотική διαμόρφωση, με συχνότητες μοναδικές για κάθε στόχο.
- t-VEP: Το ερέθισμα διαμορφώνεται κατάλληλα στο πεδίο του χρόνου, με τις χρονικές ακολουθίες του κάθε στόχου, να είναι ορθογώνιες μεταξύ τους στον χρόνο.
- c-VEP: Στην κατηγορία αυτή, χρησιμοποιείται ψευδοτυχαία κωδικοποιημένη διαμόρφωση, δηλαδή ο κάθε στόχος είναι μία ψευδοτυχαία ακολουθία.

Στην συνέχεια, λαμβάνονται τα σήματα από τα συγκεκριμένα ηλεκτρόδια όπως προαναφέρθηκε και επεξεργάζονται κατάλληλα για την εξαίεψη του θορύβου. Τελευταίο στάδιο είναι η εξαγωγή των κατάλληλων χαρακτηριστικών, έτσι ώστε να εξαχθεί το φασματικό τους περιεχόμενο. Ανάλογα με την κατανομή της ισχύς του φάσματος που εξάγουμε, μπορούμε στην συνέχεια να συμπεράνουμε σε ποια συχνότητα έχει εστιάσει ο χρήστης.



Εικόνα 2.5: BCI system using SSVEP

Εφαρμογές

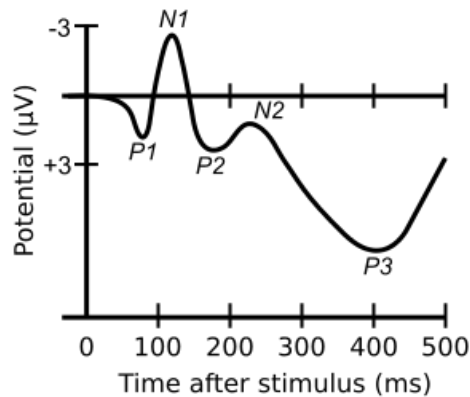
Βλέπουμε λοιπόν ότι κάθε **BCI σύστημα** αποτελείται απλοϊκά από τα εξής στάδια:

- i) Επιλογή EEG σήματος και ηλεκτροδίων λήψης (EEG paradigm)
- ii) Προεπεξεργασία για εξάλειψη θορύβου και σφαλμάτων (preprocessing methods)
- iii) Εξαγωγή χαρακτηριστικών (feature extraction methods)
- iv) Επιλογή μεθόδου ταξινόμησης (classification methods)

Έχει παρατηρηθεί ότι τα SSVEP χαρακτηρίζονται από πολύ καλό σηματοθορυβικό λόγο και από ατρωσία στον θόρυβο, γι' αυτό τον λόγο χρησιμοποιούνται ευρέως σε BCI εφαρμογές. Οι περισσότερες από αυτές τις εφαρμογές ασχολούνται με τον τηλεχειρισμό κάποιου είδους ρομπότ, το οποίο μπορεί για παράδειγμα να είναι είτε ανθρωπόμορφο ρομπότ (humanoid robot) είτε κάποιο ιπτάμενο ρομπότ όπως drone. Ο χειρισμός γίνεται αντιστοιχώντας τις διάφορες συχνότητες των EEG σημάτων στις οποίες κεντράρει το βλέμμα του ο χρήστης, σε αντίστοιχες εντολές κίνησης. Πιο συγκεκριμένα οι (Güneysu, 2013), (Zhao, 2014) και (Choi B. &, 2013) έχουν χρησιμοποιήσει τα SSVEP σήματα για τον χειρισμό του NAO, το οποίο είναι ένα humanoid robot. Όσον αφορά τον χειρισμό ιπτάμενων ρομπότ οι (Chen, 2016) και (Meriño, 2017) έχουν υλοποιήσει ανάλογες BCI εφαρμογές. Εκτός από τον χειρισμό ρομπότ, τα SSVEP χρησιμοποιούνται μέσω κατάλληλων εφαρμογών και για την χρήση του υπολογιστή, όπως για παράδειγμα ο (Hwang, 2012) έχει υλοποιήσει ένα σύστημα πληκτρολόγησης. Πέρα από την χρήση των SSVEP σημάτων που παράγονται από οπτικά ερεθίσματα λευκού φωτός, μία διαφορετική προσέγγιση που συναντιέται στην βιβλιογραφία είναι η χρήση χρωμάτων στα οπτικά ερεθίσματα. Μελέτες για την χρήση **RGB χρωμάτων** παρουσιάζονται στα άρθρα των (Mouli, 2013), (Mondada, 2016) και (Bekdash, 2015).

2.4.2 P300

Κλασικά ERP περιλαμβάνουν πολλά θετικά και αρνητικά κύματα, όπως τα P1, N1, P2, N2 και P3 (αλλιώς P300) σύμφωνα με τις προκληθείσες ακολουθίες και πολικότητες. Τα σήματα αυτά προκαλούνται μετά την εμφάνιση ενός ερεθίσματος, παρουσιάζοντας διαφορετικούς χρόνους εμφάνισης. Παρακάτω βλέπουμε ένα γράφημα των βασικότερων Event Related Potentials σε συνάρτηση με την καθυστέρηση εμφάνισης του καθενός.



Εικόνα 2.6: ERP

Παρόλο που το νευρικό υπόστρωμα του P300 σήματος παραμένει θολό, η αναπαραγωγικότητα και η πανταχού παρουσία του, το καθιστούν μία κοινή επιλογή σε εργαστηριακές μελέτες. Το P300 θεωρείται ότι είναι ένα ενδογενές δυναμικό, καθώς η εμφάνισή του δεν συνδέεται με τις φυσικές ιδιότητες ενός ερεθίσματος, αλλά με την αντίδραση του ατόμου σε αυτό. Πιο συγκεκριμένα, αντικατοπτρίζει τις διαδικασίες που εμπλέκονται στην αξιολόγηση ή ταξινόμηση των ερεθισμάτων. Η παρουσία, το μέγεθος, η τοπογραφία και ο συγχρονισμός αυτού του σήματος χρησιμοποιούνται συχνά ως μετρήσεις της γνωστικής λειτουργίας στις διαδικασίες λήψης αποφάσεων. Τα συγκεκριμένα σήματα έχουν σχετικά μεγάλο πλάτος που κυμαίνεται από τα 5-20µV και δημιουργούνται παρουσιάζοντας θετική αλλαγή της τάσης τους, μετά από παρουσία οπτικού ή ηχητικού ερεθίσματος με καθυστέρηση εμφάνισης συνήθως 300 ms (από όπου πήρε και το όνομά του) έως και 800 ms από το προκληθέν γεγονός. Η διαφορά στον χρόνο καθυστέρησης εξαρτάται από την δυσκολία της απόφασης. Τα ηλεκτρόδια από τα οποία λαμβάνονται τα συγκεκριμένα σήματα, βρίσκονται στην μέση του κεφαλιού και είναι συνήθως τα Pz, Fz, Cz και Oz.

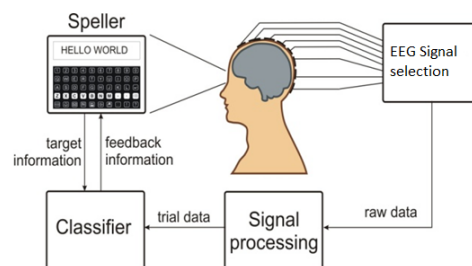
Πειραματική διαδικασία

Ένα σύστημα P300 χρησιμοποιεί συχνά ερεθίσματα με διαφορετικούς χαρακτήρες, περιεχόμενο και μεθόδους αποκωδικοποίησης για την εκτέλεση της αντίστοιχης γνωσιακής διαδικασίας του χρήστη, ανάλογα με τους σκοπούς της κάθε έρευνας. Σίγουρο είναι ότι σε κάθε τέτοιο σύστημα απαραίτητη είναι η χρήση ενός ερεθίσματος, συνήθως μίας οθόνης. Τα συστήματα P300

χρησιμοποιούν είτε μία συχνότητα για την εμφάνιση του ερεθίσματος είτε το oddball paradigm, όπου σύμφωνα με αυτό δύο ερεθίσματα παρουσιάζονται τυχαία στο χρήστη σε δύο διαφορετικές συχνότητες. Το 'κανονικό' ερέθισμα (standard stimulus) εμφανίζεται περιοδικά ενώ το δεύτερο ερέθισμα (bias stimulus) τυχαία. Το standard stimulus ονομάζεται ερέθισμα στόχου όταν το άτομο αντιδρά σε αυτό. Έτσι, το δυναμικό P300 θα παρατηρηθεί 300 ms μετά την εμφάνιση του ερεθίσματος στόχου.

Εφαρμογές

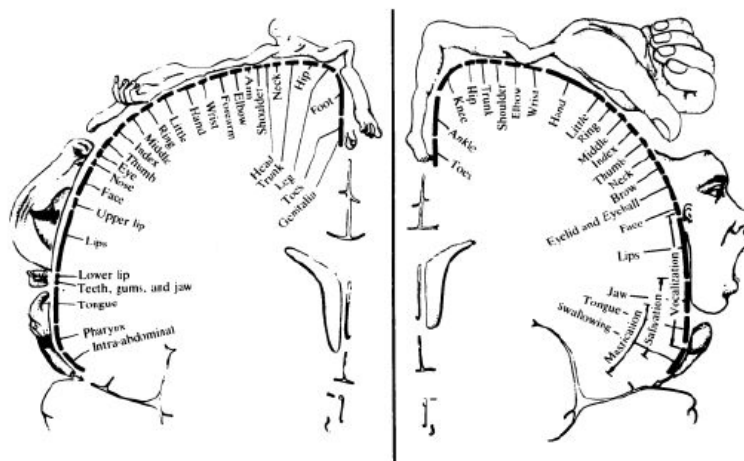
Το P300 έχει πολλές επιθυμητές ιδιότητες γι' αυτό και χρησιμοποιείται σε πολλές BCI εφαρμογές. Πρώτον, η κυματομορφή P300 είναι συνεχώς ανιχνεύσιμη και προκαλείται ως απόκριση σε ακριβή ερεθίσματα. Δεύτερον, τα P300 σήματα μπορούν να προκληθούν σε σχεδόν όλα τα άτομα με μικρές διακυμάνσεις στις τεχνικές μέτρησης, το οποίο συμβάλει στην απλοποίηση των εφαρμογών διεπαφής και επιτρέπει μεγαλύτερη χρηστικότητα. Η πιο συνήθης εφαρμογή είναι το P300 speller, για την γραφή κειμένου στον υπολογιστή. Από τους πρώτους και ο πιο διαδεδομένος αλγόριθμος που χρησιμοποιείται είναι αυτός του (Farwell, 1988), ο οποίος βασίζεται στις ασυναίσθητες διαδικασίες λήψης αποφάσεων του P300. Σύμφωνα με αυτόν, εμφανίζεται στον χρήστη ένα πλέγμα 6 × 6 χαρακτήρων και επισημαίνονται, δηλαδή αναβοσβήνουν, διάφορες στήλες ή σειρές με τυχαία σειρά. Ο χρήστης επιλέγει το γράμμα που επιθυμεί να γράψει κεντράροντας το βλέμμα του στη συγκεκριμένη θέση που βρίσκεται στο πλέγμα. Όταν μια στήλη ή μια σειρά περιέχει τον χαρακτήρα που το άτομο έχει κεντράρει το βλέμμα του, και αυτή επισημανθεί, έχουμε την απόκριση P300. Συνδυάζοντας την σειρά και την στήλη που προκάλεσε την απόκριση, εντοπίζουμε τον επιθυμητό χαρακτήρα. Σημειώνεται ότι ο αριθμός των δοκιμών πρέπει να γίνεται κατά μέσο όρο για να εξαλειφθεί ο θόρυβος που οφείλεται σε τυχαία γεγονότα. Η συχνότητα εμφάνισης των χαρακτήρων καθορίζει την ταχύτητα επεξεργασίας τους, η οποία είναι περίπου 3,4-4,3 χαρακτήρες / λεπτό. Πέρα από το speller για το οποίο έχουν υλοποιηθεί και προγενέστερες εφαρμογές του όπως του (Guger, 2009) και (Hoffmann, 2008), έχουμε εφαρμογές με την χρήση P300, για τον χειρισμό ρομπότ (Bell, 2008) και αναπηρικού αμαξιδίου (Iturrate, 2009). Εφαρμογές έχουν επίσης υλοποιηθεί για την χρήση του υπολογιστή, (Martínez-Cagigal, 2017), αλλά και για smart houses (Miralles, 2015). Παρακάτω παρατίθεται ένα σχηματικό διάγραμμα ενός P300 speller, παρατηρείται ότι τα στάδια της εφαρμογής είναι παρόμοια σε όλα τα BCI συστήματα.



Εικόνα 2.7: P300 Speller

2.4.3 MI (Motor Imagery)

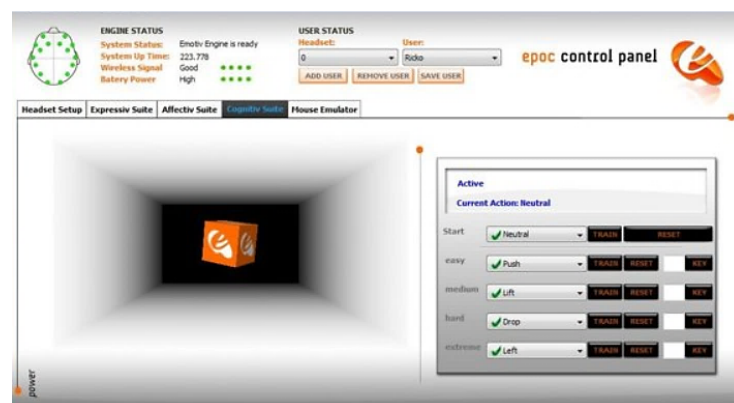
Το Motor Imagery, όπως προδίδει και το όνομά του, είναι η διαδικασία της σκέψης μίας κίνησης από έναν άνθρωπο, χωρίς όμως να την κάνει. Για παράδειγμα, ο χρήστης φαντάζεται ότι κινεί το δεξί του χέρι, χωρίς στην πραγματικότητα να το κινεί. Είναι ευρέως αποδεκτό ότι η νοητική φαντασία κινήσεων, ενεργοποιεί παρόμοιες περιοχές και λειτουργίες του εγκεφάλου, όπως εκείνες που είναι υπεύθυνες για την προετοιμασία και τον προγραμματισμό των αληθινών κινήσεων του σώματος. Οι χρήσεις του Motor Imagery ποικίλουν: χρησιμοποιείται συχνά στην αθλητική εκπαίδευση ως νοητική εξάσκηση, στην νευρολογική αποκατάσταση ασθενών καθώς έχει επίσης χρησιμοποιηθεί στη γνωστική νευροεπιστήμη και τη ψυχολογία για να ερευνηθεί το περιεχόμενο και τη δομή των συγκαλυμμένων διαδικασιών (δηλαδή των ασυνείδητων) που προηγούνται της εκτέλεσης μίας δράσης. Συγκεκριμένα έχει βρεθεί ότι, το MI συνδέεται με τα κύματα βήτα και mu που αναφέρθηκαν παραπάνω. Αναλυτικότερα, τα κύματα αυτά και ιδιαίτερα τα mu, είναι πιο εμφανή όταν ένα άτομο βρίσκεται σε ακινησία, ενώ καταστέλλονται (ERD) όταν ένα άτομο εκτελεί μια κίνηση ή οπτικοποιεί την κίνηση αυτή. Τα πιο αντιπροσωπευτικά φαινόμενα MI-ERD δημιουργούνται με οπτικοποίηση της κίνησης του αριστερού ή δεξιού χεριού και των ποδιών αντίστοιχα, και κατανέμονται στον πρωτεύοντα κινητικό φλοιό. Σε εκείνη την περιοχή βρίσκονται τα ηλεκτρόδια C3, C4 και Cz, από τα οποία λαμβάνονται συνήθως τα σήματα αυτά. Στο σχήμα που ακολουθεί φαίνονται οι περιοχές του εγκεφάλου που διεγείρονται από διάφορες ενεργοποιήσεις μυών.



Εικόνα 2.8: Περιοχές εγκεφάλου που διεγείρονται από διάφορες κινήσεις μυών

Εφαρμογές

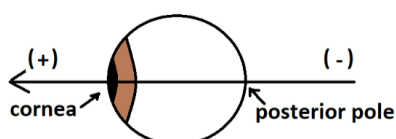
Σε κάθε εφαρμογή που χρησιμοποιεί MI, η σωστή παραγωγή του σήματος EEG είναι ένας σημαντικός παράγοντας προκειμένου να είναι επιτυχημένη. Ως εκ τούτου, τα θέματα που αφορούν την εκπαίδευση του ανθρώπου σε τέτοιες εφαρμογές αξίζουν να εξεταστούν. Συγκριτικά με το SSVEP, το MI χρειάζεται μεγαλύτερο χρόνο εκπαίδευσης ώστε να δημιουργηθούν τα φαινόμενα ERD / ERS. Το τελευταίο μπορεί να πάρει μέρες εκπαίδευσης του χρήστη πριν από την επίτευξη του επιθυμητού επιπέδου απόδοσης. Για να αποκτήσει ο χρήστης τον αυτοέλεγχο της συγκεκριμένης EEG απόκρισης, είναι απαραίτητο κάποιο είδος ανατροφοδότησης, τουλάχιστον στην αρχή, για να επιταχύνει την διαδικασία της εκμάθησης. Κάτι τέτοιο μπορεί να επιτευχθεί με την χρήση γραφικού περιβάλλοντος. Πιο συγκεκριμένα, σε πολλές μελέτες ο χρήστης προσπαθεί με νοητικές εντολές να μετακινήσει ένα αντικείμενο σε μία οθόνη προς κάποιον συγκεκριμένο στόχο, όπως στο άρθρο του (Li M. Z., 2013). Παρακάτω έχουμε ένα παράδειγμα τέτοιου περιβάλλοντος εκπαίδευσης, το οποίο χρησιμοποιείται και στο προηγούμενο άρθρο, όπου ο χρήστης προσπαθεί να κινήσει έναν κύβο με νοητικές εντολές. Συνήθεις εφαρμογές του MI έχουν να κάνουν κυρίως με τον χειρισμό κάποιου ρομπότ, όπως στα άρθρα των (Yu, 2012), (LaFleur, 2013), (Elstob, 2016) και (Li W. J., 2012), και παίξιμο παιχνιδιών στον υπολογιστή, όπως προτείνουν οι (Nijholt, 2009) και (Bonnet, 2013). Σημειώνεται ότι σε πολλές εφαρμογές BCI ο χειρισμός ρομπότ προσομοιώνει τον χειρισμό αναπηρικού αμαξιδίου για άτομα περιορισμένης κινητικότητας, κάτι ιδιαίτερα χρήσιμο.



Εικόνα 2.9: Γραφικό περιβάλλον για την εκπαίδευση BCI συστήματος

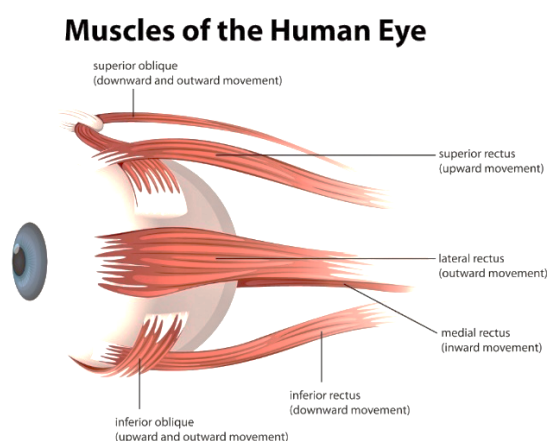
2.5 EOG signals

Ως EOG ορίζουμε την τεχνική μέτρησης του κερατοειδούς – αμφιβληστροειδούς δυναμικού μεταξύ του πρόσθιου (cornea) και οπίσθιου μέρους του ματιού (posterior pole). Η διαφορά τάσης μεταξύ των δύο αυτών περιοχών του ματιού είναι σταθερή, δηλαδή υπάρχει ανεξαρτήτως κίνησης του. Μπορούμε επομένως να θεωρήσουμε το μάτι σαν ένα δίπολο, όπως φαίνεται και στο απλουστευμένο σχήμα. Η μεταβολή στην διαφορά τάσης που μετριέται συσχετίζεται άμεσα με τις κινήσεις του ματιού και το σήμα που προκύπτει από την τεχνική αυτή ονομάζεται EOG σήμα.



Εικόνα 2.10: Γραφική απεικόνιση ματιού

Για να γίνει πιο κατανοητή η τεχνική αυτή παρουσιάζεται συνοπτικά η ανατομία του ματιού, όσον αφορά τους μύες και τα νεύρα τα οποία είναι υπεύθυνα για την κίνησή του. Βλέπουμε ότι οι μύες της κίνησης του ματιού είναι 6 συνολικά: 4 recti και 2 oblique. Οι 4 recti είναι υπεύθυνοι για την κίνηση του ματιού οριζόντια και κάθετα, ενώ οι 2 oblique για την πλάγια κίνησή του. Υπάρχει επίσης ο levator palpebrae superioris, ο οποίος κινεί το επάνω βλέφαρο του ματιού. Τα κρανιακά νεύρα τα οποία συνδέονται με τους μύες αυτούς είναι το κοινό κινητικό (oculomotor No3), το τροχλιακό (trochlear No4) και το απαγωγό νεύρο (abducens No6). Με το μάτι συνδέεται επίσης και το οπτικό νεύρο, το οποίο μεταφέρει τις οπτικές πληροφορίες.



Εικόνα 2.11: Μύες κίνησης ματιού

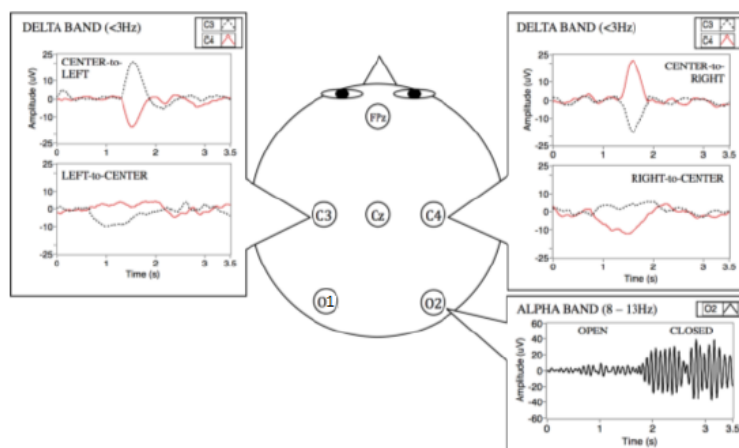
Η αποκωδικοποίηση της κίνησης του ματιού μπορεί να γίνει κατά κύριο λόγο με 3 τρόπους

i) Χρήση εξωτερικής κάμερας

Μία κάμερα καταγράφει τη κίνηση του ματιού, στην συνέχεια μέσω ειδικού λογισμικού το βίντεο επεξεργάζεται και έτσι μπορούμε να συμπεράνουμε την κίνηση του ματιού. Ο τρόπος αυτός δεν χρησιμοποιείται συχνά στις BCI εφαρμογές, λόγω της ποιότητας της κάμερας και της πολυπλοκότητας που ενέχει η επεξεργασία εικόνας.

ii) Μέσω μέτρησης EEG σημάτων

Όπως αναφέρθηκε παραπάνω η κίνηση του ματιού συνδέεται άμεσα και με τις νευρικές απολήξεις του οπτικού και κινητικού νεύρου στο κρανίο. Για αυτό τον λόγο η μέτρησή αυτών των σημάτων μπορεί να γίνει και με την τοποθέτηση ηλεκτροδίων στο κεφάλι σε συγκεκριμένες περιοχές. Τα τελευταία χρόνια, εργασίες που υιοθετούν παρόμοιες ιδέες έχουν δημοσιευτεί, και υποστηρίζουν την επιλογή των ηλεκτροδίων C3, C4, Cz, διότι βρίσκονται κοντά στην περιοχή των νευρών που είναι υπεύθυνα για την κίνηση του ματιού, αλλά και των O1, O2, τα οποία επαφίονται στον ινιακό λοβό εκεί όπου συναντάμε και το οπτικό νεύρο.

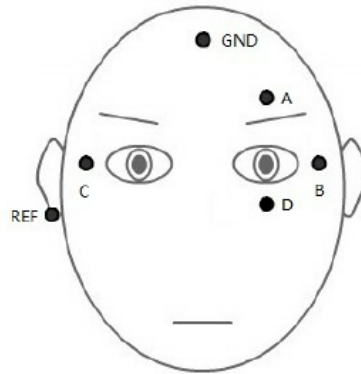


Εικόνα 2.12: Κανάλια μέτρησης EEG σημάτων που σχετίζονται με τα μάτια

iii) Μέσω μέτρησης EOG σημάτων

Όπως αναφέρθηκε η Ηλεκτροφθαλμογραφία είναι η τεχνική καταγραφής EOG σημάτων. Καθώς περιστρέφεται το μάτι, όντας δίπολο, το διάνυσμα του δυναμικού του, περιστρέφεται και αυτό αντίστοιχα. Συνεπώς η τιμή του δυναμικού εξαρτάται άμεσα από το είδος της κίνησης του ματιού. Η καταγραφή της τιμής του δυναμικού επιτυγχάνεται με την χρήση ηλεκτροδίων τα οποία τοποθετούνται συνήθως γύρω από τα μάτια, όπως φαίνεται στην παρακάτω εικόνα. Τα ηλεκτρόδια B - C που βρίσκονται δεξιά και αριστερά από τα μάτια (στους κροτάφους), συσχετίζονται με την οριζόντια κίνηση των ματιών, ενώ

αυτά που βρίσκονται πάνω και κάτω από τα μάτια, δηλαδή τα A-D, με την κατακόρυφη κίνηση. Έχουμε επίσης τα ηλεκτρόδια της γείωσης (GND) και της αναφοράς (Reference).



Εικόνα 2.13: Τοποθέτηση EOG ηλεκτροδίων

Έτσι σύμφωνα με τα παραπάνω, το δυναμικό που καταγράφεται ανάλογα με την οριζόντια κίνηση των ματιών παρουσιάζει την εξής συμπεριφορά:

- *Κίνηση του βλέμματος προς τα δεξιά*

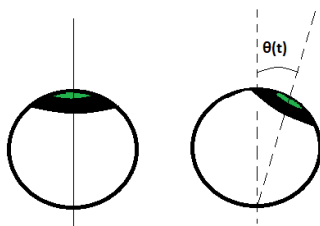
Παρουσιάζεται μία αρνητική μεταβολή της τάσης στο ηλεκτρόδιο που βρίσκεται στην θέση B, και μία αντίστοιχη θετική τιμή στο ηλεκτρόδιο C.

- *Κίνηση του βλέμματος προς τα αριστερά*

Εφόσον η κίνηση είναι αντίθετη της προηγούμενης αυτό έχει ως αποτέλεσμα μία θετική μεταβολή της τάσης στο ηλεκτρόδιο B και μία αρνητική στο ηλεκτρόδιο C αντίστοιχα.

Σε αυτό το σημείο πρέπει να σημειωθεί ότι, το πλάτος της διαφοράς δυναμικού που παράγεται με την οριζόντια κίνηση των ματιών εξαρτάται και από την γωνία στροφής του ματιού.

Συγκεκριμένα, όσο πιο μεγάλη είναι η γωνία στροφής των ματιών τόσο μεγαλύτερο είναι και το πλάτος που μετρείται. Η εξάρτηση αυτή μπορεί να περιγραφεί από την εξίσωση που ακολουθεί.



$$EOG(t) = K1 * \theta(t), \text{ όπου } K1: \text{σταθερά,}$$

$\theta(t)$: γωνία στροφής σε κάθε χρονική στιγμή

Εικόνα 2.14: Γωνία κίνησης ματιών

Εφαρμογές

Τα EOG σήματα αντιμετωπίζονται από πολλές BCI έρευνες σαν artefacts, δηλαδή σαν θόρυβος στα σήματα ενδιαφέροντος, τα οποία οφείλεται να εξαλειφθούν. Συγκεκριμένα, έχουν αναπτυχθεί πολλοί αλγόριθμοι για την ανίχνευση και απομόνωση των δυναμικών που οφείλονται στην κίνηση των ματιών. Όσον αφορά την ανίχνευση παρουσιάζονται μελέτες όπως των (Ramli, 2015), (Rösler, 2013), (Belkacem A. N., 2014) και (Samadi, 2014), με διάφορους

τρόπους κατηγοριοποίησης της κίνησης των ματιών. Παρόλα αυτά έχουν συνταχθεί μελέτες, όπως και η παρούσα διπλωματική, οι οποίες προτείνουν την εκμετάλλευση των δυναμικών που οφείλονται στην κίνηση των ματιών για BCI εφαρμογές. Ενδεικτικές εφαρμογές είναι: ο χειρισμός ρομποτικού βραχίονα όπως έχει προταθεί από τους (Χρήστος, 2017) και (Kasim, 2017), κάποιου ιπτάμενου ρομπότ (Lin, 2015), παιχνίδια (Belkacem A. N.-a., 2015), (Bulling, 2009) και χρήση του υπολογιστή (πχ του ποντικιού, πλοήγηση στο διαδίκτυο). Τέλος, μία ιδιαίτερα ενδιαφέρουσα εφαρμογή των συγκεκριμένων σημάτων είναι η παρακολούθηση του οδηγού για την αποφυγή ατυχημάτων. Πιο συγκεκριμένα, μέσω των EOG σημάτων είναι δυνατή η παρακολούθηση της κατάστασης του οδηγού, δηλαδή εάν ο οδηγός είναι ξύπνιος ή πάει να αποκοιμηθεί, και κατά συνέπεια η κατάλληλη προειδοποίηση του. Σχετικά άρθρα που έχουν δημοσιευθεί πρόσφατα είναι των (Kartsch, 2017) και (Huo, 2016).

2.6 Hybrid BCI system

Όπως έχει γίνει κατανοητό μέχρι στιγμής, μία διεπαφή εγκεφάλου υπολογιστή (BCI) επιτρέπει την επικοινωνία του εγκεφάλου με το περιβάλλον μέσω ενός υπολογιστή. Αυτό έχει αξιοποιηθεί με την υλοποίηση διαφόρων εφαρμογών όπως προαναφέρθηκε, παραδείγματος χάριν για τον έλεγχο ενός πρόσθετου άκρου. Πρόσφατα, έχουν προταθεί νέες προσεγγίσεις για την αντιμετώπιση αυτών των ζητημάτων στις τρέχουσες μελέτες BCI συνδυάζοντας ένα σύστημα BCI με άλλο σύστημα ή συστήματα που χρησιμοποιούν νευρολογικά σήματα, φυσιολογικά σήματα ή εξωτερικά σήματα. Αυτή η νέα τεχνική ονομάζεται υβριδικό BCI ή hBCI. Δεδομένου ότι κάθε σύστημα στο υβριδικό BCI είναι δυνατό να έχει διαφορετικό σήμα, ή σήματα με διαφορετικά χαρακτηριστικά, τα υβριδικά BCIs αποτελούνται από διάφορα σήματα εισόδου. Για παράδειγμα, για ένα τέτοιο σύστημα μπορεί να έχουμε ως είσοδο και εγκεφαλικά σήματα (EEG) και EOG σήματα. Έτσι με τον συνδυασμό πολλών διαφορετικών ειδών σημάτων μπορούν να αντιμετωπιστούν περιορισμοί οι οποίοι υπάρχουν στα single BCI συστήματα και να βελτιωθεί αισθητά η ακρίβεια του συστήματος. Παρόλα αυτά διαφαίνεται ότι τα υβριδικά συστήματα είναι πολύ πιο πολύπλοκα στον σχεδιασμό και στην υλοποίησή τους. Ένα πολύ ενδιαφέρον και συγκεντρωτικό άρθρο για τα υβριδικά συστήματα είναι των (Choi I. R., 2017). Όσον αφορά τον συνδυασμό τεχνικών για eye tracking, όπως ο συνδυασμός EEG και EOG σημάτων που προτείνεται και στην παρούσα διπλωματική, ενδεικτικά παραδείγματα είναι οι εφαρμογές που έχουν προταθεί από τους (Kim, 2014), (López-Gil, 2016) και (Gao, 2015).

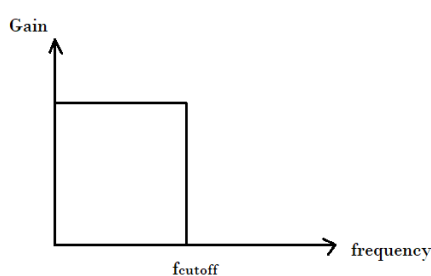
3 Κεφάλαιο: Θεωρητικό υπόβαθρο

Όπως αναφέρθηκε τα EEG σήματα χαρακτηρίζονται από χαμηλό SNR, πράγμα που σημαίνει ότι είναι ιδιαίτερα θορυβώδη. Γι' αυτό τον λόγο είναι απαραίτητη η αποθορυβοποίησή τους. Αυτό επιτυγχάνεται συνήθως με την εφαρμογή κατάλληλων φίλτρων. Στην συνέχεια θα αναφερθούμε σε 3 είδη φίλτρων, τα οποία έχουν χρησιμοποιηθεί σε αυτή την διπλωματική αλλά και γενικότερα συναντώνται στο μεγαλύτερο εύρος της βιβλιογραφίας.

3.1 Προεπεξεργασία – Αποθορυβοποίηση (Preprocessing)

3.1.1 Βαθυπερατό φίλτρο (Low pass filter)

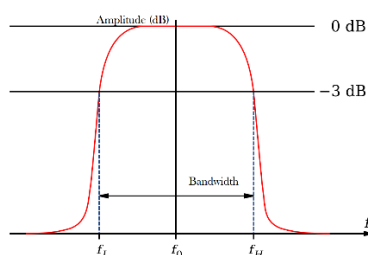
Γενικά αυτό το φίλτρο επιτρέπει να περάσουν οι χαμηλές συχνότητες του σήματος στο οποίο εφαρμόζεται, ενώ απορρίπτει τις υψηλές. Χαρακτηρίζεται από την συχνότητα αποκοπής (cutoff frequency), η οποία καθορίζει ποιες συχνότητες θα περάσουν. Ένα απλουστευμένο σχήμα ενός ιδανικού βαθυπερατού φίλτρου φαίνεται παρακάτω.



Εικόνα 3.1: Απόκριση βαθυπερατού φίλτρου

3.1.2 Ζωνοπερατό φίλτρο (Bandpass filter)

Ένα bandpass filter αφήνει να περάσουν συχνότητες ενός συγκεκριμένου εύρους, ενώ απορρίπτει τις συχνότητες που βρίσκονται απέξω από αυτό. Το εύρος καθορίζεται από 2 συχνότητες, την f_{low} και την f_{high} . Ακολουθεί ένα σχήμα ενός ζωνοπερατού φίλτρου.



Εικόνα 3.2: Απόκριση ζωνοπερατού φίλτρου στο πεδίο της συχνότητας

3.1.3 Φίλτρο διαμέσου τιμής (Median filter)

Το median filter είναι ένα μη γραμμικό φίλτρο που χρησιμοποιείται για την αφαίρεση του τυχαίου θορύβου. Η λειτουργία του είναι η εξής: Διαπερνά το σήμα τιμή προς τιμή και αντικαθιστά κάθε τιμή του σήματος από μία νέα, η οποία προκύπτει κάθε φορά ως ο διάμεσος των γειτονικών τιμών της τιμής εισόδου στο φίλτρο. Το pattern των γειτονικών τιμών ονομάζεται παράθυρο (window) και διαπερνά όπως αναφέρθηκε όλο το σήμα. Σημειώνεται ότι εάν το παράθυρο έχει περιττό αριθμό τιμών, τότε ο διάμεσος (median) είναι εύκολο να προσδιοριστεί, και είναι η μεσαία τιμή, αφού όλες οι τιμές στο παράθυρο ταξινομηθούν αριθμητικά.

3.1.4 Wiener filter

Θεωρούμε ένα σήμα στο οποίο έχει προστεθεί θόρυβος με γνωστή φασματική ισχύ. Εάν εφαρμόσουμε το φίλτρο Wiener σε αυτό το σήμα, τότε θα πάρουμε μία προσέγγιση του αρχικού μη θορυβώδους σήματος. Ο προσδιορισμός αυτός γίνεται μέσω ελαχιστοποίησης του τετραγωνικού σφάλματος, όπου το σφάλμα είναι η διαφορά του μη θορυβώδους από το θορυβώδες σήμα.

3.2 Εξαγωγή χαρακτηριστικών (feature extraction)

Ένα ακόμη σημαντικό στάδιο είναι η εξαγωγή των χαρακτηριστικών του σήματος. Συχνά κάτι τέτοιο επιτυγχάνεται πιο εύκολα όταν το σήμα μετασχηματίζεται σε ένα διαφορετικό πεδίο από εκείνο του χρόνου. Οι βασικοί μετασχηματισμοί που χρησιμοποιούνται για αυτό το σκοπό είναι συνήθως ο Discrete Fourier Transform (DFT), ο Hilbert Transform (HT) και ο Wavelet Transform (WT). Στην συγκεκριμένη διπλωματική χρησιμοποιήθηκε ο DFT, ο οποίος αναλύεται παρακάτω.

3.2.1 Δειγματοληψία

Οι τεχνικές της δειγματοληψίας μετατρέπουν μία συνεχή συνάρτηση σε μία ακολουθία διακριτών τιμών, προκειμένου να είναι δυνατή η επεξεργασία τους από ένα υπολογιστή. Θεωρούμε μία συνεχή συνάρτηση $f(t)$ η οποία εκτείνεται από το $-\infty$ έως το $+\infty$ και θέλουμε να εφαρμόσουμε την διαδικασία της δειγματοληψίας σε ομοιόμορφα διαστήματα ΔT . Ένας τρόπος να μοντελοποιήσουμε την διαδικασία στο πεδίο του χρόνου, είναι να πολλαπλασιάσουμε την $f(t)$ με την συνάρτηση δειγματοληψίας, η οποία περιγράφεται από ένα άθροισμα κρουστικών αποκρίσεων, όπως φαίνεται στην εξίσωση που ακολουθεί:

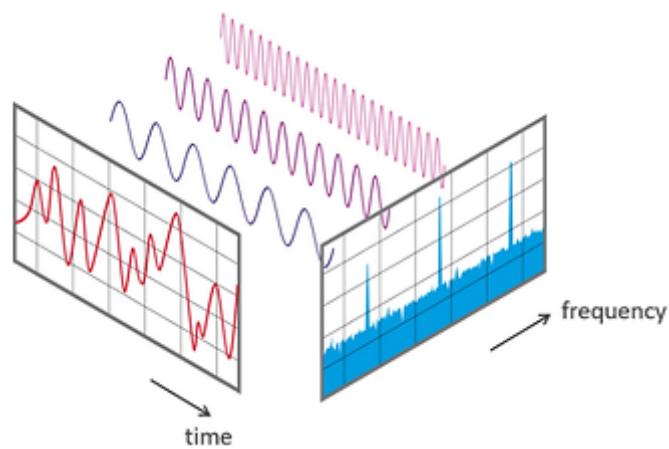
$$\tilde{f}(t) = \sum_{n=-\infty}^{+\infty} f(t)\delta(t - n\Delta T)$$

Με αυτόν τον τρόπο παίρνουμε ομοιόμορφα δείγματα του αρχικού σήματος με συχνότητα δειγματοληψίας $f_s = 1/\Delta T$. Εκτός από το πεδίο του χρόνου, μπορούμε να εφαρμόσουμε την δειγματοληψία και στο πεδίο της συχνότητας. Για αυτό τον σκοπό παίρνουμε τη συνέλιξη των μετασχηματισμών Fourier του σήματος ($X_c(j\Omega)$) και της συνάρτησης δειγματοληψίας ($S(j\Omega)$), πολλαπλασιάζοντας με $1/2\pi$. Έτσι έχουμε το δειγματοληπτούμενο σήμα:

$$X_s(j\Omega) = \frac{1}{2\pi} X_c(j\Omega) * S(j\Omega) = \frac{1}{T} \sum_{k=-\infty}^{+\infty} X_c(j(\Omega - k\Omega_s))$$

3.2.2 DFT (Discrete Fourier Transform)

Η βασική ιδέα του μετασχηματισμού Fourier είναι ότι τα σήματα μπορούν να εκφραστούν σαν άθροισμα πολλών ημιτονοειδών συναρτήσεων. Κάτι τέτοιο μας επιτρέπει να μπορέσουμε να εξάγουμε συμπεράσματα για τα χαρακτηριστικά του σήματος στο πεδίο συχνοτήτων. Στην Εικόνα 3.3 βλέπουμε μία απλή απεικόνιση ενός συνεχούς σήματος που αποτελείται από 3 ημίτονα διαφορετικών συχνοτήτων.



Εικόνα 3.3: Σχηματική περιγραφή FT

Όσον αφορά τον DFT είναι ένας μετασχηματισμός ο οποίος εφαρμόζεται σε σήματα διακριτού χρόνου, δηλαδή σήματα με διακριτές τιμές, τα οποία προκύπτουν από την διαδικασία της δειγματοληψίας. Για ένα σήμα $x[n]$, με $n = 0, 1, \dots, N - 1$, όπου N ο αριθμός των δειγμάτων του, ο DFT ορίζεται ως:

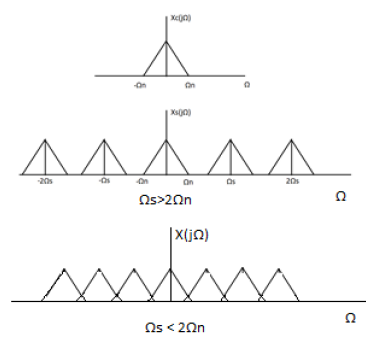
$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-\frac{i2\pi kn}{N}}, \quad \text{όπου } \omega_k = \frac{2\pi}{N} k \text{ με } k = 0, 1, \dots, N$$

Παρατηρούμε όμως ότι για τον υπολογισμό του DFT πρέπει να υπολογίσουμε πολλά γινόμενα κάτι ιδιαίτερα χρονοβόρα για μεγάλο N, συγκεκριμένα η πολυπλοκότητα είναι $O(N^2)$. Για αυτό τον λόγο χρησιμοποιούμε τον αλγόριθμο Fast Fourier Transform (FFT) για τον υπολογισμό του DFT, ο οποίος έχει πολυπλοκότητα $O(n \log n)$.

Ένα σημείο στο οποίο όμως πρέπει να σταθούμε, είναι το **θεώρημα Nyquist-Shannon**. Σύμφωνα με αυτό, προκειμένου να μπορεί ένα σήμα να ανακατασκευαστεί από τον DFT του πάλι στο πεδίο του χρόνου, πρέπει η συχνότητα δειγματοληψίας του σήματος να είναι τουλάχιστον διπλάσια του εύρους ζώνης του αρχικού σήματος, δηλαδή :

$$\Omega_s = \frac{2\pi}{T_s} \geq 2\Omega_N, \text{ όπου } \Omega_s: \text{συχνότητα δειγματοληψίας,}$$

$$\Omega_N: \text{εύρος ζώνης αρχικού σήματος}$$



Εικόνα 3.4: Διαδικασία δειγματοληψίας α) αρχικό σήμα β) ικανοποίηση της συνθήκης γ) aliasing

Εάν η παραπάνω συνθήκη δεν ικανοποιείται τότε έχουμε ένα φαινόμενο το οποίο ονομάζεται **aliasing** και προκαλεί αλλοίωση του σήματος και λόγω αυτού δεν είναι δυνατή η ανακατασκευή του σήματος στο πεδίο του χρόνου.

3.2.3 Power Spectral Density (PSD)

Το PSD περιγράφει πως η ισχύς ενός σήματος είναι μοιρασμένη στο πεδίο των συχνοτήτων. Μπορεί να βρεθεί παίρνοντας τον μετασχηματισμό Fourier της συνάρτησης αυτοσυσχέτισης του σήματος και η μονάδα μέτρησης του είναι η $\frac{\text{ενέργεια (Watts)}}{\text{συχνότητα (Hz)}}$. Ο τύπος αναλυτικά είναι:

$$S(\omega) = \sum_{k=-\infty}^{+\infty} r[k]e^{-jk\omega}, \text{ όπου } r[k] \text{ η συνάρτηση αυτοσυσχέτισης του σήματος}$$

Πιο απλά το PSD μπορεί να υπολογιστεί επίσης και μέσω του $\frac{|X(j\omega)|^2}{L}$, όπου ο παρονομαστής είναι ο DFT του σήματος και L το μέγεθός του.

3.3 Μηχανική Μάθηση

3.3.1 Επιβλεπόμενη μάθηση (Supervised learning)

Υπό μία ευρεία έννοια μπορούμε να κατηγοριοποιήσουμε τις διαδικασίες μάθησης μέσω των οποίων λειτουργεί ένα σύστημα ως εξής: σε επιβλεπόμενη και μη επιβλεπόμενη μάθηση. Στην επιβλεπόμενη μάθηση τα παραδείγματα που παρουσιάζονται στο σύστημα έτσι ώστε να εκπαιδευτεί είναι ζεύγη {είσοδος, έξοδος}, όπου η είσοδος είναι συνήθως ένα διάνυσμα χαρακτηριστικών και η έξοδος η επιθυμητή απόκριση. Τα παραδείγματα εκπαίδευσης αυτής της μορφής ονομάζονται για αυτό το λόγο και *labeled*. Στο πρόβλημα της κατηγοριοποίησης (*classification*) η επιθυμητή απόκριση είναι μία συγκεκριμένη κλάση. Χρησιμοποιώντας την επιβλεπόμενη μάθηση οι παράμετροι του συστήματος προσαρμόζονται κατάλληλα με βάση τα παραδείγματα εκπαίδευσης, έτσι ώστε το σύστημα να μπορεί να ταξινομήσει σωστά νέα άγνωστα παραδείγματα που του παρουσιάζονται. Το τελευταίο ονομάζεται γενίκευση.

3.3.2 Αλγόριθμος KNN (K- Nearest neighbors)

Ο KNN είναι ένας από τους πιο απλούς αλγόριθμους της μηχανικής μάθησης. Είναι *instance based*, δηλαδή κάθε φορά για να ταξινομήσει ένα νέο δείγμα το συγκρίνει με τα δείγματα τα οποία έχουμε χρησιμοποιήσει στην φάση της εκπαίδευσης. Η βασική ιδέα του αλγορίθμου είναι: κάθε νέο δείγμα ταξινομείται με βάση τις κλάσεις των K – γειτόνων του, δηλαδή αυτών που βρίσκονται πιο κοντά σε αυτό. Με τον όρο *κοντά*, εννοούμε συνήθως την Ευκλείδεια απόσταση του δείγματος προς εξέταση και όλων των υπάρχοντων δειγμάτων. Υπολογίζοντας όλες αυτές τις αποστάσεις, επιλέγουμε τα K – δείγματα (γείτονες) με την μικρότερη τιμή. Η επιλογή του K εξαρτάται από το δείγμα δεδομένων που έχουμε, για παράδειγμα για θορυβώδη δεδομένα συστήνεται ο μεγάλος αριθμός K - γειτόνων, κάτι το οποίο όμως μειώνει την απόσταση μεταξύ των κλάσεων.

3.3.3 Τεχνητά Νευρωνικά Δίκτυα (Neural Networks)

3.3.3.1 Ο μονοστρωματικός νευρώνας (*perceptron*)

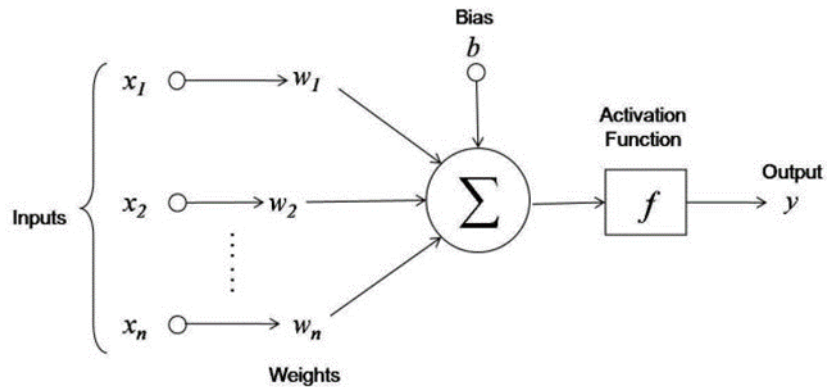
Ο νευρώνας είναι μία μονάδα επεξεργασίας πληροφορίας, η οποία είναι θεμελιώδης για την λειτουργία του Νευρωνικού Δικτύου. Το *perceptron* βασίζεται σε ένα μοντέλο ενός μη γραμμικού νευρώνα, και αποτελεί την απλούστερη δυνατή μορφή ενός Νευρωνικού Δικτύου για την ταξινόμηση προτύπων. Στην Εικόνα 3.5, απεικονίζεται η αρχιτεκτονική ενός μονοστρωματικού νευρώνα (*perceptron*). Ο τύπος που δίνει την τιμή της εξόδου (*output*) του αθροιστή είναι ο ακόλουθος:

$$z = \sum_{i=1}^N w_i X_i + b \quad (1)$$

Όπου: \mathbf{X} είναι το διάνυσμα N εισόδων

w_i είναι τα βάρη που ο νευρώνας αποδίδει σε κάθε είσοδο

b είναι ο σταθερός όρος (bias)



Εικόνα 3.5: Μονοστρωματικός νευρώνας

Όπως είναι φανερό, η εξίσωση (1) είναι γραμμική ως προς τις εισόδους X_i . Για να μπορεί ένα Νευρωνικό Δίκτυο να χειριστεί και μη-γραμμικές συναρτήσεις, η έξοδος κάθε νευρώνα τροφοδοτείται σε μία συνάρτηση ενεργοποίησης (Activation Function). Αν f είναι μία συνάρτηση ενεργοποίησης, η εξίσωση (1) γίνεται:

$$f\left(\sum_{i=1}^N w_i X_i + b\right)$$

Οι πιο χρησιμοποιούμενες συναρτήσεις ενεργοποίησης είναι :

- Γραμμική

Ουσιαστικά δεν χρησιμοποιείται συνάρτηση ενεργοποίησης και η έξοδος παραμένει γραμμική ως προς τις εισόδους, δηλαδή:

$$f(z) = z$$

- Σιγμοειδής (Sigmoid) ή Λογιστική (Logistic)

Χρησιμοποιείται κυρίως όταν η έξοδος θέλουμε να παίρνει τιμές $[0, 1]$.

$$y = \frac{1}{1 + e^{-z}}$$

- Υπερβολική εφαπτομένη (Tanh)

Η συγκεκριμένη συνάρτηση χρησιμοποιείται όταν η έξοδος θέλουμε να παίρνει τιμές [-1, 1].

$$y = \tanh (z)$$

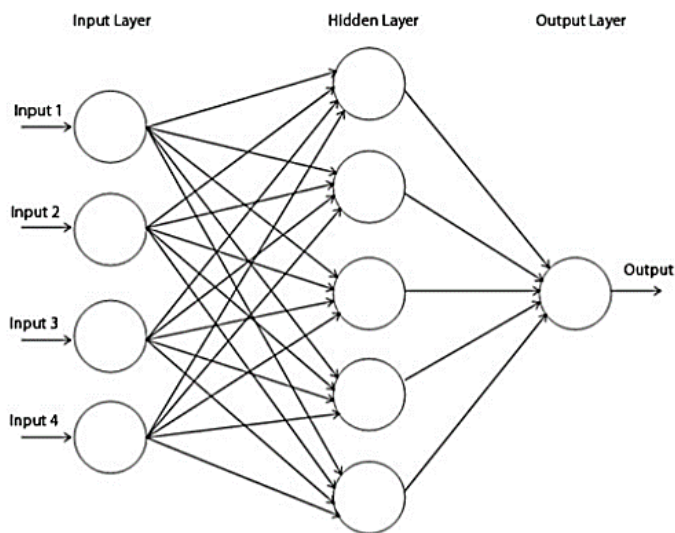
- Rectified Linear unit (ReLu)

Η ReLu είναι μία συνάρτηση ενεργοποίησης που ορίζεται ως το θετικό μέρος του ορίσμάτος της:

$$y = \max (0, z)$$

3.3.3.2 Multi - Layer Perceptron (MLP)

Το βασικό δομικό στοιχείο ενός Νευρωνικού Δικτύου είναι ο μονοστρωματικός νευρώνας (perceptron). Όπως θα δούμε και στο παρακάτω σχήμα το μοντέλο των τεχνητών νευρωνικών δικτύων αποτελείται από διαδοχικά στρώματα νευρώνων (perceptrons), καθένα συνδεδεμένο με τους νευρώνες του προηγούμενου στρώματος. Αυτό το είδος νευρωνικού ονομάζεται Multi - Layer perceptron (MLP) δίκτυο. Η αρχιτεκτονική του φαίνεται στο παρακάτω σχήμα.



Εικόνα 3.6:Multi Layer Perceptron

Οι παράμετροι που πρέπει να προσδιοριστούν είναι τα προαναφερθέντα βάρη και η μάθησή τους γίνεται με τον αλγόριθμο Back - propagation.

3.3.4 SVM (Support Vector Machines)

Τα SVM (μηχανές διανυσμάτων υποστήριξης), είναι δυαδικές μηχανές με δυνατότητα μάθησης. Η κεντρική ιδέα πίσω από μία τέτοια μηχανή είναι συνοπτικά η εξής:

Δοθέντος ενός δείγματος εκπαίδευσης η μηχανή διανυσμάτων υποστήριξης κατασκευάζει ένα υπερεπίπεδο ως επιφάνεια απόφασης, με τέτοιο τρόπο ώστε το περιθώριο διαχωρισμού μεταξύ των θετικών και αρνητικών παραδειγμάτων να μεγιστοποιείται.

Το SVM χρησιμοποιείται για την επίλυση δύο τύπων προβλημάτων: της μη γραμμικής παλινδρόμησης και της ταξινόμησης μοτίβων. Με το τελευταίο πρόβλημα ασχολείται και η παρούσα διπλωματική. Στην συνέχεια, θα παρουσιάσουμε εν συντομία την βασική θεωρία πίσω από την λειτουργία των μηχανών υποστήριξης διανυσμάτων.

Έστω ότι έχουμε ένα δείγμα εκπαίδευσης $\{(x_i, d_i)\}_{i=1}^N$, όπου x_i είναι το δείγμα εκπαίδευσης του i -οστού παραδείγματος και d_i είναι η αντίστοιχη επιθυμητή απόκριση (έξοδος). Έχουμε επίσης δύο κλάσεις (πρότυπα), η πρώτη αναπαρίσταται από το υποσύνολο $d_i = +1$ και η δεύτερη από το $d_i = -1$. Στην συνέχεια αναλόγως με το αν τα πρότυπα των δύο κλάσεων είναι γραμμικά διαχωρίσιμα ή όχι έχουμε 2 περιπτώσεις για την μορφή του υπερεπιπέδου που εκτελεί τον διαχωρισμό.

3.3.4.1 Βέλτιστο υπερεπίπεδο για γραμμικά διαχωρίσιμα πρότυπα

Η εξίσωση της επιφάνειας της απόφασης που εκτελεί τον διαχωρισμό είναι η ακόλουθη:

$$w^T x + b = 0$$

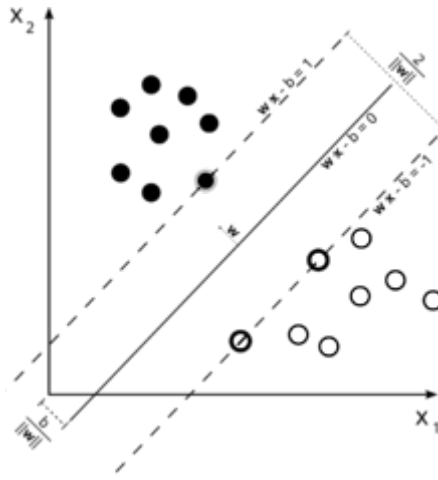
όπου: x το διάνυσμα εισόδου, w ένα προσαρμόσιμο διάνυσμα βαρών και b μία πόλωση.

Άρα μπορούμε να γράψουμε ότι

$$\begin{cases} w^T x + b \geq 0, & \text{για } d_i = +1 \\ w^T x + b < 0, & \text{για } d_i = -1 \end{cases}$$

Αν έχουμε τώρα συγκεκριμένο διάνυσμα βαρών w και συγκεκριμένη πόλωση b , ορίζουμε ως **περιθώριο διαχωρισμού ρ (margin of separation)** την απόσταση μεταξύ του υπερεπιπέδου διαχωρισμού (που ορίζεται από αυτές τις συγκεκριμένες μεταβλητές) και του πλησιέστερου σημείου δεδομένων. Επομένως, ο στόχος του SVM είναι να βρει το συγκεκριμένο υπερεπίπεδο το οποίο μεγιστοποιεί το περιθώριο διαχωρισμού. Υπό αυτή την συνθήκη η επιφάνεια διαχωρισμού αναφέρεται ως βέλτιστο υπερεπίπεδο και ορίζεται, για τις βέλτιστες τιμές w_o , b_o , ως:

$$\begin{cases} w_o^T x + b_o \geq 1, & \text{για } d_i = +1 \\ w_o^T x + b_o \leq -1, & \text{για } d_i = -1 \end{cases}$$



Εικόνα 3.7: Σχηματική αναπαράσταση του σκεπτικού ενός βέλτιστου υπερεπιπέδου

Σημειώνεται ότι τα συγκεκριμένα σημεία δεδομένων $\{(x_i, d_i)\}$ για τα οποία ισχύει η ισότητα της εξίσωσης ονομάζονται **διανύσματα υποστήριξης (support vectors)**, στο οποίο οφείλεται και η ονομασία της μηχανής αυτής. Τα διανύσματα αυτά είναι ιδιαίτερα σημαντικά, διότι είναι εκείνα τα σημεία τα οποία βρίσκονται πλησιέστερα στο βέλτιστο υπερεπίπεδο, επομένως είναι τα πλέον δύσκολα για ταξινόμηση και για αυτό τον λόγο έχουν άμεση επίδραση στην βέλτιστη θέση της επιφάνειας απόφασης. Με βάση αυτά που ειπώθηκαν για τα διανύσματα υποστήριξης, αποδεικνύεται ότι η βέλτιστη τιμή του περιθωρίου διαχωρισμού ρ είναι:

$$\rho = \frac{2}{\|w_0\|}$$

όπου: $\|w_0\|$ είναι η Ευκλείδεια νόρμα του διανύσματος βαρών του βέλτιστου υπερεπιπέδου. Από την εξίσωση αυτή καταλαβαίνουμε ότι :

$$\text{μεγιστοποίηση } \rho \Leftrightarrow \text{ελαχιστοποίηση } \|w_0\|$$

Στην Εικόνα 3.7 βλέπουμε τα διανύσματα υποστήριξης τα οποία βρίσκονται πάνω στις διακεκομμένες γραμμές καθώς επίσης και το βέλτιστο περιθώριο διαχωρισμού των 2 κλάσεων που αναφέρθηκε προηγουμένως.

3.3.4.2 Βέλτιστο υπερεπίπεδο για μη γραμμικά διαχωρίσιμα πρότυπα

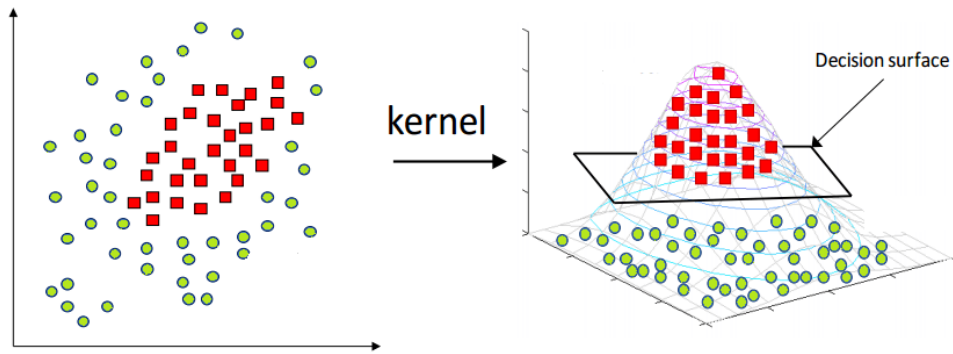
Δοθέντος ενός δείγματος δεδομένων εκπαίδευσης μη γραμμικά διαχωρίσιμων προτύπων, δεν είναι δυνατό να κατασκευάσουμε ένα υπερεπίπεδο διαχωρισμού χωρίς να έχουμε σφάλματα ταξινόμησης. Πάρα αυτά μπορούμε να βρούμε ένα βέλτιστο υπερεπίπεδο, το οποίο θα ελαχιστοποιεί την πιθανότητα σφάλματος ταξινόμησης, υπολογισμένο επί του συνόλου του δείγματος εκπαίδευσης. Γι' αυτό τον λόγο εισάγουμε ένα νέο σύνολο μη αρνητικών βαθμωτών μεταβλητών $\{\xi_i\}_{i=1}^N$, οι οποίες ονομάζονται **μεταβλητές χαλάρωσης (slack)**. Έτσι, η εξίσωση του βέλτιστου υπερεπιπέδου παίρνει την μορφή:

$$d_i(w^T x_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, N$$

Καταλαβαίνουμε ότι οι μεταβλητές χαλάρωσης μετρούν ουσιαστικά την απόκλιση ενός σημείου δεδομένων από την ιδανική συνθήκη διαχωρισιμότητας προτύπων. Έτσι συνοπτικά, η εύρεση των βέλτιστων τιμών του διανύσματος βαρών και της πόλωσης γίνεται παρόμοια με πριν, με κάποιες μικρές τροποποιήσεις λόγω των μεταβλητών χαλάρωσης, ενώ τα διανύσματα υποστήριξης ορίζονται ακριβώς με τον ίδιο τρόπο όπως και προηγουμένως.

3.3.5 SVM ως μηχανή πυρήνα

Έχουμε αναφέρει ως τώρα ότι υπολογίζοντας τις βέλτιστες τιμές του διανύσματος βαρών και πόλωσης μπορούμε να κατασκευάσουμε το βέλτιστο υπερεπίπεδο διαχωρισμού. Μία άλλη πολύ σημαντική μέθοδος εύρεσης του βέλτιστου επιπέδου διαχωρισμού είναι το τέχνασμα του πυρήνα (kernel method), η οποία είναι συνήθως υπολογιστικά φθηνότερη. Το τέχνασμα του πυρήνα οφείλει το όνομά τους στη χρήση των συναρτήσεων του πυρήνα (kernel functions), οι οποίες μετασχηματίζουν τον χώρο εισόδου των δεδομένων, σε ένα χώρο υψηλών διαστάσεων (feature space), χωρίς όμως να υπολογίζουν τις συντεταγμένες των δεδομένων σε αυτό τον χώρο, αλλά υπολογίζοντας τα εσωτερικά γινόμενα μεταξύ των εικόνων όλων των ζευγών δεδομένων στο χώρο των χαρακτηριστικών. Σημειώνεται ότι στον χώρο υψηλότερης διαστατικότητας, τα μετασχηματισμένα δεδομένα εισόδου, είναι γραμμικά διαχωρίσιμα. Άρα με το τέχνασμα του πυρήνα, επαρκεί να καθορίσουμε τον πυρήνα για την ταξινόμηση προτύπων στην έξοδο. Για να γίνει πιο εύκολα κατανοητή η μέθοδος παραθέτουμε το παρακάτω παραστατικό σχήμα.



Εικόνα 3.8: Μετασχηματισμός δεδομένων

Ο πυρήνας συμβολίζεται με $k(x, x_i)$ και είναι μία συνάρτηση η οποία υπολογίζει το εσωτερικό γινόμενο των εικόνων που παράγονται στο χώρο χαρακτηριστικών 2 σημείων δεδομένων του χώρου εισόδου.

Οι σημαντικότερες συναρτήσεις πυρήνα (kernel functions), οι οποίες δοκιμάστηκαν και σε αυτή την διπλωματική, αναφέρονται στον πίνακα που ακολουθεί:

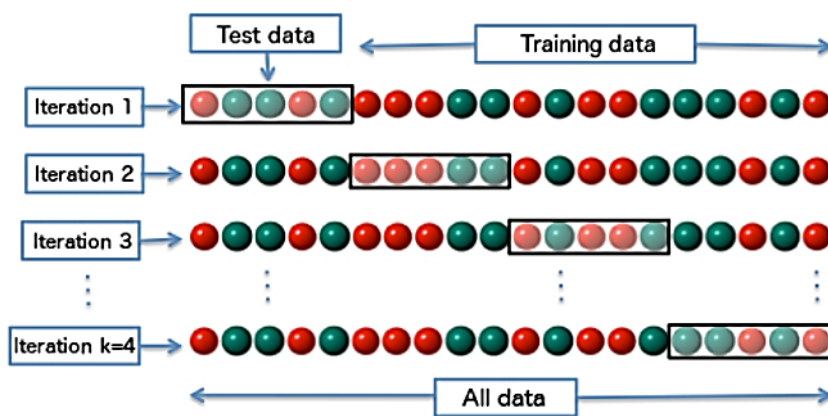
Γραμμική (Linear)	$k(x, x_i) = x^T x_i$
Πολυωνυμική (Polynomial)	$k(x, x_i) = (x^T x_i + 1)^p$, p : δύναμη που καθορίζεται από τον χρήστη
Δίκτυο RBF σ^2 : εύρος που καθορίζεται από τον χρήστη	$k(x, x_i) = \exp\left(-\frac{1}{2\sigma^2} \ x - x_i\ ^2\right)$,
Σιγμοειδής (Sigmoid)	$k(x, x_i) = \tanh(\beta_0 x^T x_i + \beta_1)$, για συγκεκριμένες τιμές των β_0, β_1

Εικόνα 3.9: Πυρήνες διανυσμάτων υποστήριξης

3.4 Επαλήθευση με την μέθοδο Cross – Validation

Είναι γνωστό ότι η διαδικασία μάθησης ενός συστήματος έχει ως αποτέλεσμα μία επιλογή των παραμέτρων του για ένα ορισμένο σύνολο δεδομένων. Με αυτόν τον τρόπο ελπίζουμε το σύστημα να καταλήξει καλά εκπαιδευμένο, ώστε μαθαίνοντας καλά από το παρελθόν, να μπορεί να γενικεύει στο μέλλον. Ένα καθιερωμένο εργαλείο της στατιστικής που χρησιμοποιείται για την εκπαίδευση και επαλήθευση, είναι η *διασταυρωμένη επικύρωση (cross validation)*. Σύμφωνα με αυτό, το διαθέσιμο σύνολο δεδομένων χωρίζεται τυχαία σε ένα σύνολο *δεδομένων εκπαίδευσης (training data)* και ένα σύνολο *δεδομένων ελέγχου (testing data)*. Το δείγμα εκπαίδευσης διαμερίζεται περαιτέρω σε δύο ξένα υποσύνολα: i) υποσύνολο εκτίμησης, το οποίο χρησιμοποιείται για την επιλογή του μοντέλου ii) υποσύνολο επικύρωσης, το οποίο χρησιμοποιείται για τον έλεγχο του μοντέλου. Με αυτό τον τρόπο μπορούμε χρησιμοποιώντας το δείγμα εκπαίδευσης να αξιολογήσουμε διάφορα μοντέλα και να επιλέξουμε το βέλτιστο. Έχοντας επιλέξει το μοντέλο με την καλύτερη απόδοση, για να αποφύγουμε την πιθανότητα υπερπροσαρμογής σε συγκεκριμένα δεδομένα, μετράμε την δυνατότητα γενίκευσης του μοντέλου με βάση το σύνολο ελέγχου. Υπάρχουν και άλλες παραλλαγές της μεθόδου του cross validation, μία από αυτές που χρησιμοποιείται και στην παρούσα διπλωματική είναι η **πολλαπλή**

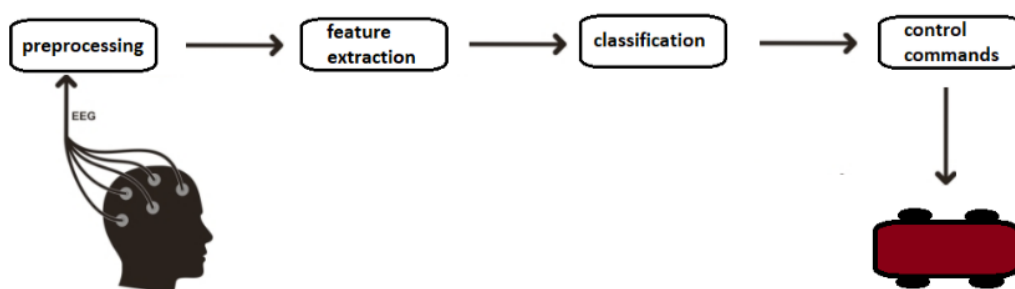
διασταυρωμένη επικύρωση (K- fold cross validation). Στην περίπτωση αυτή το σύνολο δεδομένων εκπαίδευσης N , χωρίζεται σε K - υποσύνολα, όπου $K > 1$, υποθέτοντας ότι το σύνολο N μπορεί να διαιρεθεί με το K . Έπειτα το μοντέλο εκπαιδεύεται με όλα τα υποσύνολα εκτός από ένα και το σφάλμα επικύρωσης μετριέται ελέγχοντας το ως προς το υποσύνολο με το οποίο δεν έχει εκπαιδευτεί. Αυτή η διαδικασία επαναλαμβάνεται K φορές, χρησιμοποιώντας κάθε φορά διαφορετικό υποσύνολο για την επικύρωση. Έτσι, η απόδοση του μοντέλου υπολογίζεται ως ο μέσος όρος του τετραγωνικού σφάλματος της επικύρωσης επί του συνόλου των δοκιμών. Στο παρακάτω σχήμα βλέπουμε μία αναπαράσταση της πολλαπλής διασταυρωμένης επικύρωσης για $K=4$.



Εικόνα 3.10: Μέθοδος cross validation

4 Κεφάλαιο: Περιγραφή συστήματος

Στην ενότητα αυτή περιγράφονται τα στάδια του συστήματος που υλοποιήσαμε, προκειμένου να αξιοποιήσουμε τα EEG και EOG σήματα και να υλοποιήσουμε την επιθυμητή εφαρμογή. Συνοπτικά το σύστημά μας χωρίζεται στα εξής στάδια: καταγραφή σημάτων (recording), προεπεξεργασία (preprocessing), εξαγωγή χαρακτηριστικών (feature extraction) και ταξινόμηση (classification). Όπως έχει αναφερθεί παραπάνω, σκοπός μας είναι η ταξινόμηση 4 κινήσεων των ματιών: άνοιγμα/κλείσιμο και κίνηση δεξιά/αριστερά. Για να επιτευχθεί όμως αυτό θα πρέπει να έχουμε συλλέξει αρκετά δεδομένα για τις 4 κινήσεις, έτσι ώστε το σύστημά μας να είναι κατάλληλα εκπαιδευμένο, προκειμένου στην συνέχεια να μπορεί να ταξινομεί καινούργια άγνωστα δεδομένα εισόδου. Για τον σκοπό αυτό έχουμε δημιουργήσει ένα γραφικό περιβάλλον διεπαφής για την ομοιόμορφη καταγραφή δειγμάτων από τον χρήστη, το οποίο περιγράφεται στην συνέχεια. Καταλήγοντας, έχοντας εκπαιδεύσει κατάλληλα τα σύστημά μας, τελευταίο στάδιο της εφαρμογής είναι η αντιστοίχιση των 4 κινήσεων των ματιών με 4 βασικές κινήσεις ενός demo car που κατασκευάστηκε.



Εικόνα 4.1: Στάδια υλοποίησης της εφαρμογής

4.1 Πειραματικό μέρος

4.1.1 Καταγραφή δεδομένων (Recording)

4.1.1.1 Περιγραφή EPOC+

Σε αυτό το κεφάλαιο θα περιγράψουμε το σύστημα που χρησιμοποιήθηκε για την καταγραφή EEG και EOG σημάτων. Αρχικά, για την καταγραφή των σημάτων χρησιμοποιήθηκε η συσκευή του εργαστηρίου η οποία είναι το EPOC+ Neuroheadset της εταιρείας EMOTIV. Η συσκευή είναι ένα εμπορικό προϊόν το οποίο χρησιμοποιείται και σε πολλές ερευνητικές εφαρμογές. Μαζί με την συσκευή παρέχεται και το hydration pack, μέσα στο οποίο φυλάσσονται οι αισθητήρες για

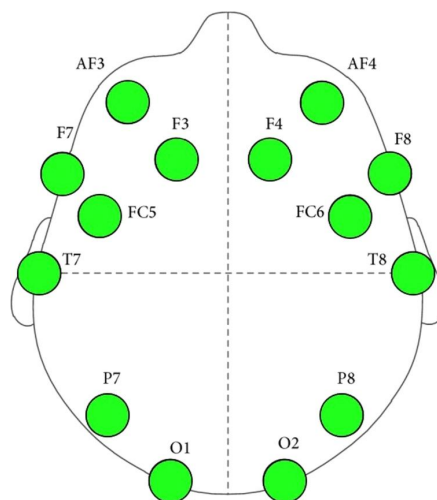
την καταγραφή EEG σημάτων. Οι μετρήσεις στέλνονται ασύρματα μέσω Bluetooth και λαμβάνονται από τον υπολογιστή μέσω ενός universal USB receiver.



Εικόνα 4.2: EMOTIV EPOC

4.1.1.1.1 Τεχνικά χαρακτηριστικά

Το headset έχει συνολικά 16 κανάλια, οι αισθητήρες καταγραφής όμως είναι ουσιαστικά 14, διότι οι άλλοι 2 χρησιμοποιούνται ως γείωση (GND) και ως αναφορά (Reference). Όπως φαίνεται στο παρακάτω σχήμα, τα 14 χρησιμοποιούμενα κανάλια είναι τα AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, O1, O2, ενώ η γείωση και η αναφορά είναι ή τα T7, T8 ή εναλλακτικά τα P3, P4 αντίστοιχα.



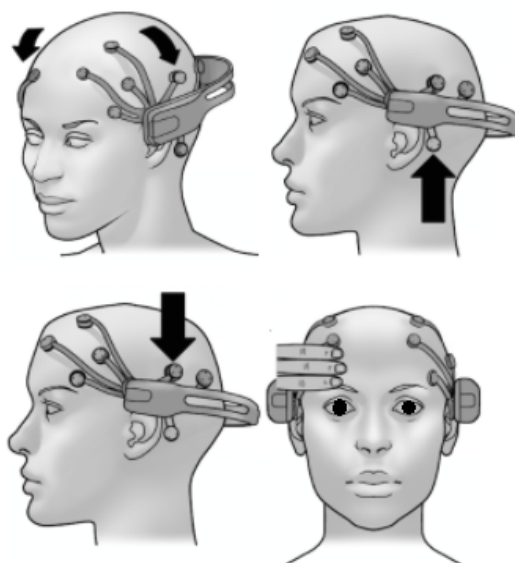
Εικόνα 4.3: Ηλεκτρόδια EPOC

Η συχνότητα δειγματοληψίας των EEG σημάτων είναι $f_s = 128 \text{ Hz/channel}$ και προεπεξεργάζονται κατάλληλα για την απομάκρυνση θορύβου και παρεμβολών.

4.1.2 Experimental setup

4.1.2.1 Προετοιμασία headset για χρήση

Είναι σημαντική η σωστή προετοιμασία και χρήση του EPOC, έτσι ώστε να έχουμε τα καλύτερα δυνατά αποτελέσματα. Πρώτο βήμα είναι να εμποτίσουμε τους αισθητήρες οι οποίοι βρίσκονται μέσα στο hydration rack με αλατούχο διάλυμα για να έχουμε καλύτερη επαφή του αισθητήρα με το δέρμα του κεφαλιού και επομένως καλύτερο σήμα. Στην συνέχεια, τοποθετούμε τους αισθητήρες στο headset και σύμφωνα με το σύστημα 10-20 που αναφέρθηκε παραπάνω το προσαρμόζουμε στο κεφάλι μας, προκειμένου να βρίσκονται στην σωστή θέση. Πιο πρακτικά μπορούμε να διορθώσουμε την θέση του με τα δάχτυλά μας όπως φαίνεται στο σχήμα που ακολουθεί.



Εικόνα 4.4: Τοποθέτηση και προσαρμογή συσκευής

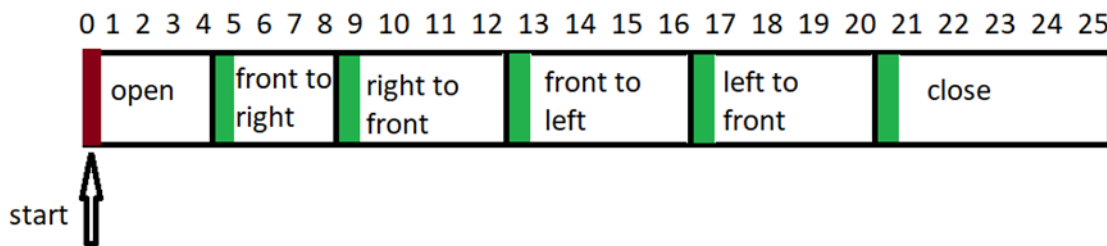
4.1.2.2 Επιλογή καναλιών

Είναι προφανές ότι η επιλογή των καναλιών γίνεται ανάλογα με το ποια EEG σήματα θέλουμε να χρησιμοποιήσουμε. Αυτό διότι όπως γνωρίζουμε κάθε κανάλι - ηλεκτρόδιο βρίσκεται σε συγκεκριμένη περιοχή του κεφαλιού, η οποία συνδέεται με συγκεκριμένες λειτουργίες του εγκεφάλου. Στην συγκεκριμένη περίπτωση που μελετάμε τις κινήσεις των ματιών, αφού κάναμε διάφορες δοκιμές και συμβουλευτήκαμε την σχετική βιβλιογραφία, καταλήξαμε στην χρήση των

ηλεκτροδίων O1, O2 και F7, F8. Τα κανάλια O1, O2, συνδέονται άμεσα με το άνοιγμα/κλείσιμο των ματιών, ενώ τα F7, F8 τα επιλέξαμε γιατί εμφανίζουν συγκεκριμένη συμπεριφορά στις οριζόντιες κινήσεις των ματιών. Παραπάνω πληροφορίες θα αναφερθούν στο στάδιο της προεπεξεργασίας όπου και θα δούμε αναλυτικά τα σήματα που καταγράφονται.

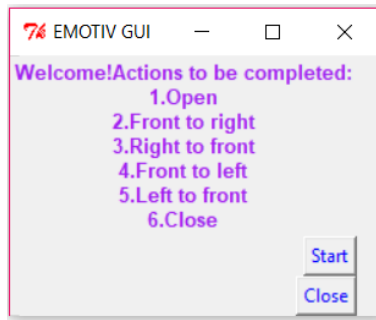
4.1.2.3 Μέθοδος καταγραφής δεδομένων - GUI (Graphic User Interface)

Αφού έχουμε τοποθετήσει κατάλληλα το headset, το συνδέουμε μέσω Bluetooth με τον υπολογιστή και με το λογισμικό που έχουμε υλοποιήσει. Τα προγράμματα που συνδέουν το headset με τον υπολογιστή τα έχουμε γράψει σε Python έτσι ώστε να έχουμε real time καταγραφή και στην συνέχεια αποθήκευση των δεδομένων από όλα τα κανάλια. Σημειώνεται ότι τα δεδομένα λαμβάνονται σε μορφή ζεύγους, δηλαδή {τιμή, ποιότητα} για κάθε κανάλι. Η τιμή της ποιότητας είναι μία ένδειξη για το πόσο καλό είναι το λαμβανόμενο σήμα, και επιθυμούμε να είναι συνήθως τουλάχιστον τριψήφιος αριθμός. Προκειμένου να έχουμε ομοιόμορφη καταγραφή των δεδομένων από τον χρήστη δημιουργήσαμε ένα **πρωτόκολλο καταγραφής** τους, με την βοήθεια ενός γραφικού περιβάλλοντος (GUI). Με αυτό τον τρόπο ο χρήστης έχει την δυνατότητα να καταγράφει τα δεδομένα από τις 4 κινήσεις των ματιών, δηλαδή άνοιγμα/κλείσιμο και κίνηση των ματιών δεξιά/αριστερά. Τα δεδομένα αποθηκεύονται σε .csv αρχεία για μελλοντική χρήση. Σημειώνεται επίσης ότι το GUI μπορεί να μεταβληθεί έτσι ώστε ο χρήστης να καταγράφει όποια κίνηση από τις 4 επιθυμεί. Πιο αναλυτικά, ο χρόνος καταγραφής των δεδομένων είναι συνολικά 25 sec και η αλληλουχία των κινήσεων είναι αυτές που αναγράφονται συνοπτικά στην αρχική οθόνη και είναι 6 συνολικά.



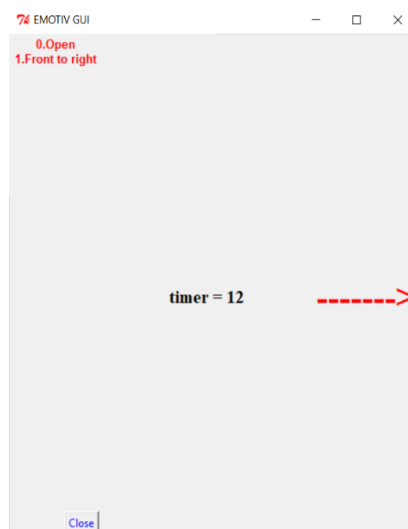
Εικόνα 4.5: Αλληλουχία κινήσεων GUI

Ως 'front' ορίζουμε την ευθεία θέση των ματιών, δηλαδή όταν ο χρήστης κοιτάει μπροστά στην ίδια κατεύθυνση με το κεφάλι. Παρακάτω φαίνεται η αρχική οθόνη της διεπαφής.

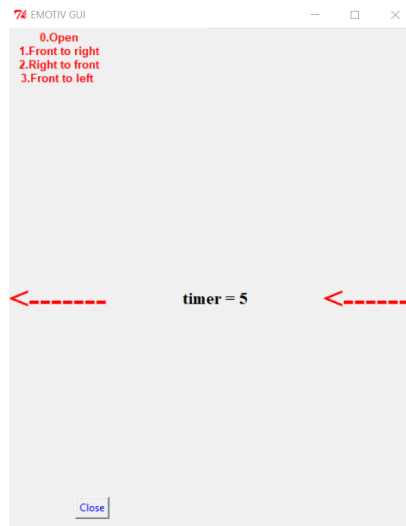


Εικόνα 4.6: Αρχική οθόνη διεπαφής

Είναι σημαντικό να ειπωθεί ότι οι κινήσεις {front to right, left to front} παράγουν ουσιαστικά το ίδιο αποτέλεσμα καθώς το βλέμμα ακολουθεί την ίδια κυκλική πορεία, δηλαδή από τα δεξιά προς τα αριστερά όπως και οι κινήσεις {front to left, right to front}, στις οποίες το βλέμμα κινείται από αριστερά προς τα δεξιά. Η κάθε κίνηση διαρκεί 4 δευτερόλεπτα, με εξαίρεση την τελευταία που διαρκεί 5. Η αλληλουχία των κινήσεων είναι αρκετά εύκολο να απομνημονευθεί, καθώς αποτελεί μία κυκλική πορεία των ματιών ξεκινώντας από το κέντρο, στην συνέχεια δεξιά, μετά πάλι στο κέντρο και επαναλαμβάνοντας την ίδια κίνηση από το κέντρο προς τα αριστερά. Η εναλλαγή των κινήσεων ακολουθείται από την εμφάνιση ενός βέλος προς την αντίστοιχη κατεύθυνση και παράλληλα από ένα προειδοποιητικό σύντομο ήχο του ενός δευτερολέπτου. Έτσι, ο χρήστης έχει την δυνατότητα είτε να κοιτά την οθόνη είτε να περιμένει να ακούσει τον ήχο προκειμένου να μεταβεί στην επόμενη κίνηση. Καθόλη την διάρκεια της καταγραφής εμφανίζεται ο απομεινάντας χρόνος μέσω της μεταβλητής timer καθώς επίσης και η κίνηση η οποία εκτελείται μαζί με όσες έχουν ήδη εκτελεστεί, προκειμένου να είναι ενήμερος ο χρήστης. Όταν ο χρήστης είναι έτοιμος πατάει το κουμπί 'Start' και ξεκινάει αυτόματα η καταγραφή. Στην συνέχεια, παρουσιάζονται δύο περιπτώσεις της διεπαφής για δεξιά και αριστερή κίνηση αντίστοιχα.



Εικόνα 4.7: GUI: Κίνηση προς τα δεξιά



Εικόνα 4.8: GUI: Κίνηση προς τα αριστερά

4.2 Προεπεξεργασία και εξαγωγή χαρακτηριστικών

Όπως έχουμε αναφέρει σε προηγούμενο κεφάλαιο τα EEG σήματα είναι ιδιαίτερα θορυβώδη και γι' αυτό τον λόγο χαρακτηρίζονται και από χαμηλό σηματοθορυβικό λόγο (SNR). Επομένως απαραίτητο στάδιο για την περαιτέρω χρήση τους είναι η κατάλληλη επεξεργασία, έτσι ώστε να απομακρυνθεί όσο το δυνατό περισσότερο ο θόρυβος. Στην παρούσα διπλωματική δοκιμάσαμε διάφορα φίλτρα για την αποθορυβοποίηση των καταγραφόντων σημάτων. Γενικά εφαρμόσαμε σε όλα τα κανάλια κάποιο υψιπερατό φίλτρο για να αφαιρέσουμε την DC συνιστώσα από τα σήματά, αλλά κατά τα άλλα η επιλογή των φίλτρων για κάθε κανάλι έγινε ξεχωριστά ανάλογα με το φασματικό περιεχόμενο που θέλαμε να κρατήσουμε. Παρακάτω θα αναφέρουμε αναλυτικά τα φίλτρα που εφαρμόσαμε για τα ζεύγη καναλιών {O1, O2} και {F7, F8}.

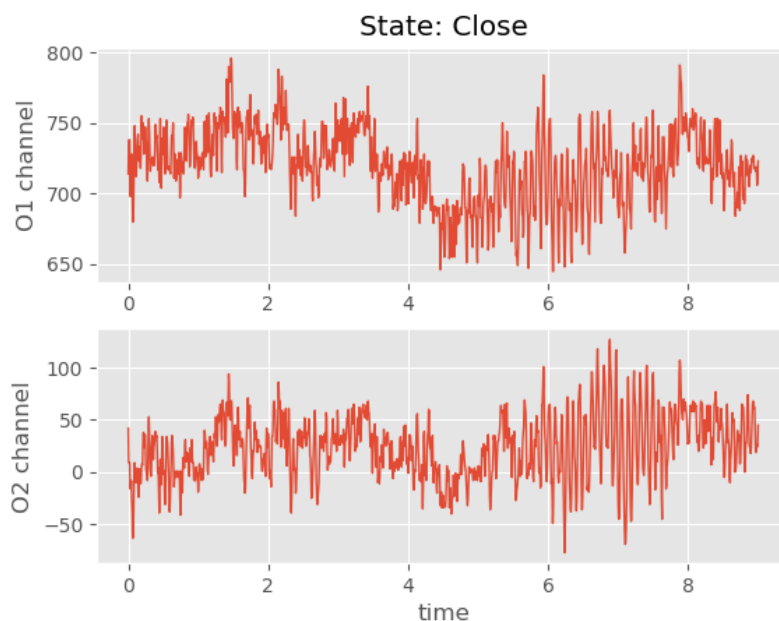
4.2.1 Κανάλια O1-O2

Βάση των ανωτέρω, τα ηλεκτρόδια O1, O2 συνδέονται άμεσα με το άνοιγμα / κλείσιμο των ματιών. Το τελευταίο γίνεται εμφανές από τις διακυμάνσεις των άλφα κυμάτων, τα οποία εντοπίζονται στο εύρος συχνοτήτων από 8-14 Hz. Εξαιτίας αυτού, καταλήξαμε ότι η καλύτερη επιλογή ήταν να εφαρμόζουμε ένα **ζωνοπερατό φίλτρο Butterworth 4^{ης} τάξης με συχνότητες $f_{low} = 5Hz$, $f_{high} = 20Hz$** . Ειδικότερα, παρατηρείται ότι όταν κλείνουμε τα μάτια μας παρουσιάζεται μία απότομη αύξηση των άλφα κυμάτων, η οποία μοιάζει με μία Dirac συνήθως στα 10 - 11 Hz. Για να γίνει ορατή αυτή η μεταβολή είναι αναγκαία η μεταφορά του σήματος στο πεδίο των συχνοτήτων. Αυτή η μεταφορά του σήματος από το πεδίο του χρόνου επιτυγχάνεται με την εφαρμογή του **διακριτού μετασχηματισμού Fourier (DFT)**, με συχνότητα δειγματοληψίας $f_s = 128Hz$. Έτσι, με την χρήση του DFT εξάγονται τα απαραίτητα φασματικά χαρακτηριστικά του σήματος, τα οποία παρουσιάζονται στην συνέχεια. Ακολουθούν οι γραφικές παραστάσεις των σημάτων με χρόνο καταγραφής 10 δευτερόλεπτα στο πεδίο του χρόνου και της συχνότητας,

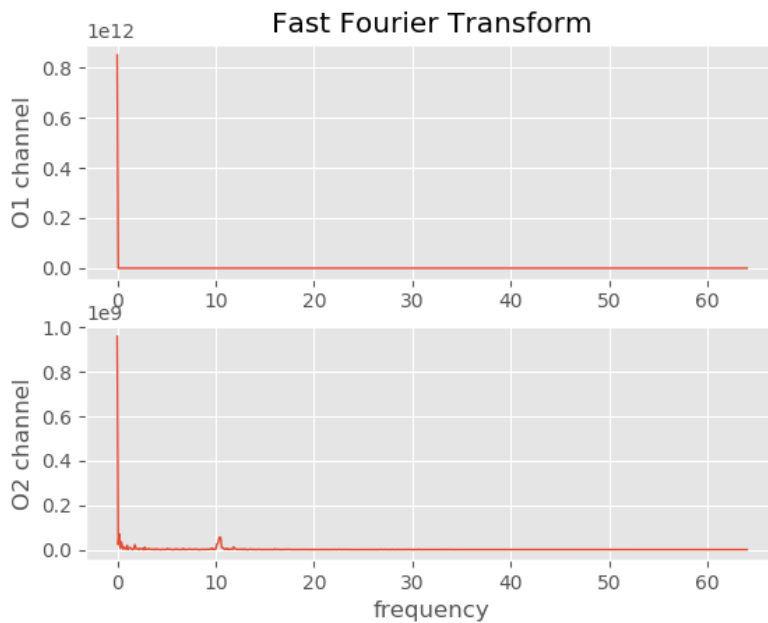
επεξεργασμένα και μη, προκειμένου να είναι εμφανής η διαφορά. Σημειώνεται ότι το κλείσιμο των ματιών πραγματοποιείται από τα 5-8sec. Επίσης, για τον λόγο ότι μιλάμε για την ισχύ των σημάτων στο πεδίο των συχνοτήτων απεικονίζεται το τετράγωνο των τιμών των καναλιών O1, O2.

4.2.1.1 Για την κατάσταση ανοιχτών-κλειστών ματιών (State:Close)

Αρχικά όσον αφορά τα αρχικά σήματα τα οποία περιλαμβάνουν και άνοιγμα και κλείσιμο των ματιών, είναι εμφανές ότι περιέχουν μεγάλη dc συνιστώσα. Μιλώντας για το πεδίο του χρόνου παρατηρείται στην περιοχή όπου ο χρήστης κλείνει τα μάτια του μία αύξηση των τιμών και της συχνότητας του σήματος. Παράλληλα βλέπουμε και στο πεδίο της συχνότητας ότι η dc συνιστώσα είναι πολύ μεγάλη, γεγονός που μας εμποδίζει να βγάλουμε κάποιο συμπέρασμα για το υπόλοιπο φάσμα.



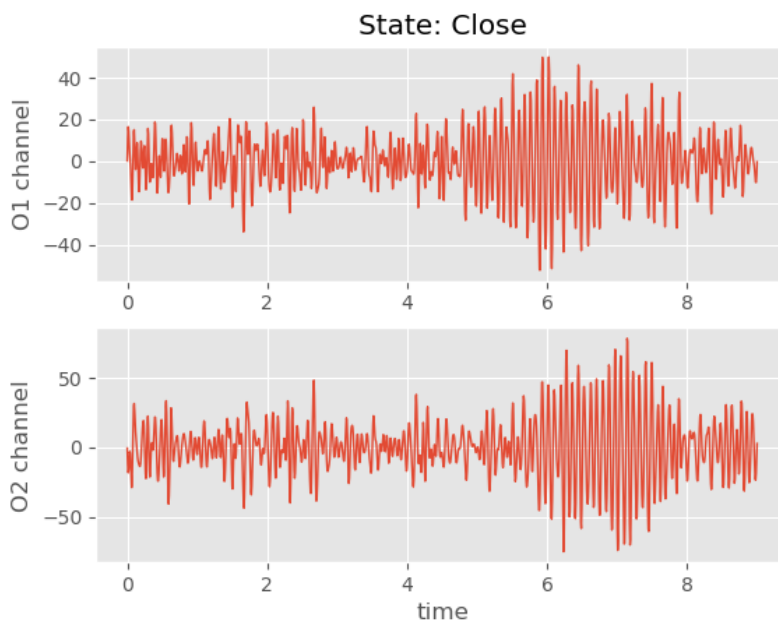
Εικόνα 4.9: Αρχικά σήματα α) πεδίο του χρόνου



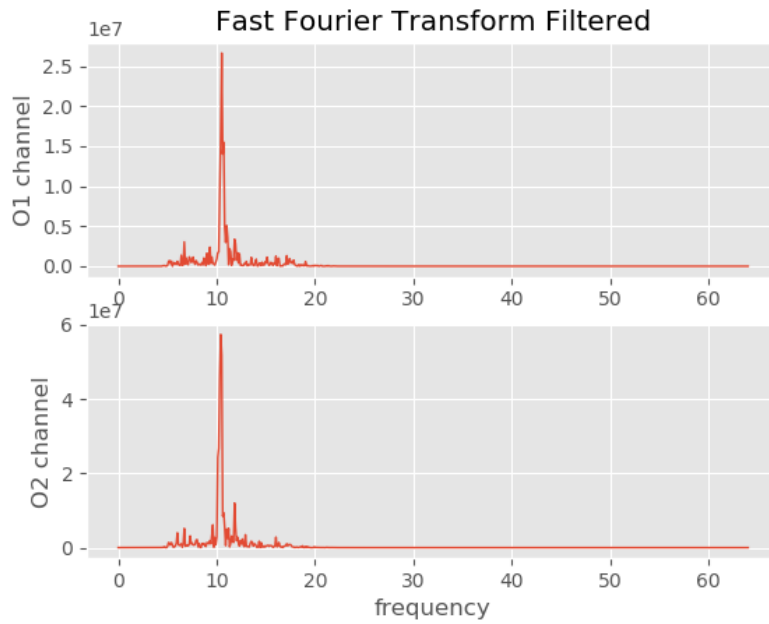
Εικόνα 4.10: Αρχικά σήματα β) πεδίο της συχνότητας

4.2.1.1.1 Εφαρμογή ζωνοπερατού φίλτρου

Εάν εφαρμόσουμε το φίλτρο που προαναφέρθηκε παίρνουμε τα παρακάτω αποτελέσματα. Φαίνεται ότι η dc συνιστώσα έχει εξαλειφθεί, κάτι το οποίο μας επιτρέπει να έχουμε μία πιο ξεκάθαρη εικόνα κυρίως στο πεδίο της συχνότητας, όπου είναι εμφανής η 'dipac' γύρω από τα 10-11Hz, η οποία είναι της τάξης των 10^7 (V^2).



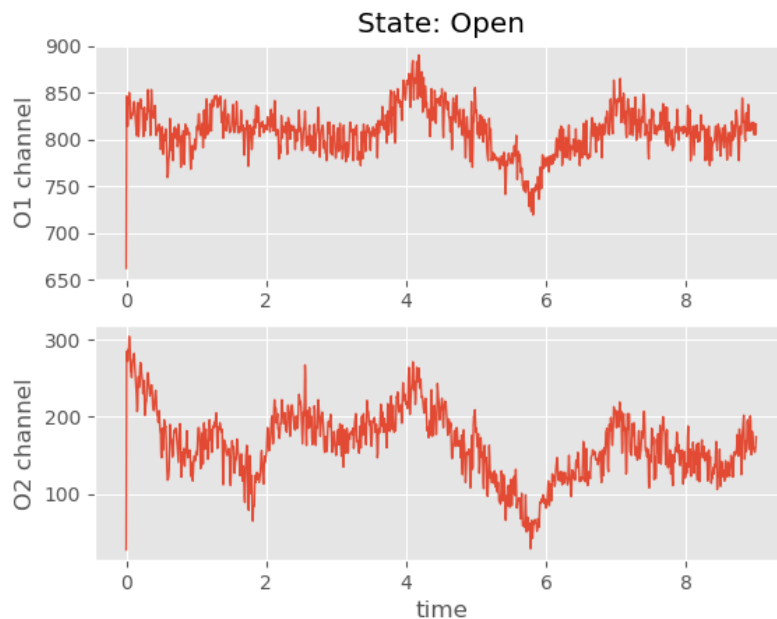
Εικόνα 4.11: Φιλτραρισμένα σήματα α) πεδίο του χρόνου



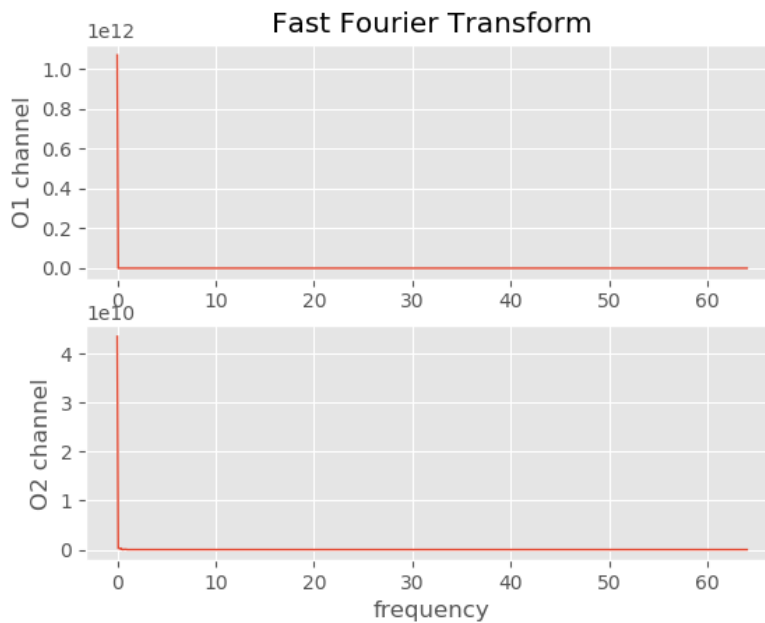
Εικόνα 4.12: Φιλτραρισμένα σήματα β) πεδίο της συχνότητας

4.2.1.2 Για την κατάσταση των ανοιχτών ματιών (State:Open)

Ακολουθούν οι αντίστοιχες γραφικές παραστάσεις οι οποίες είναι τα αποτελέσματα καταγραφής 10 δευτερολέπτων όπου ο χρήστης έχει τα μάτια του συνεχώς ανοιχτά. Όπως και προηγουμένως παρατηρούμε την μεγάλη dc συνιστώσα των σημάτων, τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας.



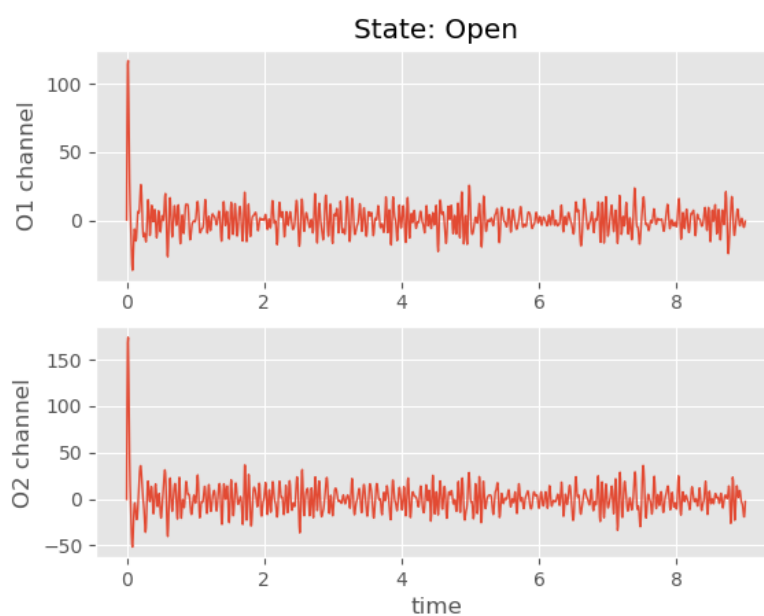
Εικόνα 4.13: Αρχικά σήματα α) πεδίο του χρόνου



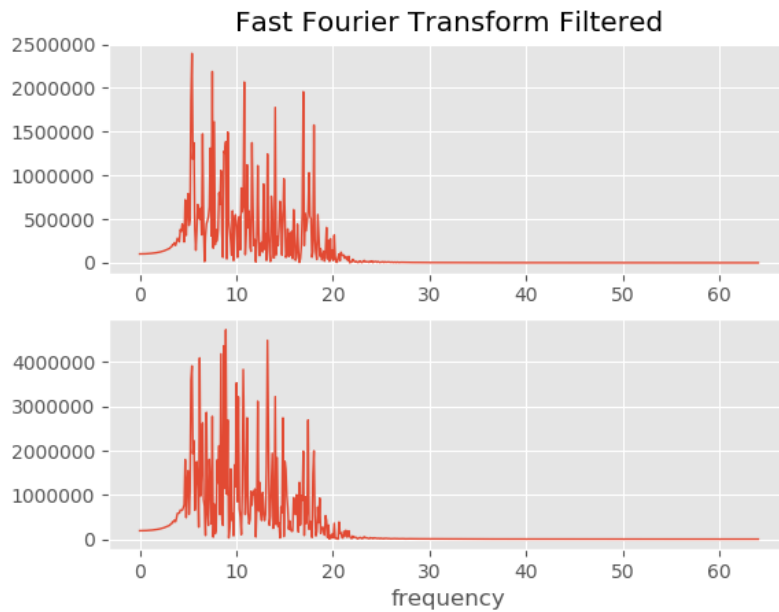
Εικόνα 4.14: Αρχικά σήματα θ) πεδίο της συχνότητας

4.2.1.2.1 Εφαρμογή ζωνοπερατού φίλτρου

Με την εφαρμογή του φίλτρου, δεν παρατηρούμε καμία ιδιαίτερη μεταβολή των τιμών του σήματος όσο ο χρήστης έχει ανοιχτά τα μάτια του. Όσον αφορά το πεδίο της συχνότητας η μέγιστη τιμή της ισχύος είναι μία τάξη μεγέθους κάτω σε σχέση με την περίπτωση του κλεισίματος των ματιών και παράλληλα το φάσμα που μας ενδιαφέρει παρουσιάζει σχετική ομοιομορφία ως προς τις τιμές.



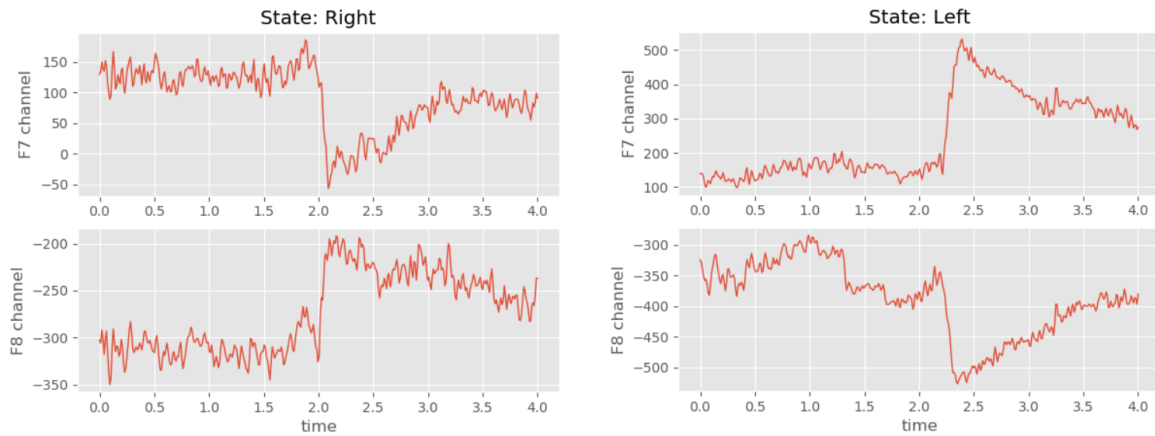
Εικόνα 4.15: Φιλτραρισμένα σήματα O1, O2 α) πεδίο του χρόνου



Εικόνα 4.16: Φιλτραρισμένα σήματα O1, O2 θ) πεδίο της συχνότητας

4.2.2 Κανάλια F7-F8

Τα κανάλια που επιλέγονται κανονικά και συχνότερα με βάση την βιβλιογραφία για τις κινήσεις των ματιών είναι τα C3 - C4, αλλά όπως φαίνεται το EPOC δεν διαθέτει την επιλογή αυτή. Ως επακόλουθο αναζητήσαμε άλλες λύσεις έτσι ώστε να πάρουμε τα κατάλληλα χαρακτηριστικά που χρειαζόμαστε για τον διαχωρισμό των κινήσεων. Τελικά καταλήξαμε ότι ορθότερη επιλογή είναι τα κανάλια F7 - F8, διότι με αυτά μπορούμε να καταγράψουμε τα ζητούμενα EOG σήματα και να εξάγουμε χαρακτηριστικά τα οποία βοηθούν στον διαχωρισμό των οριζόντιων κινήσεων. Πιο αναλυτικά με βάση τις κινήσεις των ματιών τα 2 σήματα παρουσιάζουν τις εξής συμπεριφορές στο πεδίο του χρόνου: **για την κίνηση του βλέμματος από το κέντρο προς τα δεξιά (State: Right)**, παρατηρείται μείωση των τιμών του F7 και αύξηση αυτών του F8 αντίστοιχα. Αντιθέτως **για κίνηση των ματιών από το κέντρο προς τα αριστερά (State: Left)** παρουσιάζεται αύξηση στις τιμές του F7 και αντίστοιχη μείωση των τιμών του F8. Κάτι τέτοιο είναι αναμενόμενο καθώς οι δύο κινήσεις είναι αντίθετες μεταξύ τους. Ακολουθούν οι γραφικές παραστάσεις των αρχικών σημάτων διάρκειας 4 δευτερολέπτων. Σημειώνεται ότι η εκάστοτε κίνηση πραγματοποιείται περίπου στην μέση της καταγραφής, όπου παρατηρείται και αυτή η απότομη μεταβολή των τιμών.

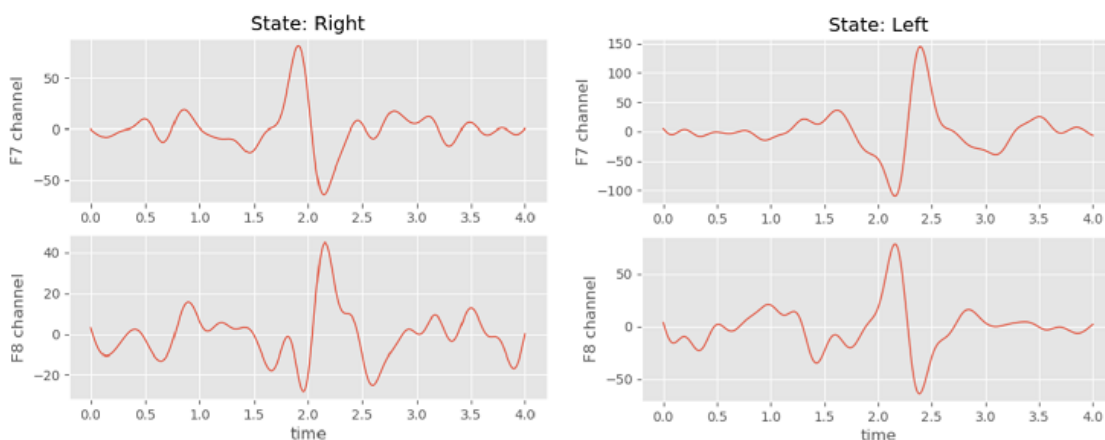


Εικόνα 4.17: Αρχικά σήματα των F7, F8 καναλιών

Παρόλου που τα παραπάνω σήματα παρουσιάζουν τις επιθυμητές συμπεριφορές που περιεγράφηκαν προηγουμένως, βλέπουμε ότι είναι ιδιαίτερα θορυβώδη. Έτσι προκειμένου να απομακρυνθεί ο πρόσθετος θόρυβος, εφαρμόζονται διάφορες τεχνικές αποθορυβοποίησης, οι οποίες παρατίθενται παρακάτω.

4.2.2.1 Ζωνοπερατό φίλτρο με $f_{low} = 1\text{Hz}$, $f_{high} = 5\text{Hz}$.

Γνωρίζουμε ότι το φασματικό περιεχόμενο που μας ενδιαφέρει βρίσκεται στις χαμηλές συχνότητες, έτσι με την εφαρμογή αυτού του φίλτρου κόβεται όλος ο υψίσυχνος θόρυβος, αλλά και η dc συνιστώσα η οποία μετατοπίζει το αρχικό σήμα. Κατά συνέπεια λαμβάνονται τα παρακάτω γραφικά αποτελέσματα για τις 2 κινήσεις (Right, Left).

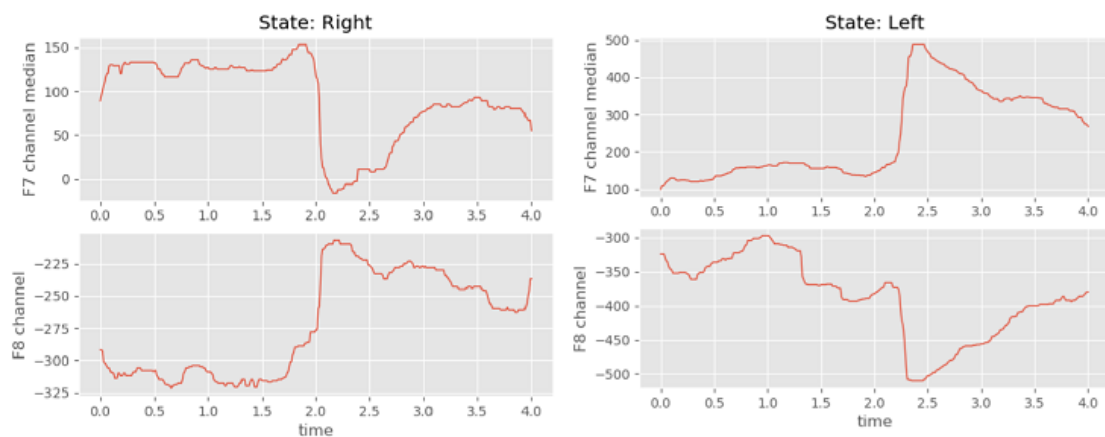


Εικόνα 4.18: Αποτέλεσμα εφαρμογής ζωνοπερατού φίλτρου στα κανάλια F7, F8

Παρατηρείται ότι λόγω του ζωνοπερατού φίλτρου έχουμε κατάλληλη μεταβολή του φασματικού περιεχομένου, πράγμα χρήσιμο διότι είναι πιο εμφανείς οι μεταβολές του σήματος που συσχετίζονται με κάποια κίνηση των ματιών.

4.2.2.2 Median φίλτρο

Το συγκεκριμένο φίλτρο χρησιμοποιείται συχνά για εξομάλυνση (smoothing) του σήματος, για αυτό τον λόγο δοκιμάστηκε και τα αποτελέσματα παρουσιάζονται ακολούθως.

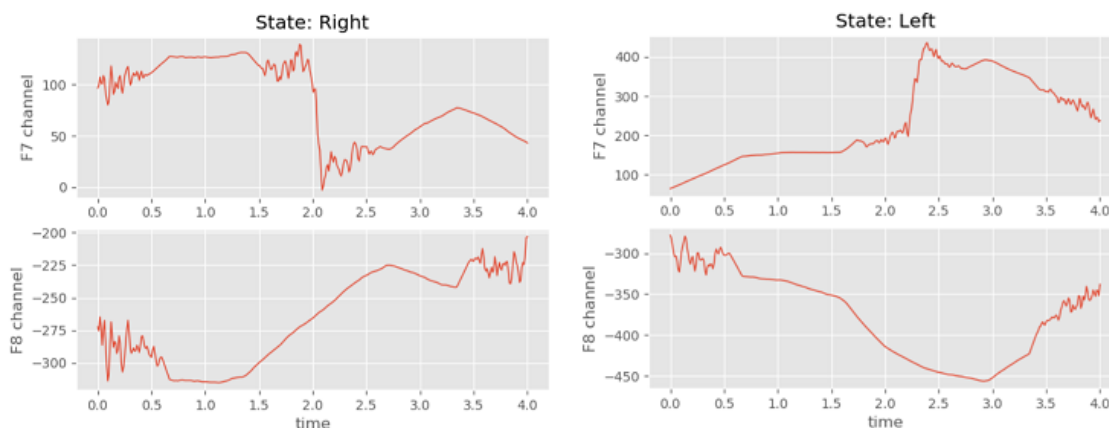


Εικόνα 4.19: Αποτέλεσμα εφαρμογής median φίλτρου στα κανάλια F7, F8

Βλέπουμε ότι το συγκεκριμένο φίλτρο μειώνει αισθητά τον πρόσθετο θόρυβο παρόλα αυτά δεν κάνει πιο έντονη την μεταβολή του σήματος κάτι το οποίο επιθυμούμε στην συγκεκριμένη περίπτωση, ούτε αφαιρεί την dc συνιστώσα.

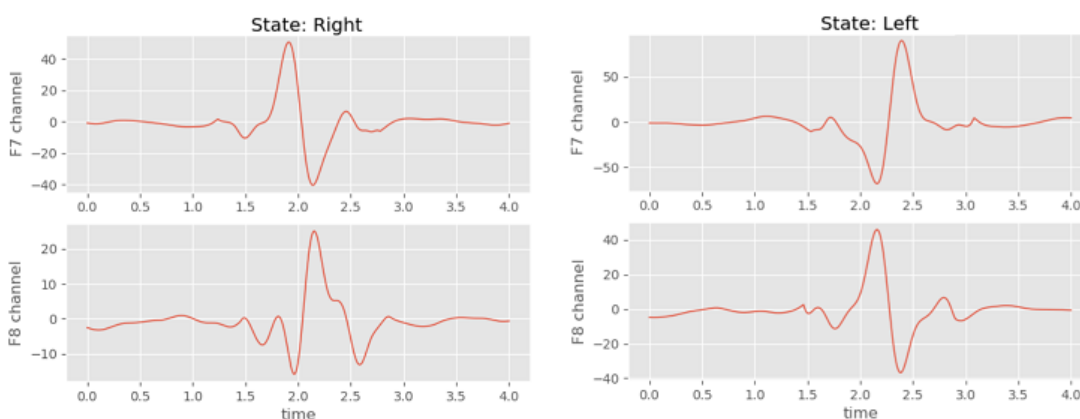
4.2.2.3 Wiener φίλτρο

Παρατηρούμε ότι με την χρήση του φίλτρου αυτού εξομαλύνεται σε μεγάλο βαθμό το σήμα το οποίο μας είναι χρήσιμο στην προκειμένη περίπτωση, αλλά και σε αυτή την περίπτωση δεν υπάρχει αύξηση της μεταβολής του σήματος αντιθέτως σημειώνεται εξομάλυνση και της διακύμανσης αυτής. Επίσης όπως και προηγουμένως η dc συνιστώσα παραμένει.



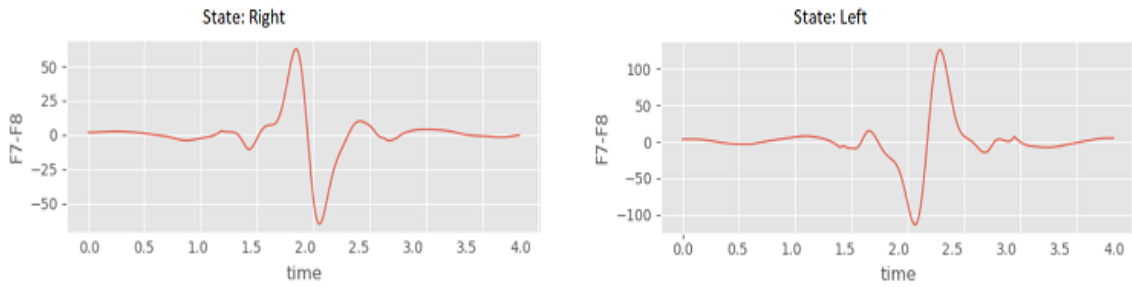
Εικόνα 4.20: Αποτέλεσμα εφαρμογής wiener φίλτρου στα κανάλια F7, F8

Έπειτα δοκιμάσαμε συμπληρωματικά τους συνδυασμούς των παραπάνω φίλτρων για να βρεθεί εν τέλει ο πιο αποτελεσματικός τρόπος φιλτραρίσματος των σημάτων F7, F8, για να μπορούμε στην συνέχεια να εξάγουμε τα επιθυμητά χαρακτηριστικά. Έτσι καταλήξαμε ότι η εφαρμογή πρώτα ενός ζωοπερατού φίλτρου και στην συνέχεια του wiener είναι η καλύτερη μέθοδος αποθορυβοποίησης για τα συγκεκριμένα σήματα, καθώς ο συνδυασμός τους έχει ως αποτέλεσμα την εξάλειψη του θορύβου και την ενίσχυση της μεταβολής που θέλουμε. Ακολουθούν τα τελικά αποτελέσματα.



Εικόνα 4.21: Αποτέλεσμα εφαρμογής συνδυασμού βαθυπερατού και wiener φίλτρου στα κανάλια F7, F8

Συμπερασματικά, είναι φανερό ότι οι παραπάνω γραφικές παραστάσεις είναι αντίθετες μεταξύ τους και παρουσιάζουν πρώτα ένα μέγιστο και μετά ένα ελάχιστο στην περιοχή όπου πραγματοποιείται η εκάστοτε κίνηση και αντίστροφα. Τα χαρακτηριστικά επομένως που επιθυμούμε να εξάγουμε για κάθε κίνηση είναι αυτή η απότομη μεταβολή, η οποία είναι σαφώς πιο εύκολα διαχωρίσιμη σε σχέση με το αρχικό σήμα. Εφόσον λοιπόν θέλουμε όσο το δυνατό εντονότερη μεταβολή, εφαρμόζουμε το ίδιο φιλτράρισμα με πριν, τώρα όμως **στην διαφορά των δύο σημάτων, δηλαδή στο F7-F8.**



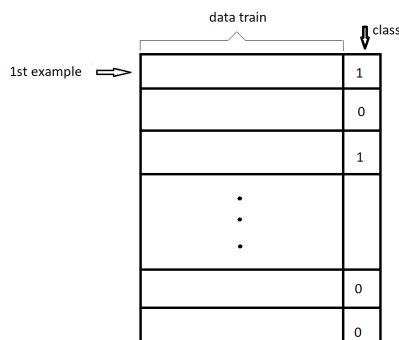
Εικόνα 4.22: Διαφορά των καναλιών F7, F8

Η εφαρμογή φιλτραρίσματος στην διαφορά των δύο σημάτων μπορεί να βοηθήσει επίσης και στην εξάλειψη διαφόρων τυχαίων υψηλών τιμών, οι οποίες δεν σχετίζονται με κάποια κίνηση, αλλά αποτελούν σφάλματα στο μετρούμενο σήμα και εμφανίζονται και στα δύο κανάλια.

4.3 Δημιουργία δεδομένων εκπαίδευσης (training dataset)

Πριν προχωρήσουμε στην περαιτέρω περιγραφή και ανάλυση του συστήματος ταξινόμησης αξίζει να σημειωθεί ο τρόπος δημιουργίας του κάθε συνόλου δεδομένων εκπαίδευσης. Ένα training dataset είναι ένας πίνακας δεδομένων ο οποίος είναι διαστάσεων $M \times N$. Ο αριθμός των γραμμών M αποτελεί το πλήθος των δεδομένων που καταγράφηκαν, ενώ το πλήθος στηλών $N-1$ εκφράζει το μέγεθος, δηλαδή τον χρόνο της κάθε καταγραφής. Η N -οστή στήλη του πίνακα αποτελεί την κλάση στην οποία βρίσκεται το συγκεκριμένο δείγμα δεδομένων. Ουσιαστικά κάθε δείγμα δεδομένων (κάθε παράδειγμα) είναι ένα ζεύγος του δείγματος δεδομένων με την αντίστοιχη κλάση στην οποία ανήκει. Υπενθυμίζεται ότι τα δεδομένα εκπαίδευσης καταγράφηκαν με βάση το πρωτόκολλο που δημιουργήθηκε. Επίσης είναι φανερό ότι, τα δεδομένα εκπαίδευσης έχουν περάσει πρώτα από τα παραπάνω στάδια επεξεργασίας, και στην συνέχεια τους έχει προστεθεί η κλάση στην οποία ανήκουν. Με αυτόν το τρόπο δημιουργήθηκαν τα 3 training datasets που περιγράφονται παρακάτω για την εκπαίδευση των ταξινομητών.

Εικόνα 4.23: Πίνακας δεδομένων εκπαίδευσης



4.4 Artefacts (Σφάλματα)

Είναι φανερό ότι τα EEG και EOG σήματα είναι ιδιαίτερα επιρρεπή στο θόρυβο, ο οποίος δημιουργείται συνήθως είτε από παρεμβολές (πχ άλλων ηλεκτρικών συσκευών) είτε από διάφορα είδη σφαλμάτων. Για αυτό τον λόγο είναι ιδιαίτερα σημαντικός ο εντοπισμός και η γνώση των αιτιών του θορύβου έτσι ώστε να λαμβάνονται συγκεκριμένα μέτρα για να καταγράφονται τα καλύτερα δυνατά σήματα. Θα μπορούσαμε επομένως να χωρίσουμε τα σφάλματα ανάλογα με την προέλευσή τους στις εξής κατηγορίες:

Της συσκευής

Όπως έχει προαναφερθεί η συσκευή EMOTIV EPOC είναι μία εμπορική συσκευή καταγραφής ηλεκτροεγκεφαλικών σημάτων. Κατά συνέπεια είναι λογικό επακόλουθο να μην παρουσιάζει την ίδια ακρίβεια με αντίστοιχες συσκευές οι οποίες είναι αποκλειστικά εργαστηριακής χρήσης. Πιο συγκεκριμένα η ποιότητα των αισθητήρων και ο τρόπος τοποθέτησης του headset, ο οποίος δεν είναι απόλυτα προκαθορισμένος, συμβάλλουν αρνητικά στην ακρίβεια των μετρήσεων. Επιπλέον είναι σημαντικό να σημειωθεί ότι η σύνδεση της συσκευής και η μετάδοση των δεδομένων στον υπολογιστή είναι πολλές φορές κακή, με αποτέλεσμα να χάνεται η σύνδεση και να μην λαμβάνονται όλα τα δεδομένα από την συσκευή. Για αυτούς τους λόγους κατά την διάρκεια της προετοιμασίας και καταγραφής θα πρέπει να είμαστε ιδιαίτερα προσεχτικοί όσον αφορά την ποιότητα των αισθητήρων αλλά και της σύνδεσης. Ακόμα και αν εξασφαλίσουμε την ποιότητα των δύο προηγούμενων παραγόντων, παρατηρούνται πολλές φορές τυχαίες μεγάλες τιμές στα δεδομένα που καταγράφονται χωρίς προφανή λόγο, οι οποίες πρέπει να εξαιρεθούν όσο γίνεται στο στάδιο της αποθρομβοποίησης. Παρόλα αυτά είναι ιδιαίτερα ενδιαφέρουσα και χρήσιμη η μελέτη και στην συνέχεια η αξιοποίηση των δυνατοτήτων μίας τέτοιας συσκευής, καθώς είναι πιο οικονομική και προσιτή στο ευρύτερο κοινό.

Του χρήστη

Έκτος από τα σφάλματα λόγω της συσκευής σημαντικός παράγοντας είναι και ο τρόπος χρήσης της. Αρχικά, ο χρήστης θα πρέπει να γνωρίζει τους παραπάνω περιορισμούς και τις δυνατότητες της συσκευής έτσι ώστε να την χρησιμοποιεί σωστά. Όπως αναφέρθηκε προηγουμένως η τοποθέτηση της συσκευής επακριβώς σύμφωνα με το σύστημα 10-20 είναι λίγο δύσκολη με αποτέλεσμα πολλές φορές να αποκλίνει από την σωστή της θέση, γεγονός που επηρεάζει την ακρίβεια των μετρήσεων. Επίσης, είναι γεγονός ότι τα EEG σήματα επηρεάζονται ιδιαίτερα από άλλου είδους σήματα, ιδιαίτερα τα μυϊκά τα οποία παράγονται από την κίνηση μυών του προσώπου. Για αυτό τον λόγο ο χρήστης θα πρέπει κατά την διάρκεια των μετρήσεων να είναι χαλαρός και εκτελεί μόνο τις κινήσεις εκείνες οι οποίες είναι απαραίτητες για την διεξαγωγή του πειράματος.

Του περιβάλλοντος

Τελευταίος και εξίσου σημαντικός παράγοντας που επηρεάζει την ποιότητα των δεδομένων είναι το περιβάλλον στο οποίο λαμβάνουν χώρα οι μετρήσεις. Για να έχουμε το καλύτερο δυνατό αποτέλεσμα είναι αναγκαίο τα πειράματα να γίνονται σε απόλυτα ελεγχόμενο περιβάλλον, δηλαδή να είναι για παράδειγμα ελεγχόμενη η θερμοκρασία, η φωτεινότητα και η παρουσία άλλων ηλεκτρικών συσκευών οι οποίες μπορεί να παρεμβάλλουν στα EEG σήματα.

4.5 Μηχανική Μάθηση

Στο σημείο αυτό θα παραθέσουμε τις διάφορες τεχνικές μηχανικής μάθησης που χρησιμοποιήσαμε για τον διαχωρισμό των σημάτων 01, 02 και F7, F8. Είναι γνωστό όμως ότι τα συστήματα αυτά προκειμένου να μπορούν να ταξινομήσουν νέα δείγματα δεδομένων πρέπει πρώτα να έχουν εκπαιδευθεί κατάλληλα. Για αυτό τον λόγο δημιουργήσαμε ένα δικό μας σύνολο δεδομένων εκπαίδευσης (training dataset) για κάθε κατηγορία σήματος. Πιο αναλυτικά, για την καταγραφή των σημάτων και για τις 4 κατηγορίες συμμετείχαν 6 εθελοντές. Για τις κατηγορίες open/close καταγράφηκαν συνολικά 101 παραδείγματα των 4 δευτερολέπτων (256 samples) και για τις κατηγορίες κίνησης των ματιών δεξιά/αριστερά καταγράφηκαν επίσης 103 παραδείγματα των 4 δευτερολέπτων (512 samples). Είναι σημαντικό να σημειωθεί ότι λόγω θορύβου και σφαλμάτων τα δείγματα τα οποία επιλέχθηκαν να συμπεριληφθούν, ελέγχθηκαν πρώτα ποιοτικά έτσι ώστε να αντιπροσωπεύουν την κάθε κλάση.

Προκειμένου να επιλεγεί ο βέλτιστος αλγόριθμος για την ταξινόμηση των σημάτων, ξεκινήσαμε δοκιμάζοντας διάφορους τρόπους μηχανικής μάθησης για τον διαχωρισμό των δεδομένων που πάρθηκαν από τα 01, 02, για τις κλάσεις: κλάση0: open (ανοιχτά μάτια) και κλάση1:close (κλειστά μάτια). Έχοντας φτιάξει τα σύνολα των δεδομένων εκπαίδευσης δοκιμάσαμε τις παρακάτω τεχνικές ταξινόμησης για τον διαχωρισμό των σημάτων αυτών:

4.5.1 Αλγόριθμος KNN

Αρχικά, χρησιμοποιήθηκε ο αλγόριθμος KNN, η εκπαίδευση και η επικύρωση του οποίου έγινε με την μέθοδο cross validation για $k=9$. Δοκιμάστηκαν διαφορετικές τιμές για τις μεταβλητές:

- **n_neighbors**: ορίζει τον αριθμό των γειτόνων που θα λάβει υπόψιν ο αλγόριθμος. Συγκεκριμένα δοκιμάστηκε το σύνολο τιμών [2, 5, 10, 20].
- **weights**: καθορίζει τα βάρη των γειτόνων και για την οποία έχουμε 2 επιλογές:
 - i) uniform: όλα τα σημεία της γειτονικής περιοχής είναι ισοσταθμισμένα
 - ii) distance: τα σημεία της γειτονιάς έχουν βάρος αντιστρόφως ανάλογο της απόστασης από το σημείο που βρίσκεται προς ταξινόμηση. Με άλλα λόγια οι γείτονες οι οποίοι βρίσκονται πιο κοντά στο σημείο που πρέπει να κατηγοριοποιηθεί έχουν μεγαλύτερη επιρροή από αυτούς που βρίσκονται πιο μακριά.

έτσι ώστε να ελέγξουμε την διακύμανση της ακρίβειας και να επιλεγεί τελικά ο βέλτιστος συνδυασμός παραμέτρων, ο οποίος παρουσιάζεται στον παρακάτω πίνακα. Σημειώνεται ότι η ακρίβεια (accuracy) υπολογίζεται από την μέθοδο cross validation, ως ο μέσος όρος των αποτελεσμάτων που προκύπτουν από κάθε K-fold.

Παράμετροι	Βέλτιστη επιλογή
n_neighbors	3
weights	distance
accuracy	93%
deviation	0.13%

Παρατηρούμε ότι η ακρίβεια που επιτυγχάνεται με τον συγκεκριμένο αλγόριθμο είναι αρκετά καλή, με μόνο πρόβλημα την σχετικά μεγάλη απόκλιση που παρουσιάζει. Κάτι τέτοιο ίσως είναι και λόγω των δεδομένων εκπαίδευσης που συλλέχθηκαν, αλλά παρόλα αυτά το αποτέλεσμα είναι ικανοποιητικό.

4.5.2 SVM

Στην συνέχεια, εκπαιδύσαμε μία μηχανή διανυσμάτων υποστήριξης ομοίως με την μέθοδο cross validation για k=10. Στην συγκεκριμένη περίπτωση έπρεπε να βρεθούν οι βέλτιστες τιμές των παραμέτρων του SVM έτσι ώστε να έχουμε την καλύτερη δυνατή απόδοση για τον διαχωρισμό των κλάσεων. Οι παράμετροι και οι επιλογές που χρησιμοποιήθηκαν είναι οι παρακάτω:

Παράμετροι	Πιθανές επιλογές
C	0.5, 1, 10 : καθορίζει την απόσταση των κλάσεων
kernel	linear, poly, rbf, sigmoid
coefo	0, 1, 2 : παράμετρος μόνο για τις επιλογές poly, sigmoid
max_iter	10, 20, 50, 70 : όριο επαναλήψεων μέσα στον solver (default:-1, κανένα όριο)
decision function shape	ovo, onr
gamma	0.01, 0.001, 0.0001

Πίνακας 1: Παράμετροι SVM

Η βελτιστοποίηση των παραμέτρων έγινε με την χρήση του αλγόριθμου **grid search**. Ο αλγόριθμος αυτός λειτουργεί κάνοντας exhaustive search μέσα σε ένα προκαθορισμένο υποσύνολο των τιμών των παραμέτρων, οι οποίες χρησιμοποιούνται στον συγκεκριμένο αλγόριθμο μηχανικής μάθησης. Η μέθοδος του exhaustive search είναι ένας τρόπος επίλυσης προβλημάτων κατά τον οποίο υπολογίζονται όλοι οι δυνατοί συνδυασμοί των παραμέτρων και επιλέγεται με βάση κάποιο performance metric ο καλύτερος. Αυτό το κριτήριο επιλογής, δηλαδή το performance metric, υπολογίζεται από την εκπαίδευση και επικύρωση του εκάστοτε πιθανού SVM από το cross validation. Στην συγκεκριμένη περίπτωση χρησιμοποιείται ως μετρική η ακρίβεια (accuracy).

Παράμετροι	Βέλτιστη επιλογή
C	0.5
kernel	linear
coef0	0
max_iter	-1
decision function shape	ovo
accuracy	94%
deviation	0.11%

Πίνακας 2: Βέλτιστο SVM

Σύμφωνα με τον παραπάνω πίνακα τα αποτελέσματα που λαμβάνονται από το βελτιστοποιημένο SVM είναι αρκετά ικανοποιητικά, καθώς επίσης παρατηρούμε ότι η απόκλιση της ακρίβειας σε αυτή την περίπτωση είναι λίγο μικρότερη.

4.5.3 Neural Networks

Ο τελευταίος τρόπος μηχανικής μάθησης που δοκιμάστηκε ήταν τα Νευρωνικά Δίκτυα. Ως διάνυσμα χαρακτηριστικών στην είσοδο του Νευρωνικού, πήραμε τα δείγματα που είχαν καταγραφεί για τις κλάσεις open/close. Οι παράμετροι οι οποίοι μεταβλήθηκαν κατά την διάρκεια της εκπαίδευσης του δικτύου παρατίθενται στον παρακάτω πίνακα:

Παράμετροι	Πιθανές επιλογές
hidden_layer_sizes	10, 20, 80, 100
activation	logistic, relu, sigmoid
alpha	0.1, 0.001, 0.0001, 0.0001, 0.00001, 0.000001
learning_rate	0.001
max_iter	200, 300, 350
tol	0.0001

Πίνακας 3: Παράμετροι Νευρωνικού Δικτύου

Ομοίως με προηγουμένως η βελτιστοποίηση των παραμέτρων έγινε με τον αλγόριθμο grid search. Συνεπώς το βέλτιστο Νευρωνικό δίκτυο για τα συγκεκριμένα δεδομένα είναι :

Παράμετροι	Βέλτιστες τιμές
hidden_layer_sizes	200
activation	tanh
alpha	0.1
learning_rate	0.001
max_iter	200
tol	0.0001
accuracy	90%
deviation	0.12%

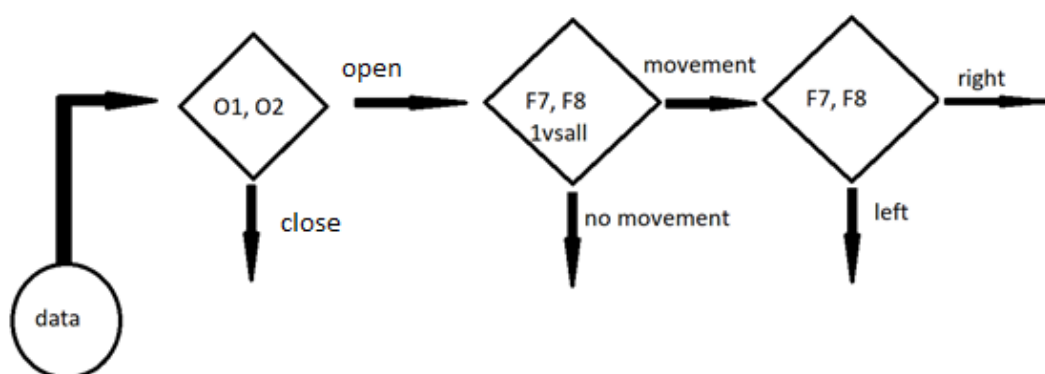
Πίνακας 4: Βέλτιστο Νευρωνικό Δίκτυο

Άξιο αναφοράς είναι το γεγονός ότι ο χρόνος εκπαίδευσης του δικτύου σε σχέση με τις δύο προηγούμενες μεθόδους ήταν πολύ μεγαλύτερος. Κάτι τέτοιο ήταν αναμενόμενο καθώς η πολυπλοκότητα του αλγορίθμου εκπαίδευσης και η βελτιστοποίηση των παραμέτρων είναι υπολογιστικά δυσκολότερη. Όσον αφορά την ακρίβεια του Νευρωνικού Δικτύου βλέπουμε ότι είναι και αυτή αρκετά υψηλή, αλλά χαμηλότερη σε σχέση με τους άλλους 2 τρόπους ταξινόμησης.

5 Κεφάλαιο: Συνολικό σύστημα απόφασης

5.1 Γενική περιγραφή συστήματος

Λαμβάνοντας όλα τα παραπάνω αποτελέσματα του Κεφαλαίου 4 υπόψιν, συμπεραίνουμε ότι το SVM έχει μεγαλύτερη ή τουλάχιστον ίση απόδοση για το συγκεκριμένο πρόβλημα που θέλουμε να επιλύσουμε και παρουσιάζει επίσης την μικρότερη απόκλιση από τις άλλες δύο μεθόδους. Για αυτό το λόγο η δημιουργία του συνολικού συστήματος θα γίνει με βάση τα SVM. Όπως έχουμε αναφέρει σκοπός του συστήματος είναι η ταξινόμηση 4 κινήσεων : άνοιγμα/κλείσιμο των ματιών (open/close) και δεξιά/αριστερή κίνηση τους (right/left). Για αυτό τον λόγο υλοποιήσαμε 3 SVM για την ταξινόμηση των 4 προαναφερθέντων κλάσεων, συνδεδεμένα εν σειρά. Αρχικά, το πρώτο SVM λαμβάνει την απόφαση για το αν τα δεδομένα εισόδου ανήκουν στην κατηγορία των ανοιχτών ή κλειστών ματιών. Εάν αποφασιστεί ότι ανήκουν στην πρώτη κλάση, τότε ο έλεγχος μεταβαίνει στο δεύτερο SVM, το οποίο αποφασίζει για το αν υπάρχει κίνηση ή όχι των ματιών. Στην συνέχεια με την ίδια λογική εάν το 2^ο SVM βγάλει την απόφαση ότι υπάρχει κίνηση, ο έλεγχος μεταβαίνει στο 3^ο και τελευταίο SVM της αλυσίδας το οποίο ταξινομεί την κίνηση σε δεξιά και αριστερή. Για να είναι πιο κατανοητή η αρχιτεκτονική του συστήματος, παρατίθεται παρακάτω το σχηματικό διάγραμμα του, το οποίο απεικονίζει τον τρόπο σύνδεσης των ταξινομητών.



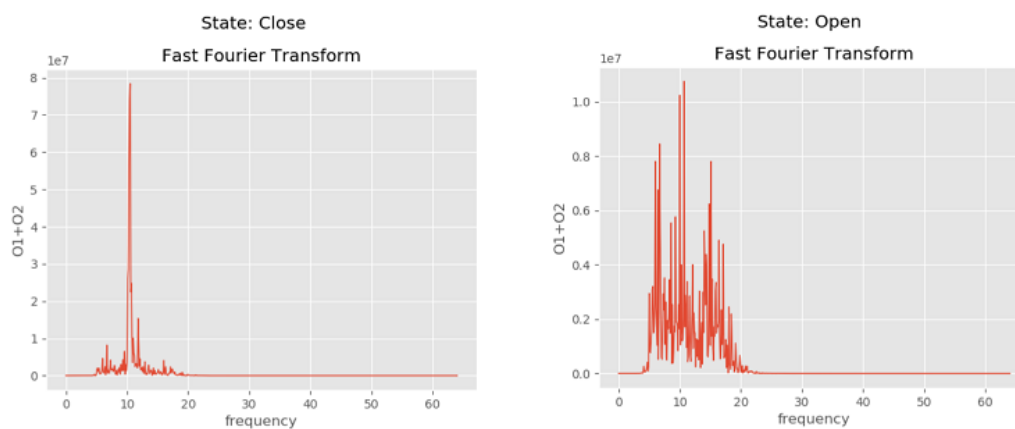
Εικόνα 5.1: Αρχιτεκτονική συστήματος απόφασης

Σημειώνεται ότι κάθε SVM υλοποιήθηκε πρώτα ξεχωριστά και έπειτα δημιουργήθηκε το συνολικό σύστημα απόφασης. Η υλοποίηση κάθε ταξινομητή ξεχωριστά ήταν απαραίτητη, έτσι ώστε να γίνει σωστά η εκπαίδευση και η επικύρωση του, για τα εκάστοτε δεδομένα εκπαίδευσης που είχαν καταγραφεί για την συγκεκριμένη κίνηση. Όσον αφορά την εκπαίδευση του κάθε SVM σημαντικό ήταν να έχουμε επιλέξει τα κατάλληλα δεδομένα τα οποία θα περιέχουν τα επιθυμητά χαρακτηριστικά για την σωστή ταξινόμηση. Για να ελεγχθεί το τελευταίο κάναμε για κάθε υποψήφιο σύνολο δεδομένων εκπαίδευσης grid search έτσι ώστε να γίνει βελτιστοποίηση των παραμέτρων του ταξινομητή. Δοκιμάστηκαν πολλά training sets τα οποία διαφοροποιούνταν ως

προς τα χαρακτηριστικά που επιλέγαμε και ως προς το πλήθος των δειγμάτων καταγραφής. Έτσι, συγκρίνοντας τα αποτελέσματα του ταξινομητή με την μέθοδο grid search που προαναφέρθηκε, επιλέξαμε αυτά που μας δίνουν την καλύτερη ακρίβεια για κάθε κίνηση. Η επιλογή των χαρακτηριστικών του σήματος της κάθε κλάσης και οι βέλτιστοι παράμετροι του κάθε SVM παρατίθενται ακολούθως.

5.1.1 1^ο SVM

Το πρώτο SVM στην αλυσίδα του συστήματος παίρνει ως είσοδο δεδομένα από τα κανάλια O1, O2. Όπως έχει προαναφερθεί το SVM πρέπει να είναι κατάλληλα εκπαιδευμένο έτσι ώστε να μπορεί να γίνει σωστά η ταξινόμηση. Η εκπαίδευση του συγκεκριμένου SVM έγινε για 2 κλάσεις: **class0**-> **open** και **class1**-> **close**. Για την κλάση 0 καταγράφηκαν 46 παραδείγματα των 4 δευτερολέπτων, δηλαδή $4 \cdot 128 = 512$ δείγματα (samples). Για την κλάση 1 καταγράφηκαν 55 παραδείγματα των 4 δευτερολέπτων, δηλαδή 512 δείγματα. Τα χαρακτηριστικά των δεδομένων εκπαίδευσης έχουν εξαχθεί από το φασματικό περιεχόμενο των σημάτων όπως έχει αναφερθεί παραπάνω και έτσι τα δείγματα που τελικά εισάγονται στο dataset για κάθε καταγραφή είναι 256 samples. Σημειώνεται ότι εν τέλει ως διάνυσμα χαρακτηριστικών επιλέξαμε το άθροισμα των καναλιών O1, O2 στο πεδίο της συχνότητας, διότι αυτό μας έδωσε την καλύτερη ακρίβεια. Παρατηρείται ότι το training set είναι σχεδόν ισοσταθμισμένο.



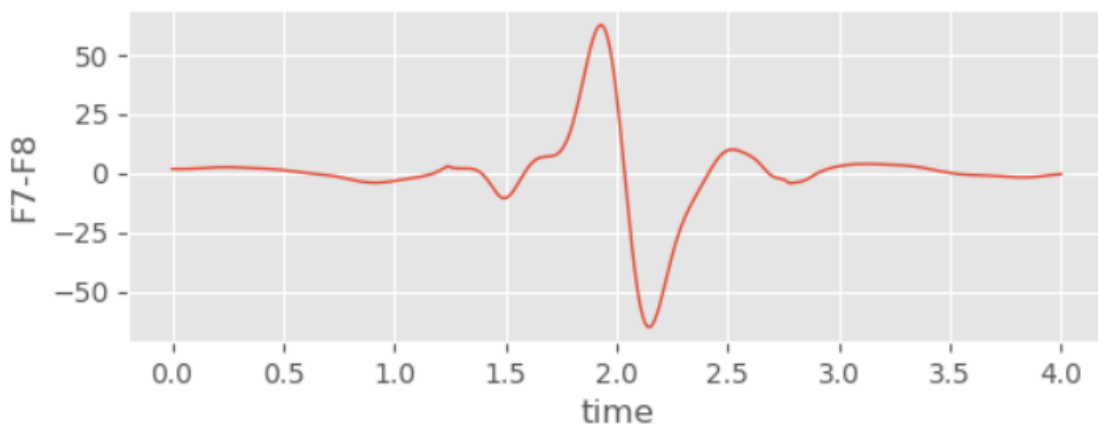
Εικόνα 5.2: Χαρακτηριστικά δεδομένων εκπαίδευσης 1ου SVM

Σχετικά με την μορφή των 256 δεδομένων κάθε παραδείγματος που εισάγονται στο σύνολο δεδομένων εκπαίδευσης, βλέπουμε ότι έχουν αρκετά διαφορετική κατανομή των τιμών και για αυτό τον λόγο, όπως συμπεράναμε προηγουμένως διαχωρίζονται γραμμικά από τον ταξινομητή.

5.1.2 2° SVM

Όπως φαίνεται παραπάνω το δεύτερο SVM λαμβάνει ως είσοδο δεδομένα από τα κανάλια F7, F8. Η εκπαίδευση αυτού του ταξινομητή έγινε για 2 κλάσεις : **class0-> no movement** και **class1-> movement**. Όπως προδίδουν και τα ονόματα των κλάσεων, η πρώτη κλάση αφορά δεδομένα τα οποία έχουν παραχθεί από ακινησία των ματιών του χρήστη, ενώ η δεύτερη από οποιαδήποτε οριζόντια κίνηση των ματιών (δεξιά/αριστερά). Για την κλάση 0 καταγράφηκαν 38 παραδείγματα των 4 δευτερολέπτων, δηλαδή $4 \cdot 128 = 512$ δείγματα (samples) για κάθε παράδειγμα. Για την κλάση 1 καταγράφηκαν 52 παραδείγματα των 4 δευτερολέπτων, δηλαδή 512 δείγματα (samples). Στο συγκεκριμένο SVM παρατηρούμε ότι στην κλάση1 έχουμε ενοποιησει ουσιαστικά 2 κλάσεις μαζί (δεξιά/αριστερά κίνηση). Σημειώνεται επίσης ότι το συγκεκριμένο πρόβλημα ταξινόμησης θα μπορούσε να επιλυθεί και με ένα ταξινομητή πολλαπλών κλάσεων (multiclass svm), ο οποίος λειτουργεί με την τεχνική **1 vs all**.

Βλέπουμε λοιπόν ότι το training set είναι σχετικά ισοσταθμισμένο. Όσον αφορά τα χαρακτηριστικά των δεδομένων εκπαίδευσης έχουν εξαχθεί από το πεδίο του χρόνου των σημάτων όπως έχει αναφερθεί παραπάνω, δηλαδή παίρνουμε την διαφορά των δύο καναλιών. Σε αυτό το σημείο είναι σημαντικό να σημειωθεί ότι ο αριθμός των δειγμάτων που τελικά εισάγονται στο dataset για κάθε καταγραφή είναι 128 samples. Ο τρόπος με τον οποίο γίνεται η επιλογή αυτών των 128 samples είναι ιδιαίτερα σημαντικός και περιγράφεται παρακάτω.



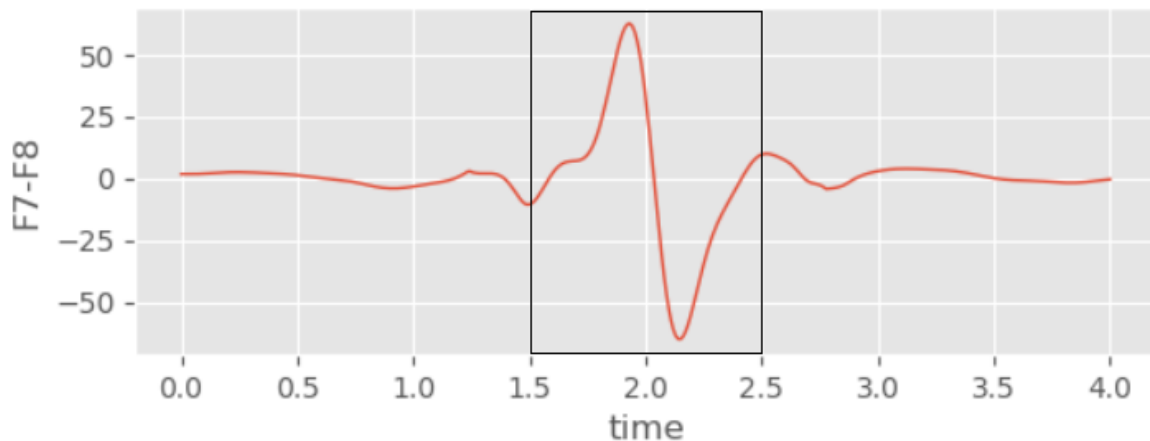
Εικόνα 5.3: Αρχικό σήμα 512 δειγμάτων

Sliding window method

Είναι φανερό από την παραπάνω γραφική παράσταση ότι, η διαφορά των δεδομένων που αντιστοιχούν σε οριζόντια κίνηση των ματιών και αυτών της ακινησίας, είναι η απότομη μεταβολή των τιμών που παρουσιάζουν όταν ανήκουν στην class1. Για αυτό τον λόγο καταλαβαίνουμε ότι απαραίτητα χαρακτηριστικά για τον διαχωρισμό της κλάσης1 από την κλάση0 είναι μόνο αυτή η απότομη μεταβολή, η οποία διαρκεί λίγο λιγότερο από 1 sec, δηλαδή 128 samples. Έτσι για να βρεθούν τα κατάλληλα 128 δείγματα, χρησιμοποιείται η μέθοδος του sliding window. Σύμφωνα με αυτήν δημιουργείται ένα παράθυρο το οποίο μετακινείται κατά

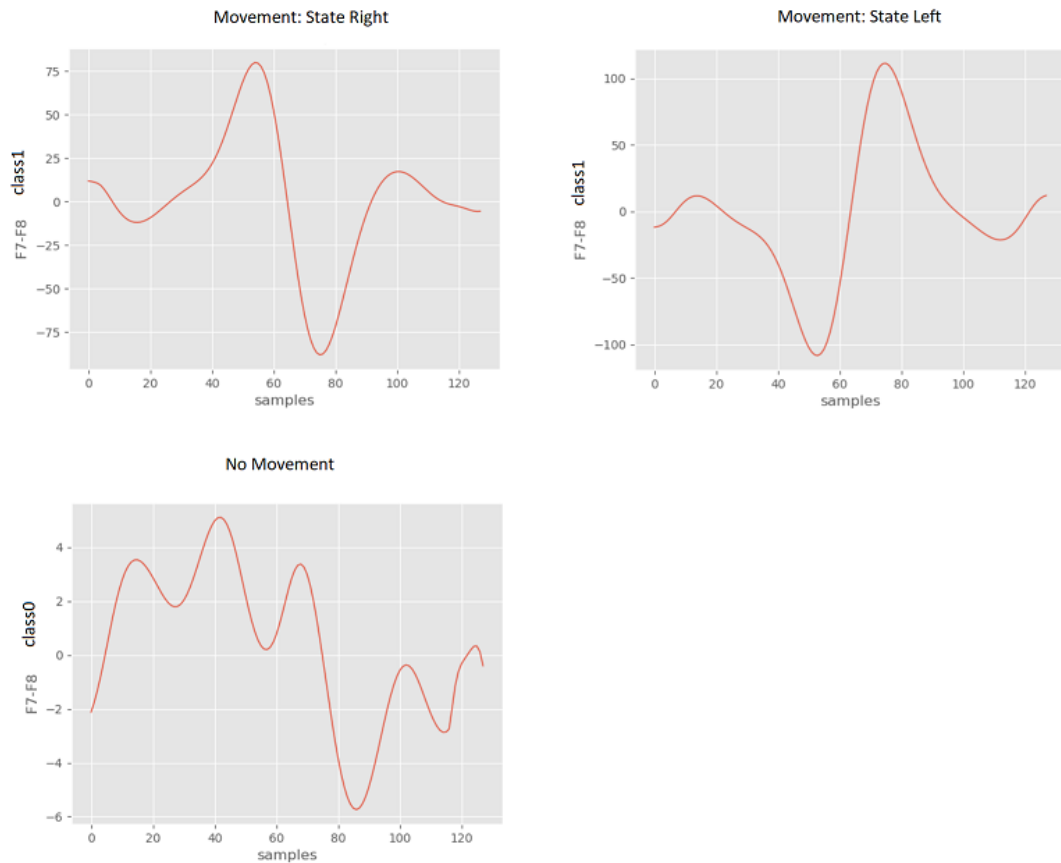
μήκος του εξεταζόμενου σήματος, μέχρι να βρεθούν τα επιθυμητά δείγματα. Στην συγκεκριμένη περίπτωση το παράθυρο αναζήτησης είναι μεγέθους 128 samples και ο αλγόριθμος εύρεσης των δειγμάτων είναι ο εξής:

1. Create sliding window (size: 128 samples)
2. For i in range of the signal:
 - find (max & min)
 - find indexes of (max & min)
 - calculate $index(mean) = \frac{index(max)+index(min)}{2}$
 - center window in correlation with $index(mean)$
3. Crop selected window
4. Return window values to be inserted to the dataset



Εικόνα 5.4: Διαδικασία εύρεσης παραθύρου στο αρχικό σήμα

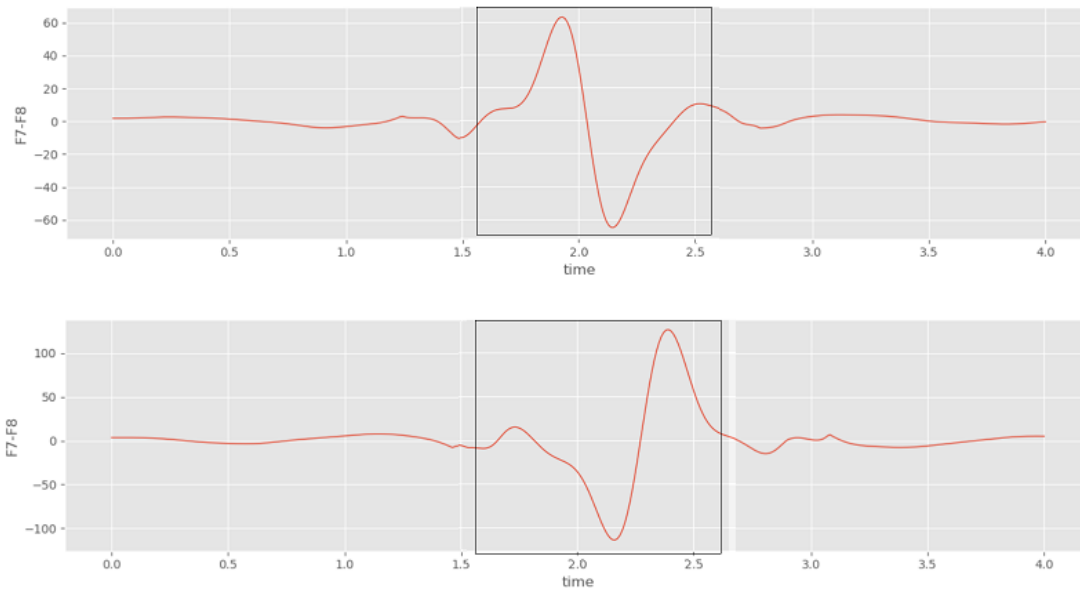
Ουσιαστικά τα 128 δεδομένα του παραθύρου κεντράρονται βάση του $index(mean)$, το οποίο είναι η μέση μεταξύ των θέσεων όπου εντοπίζεται το μέγιστο και το ελάχιστο για το συγκεκριμένο παράθυρο. Συνεπώς με αυτό τον τρόπο παίρνουμε τα δεδομένα που σχετίζονται με την απότομη μεταβολή που προαναφέραμε, τα οποία εισάγονται στο training set και αντιστοιχίζονται σε κάθε κλάση αντίστοιχα. Στις γραφικές που ακολουθούν φαίνονται οι μορφές των δεδομένων εκπαίδευσης για τις 2 κλάσεις. Παρατηρείται ότι στην κλάση1 σε σχέση με την κλάση0, οι τιμές των μεγίστων και ελαχίστων είναι σαφώς μεγαλύτερες. Επίσης βλέπουμε ότι η κυματομορφή της πρώτης κλάσης είναι συγκεκριμένη, θα μπορούσε να παρομοιαστεί με ημιτονοειδή μορφή, και όχι ακανόνιστη όπως στην μηδενική κλάση.



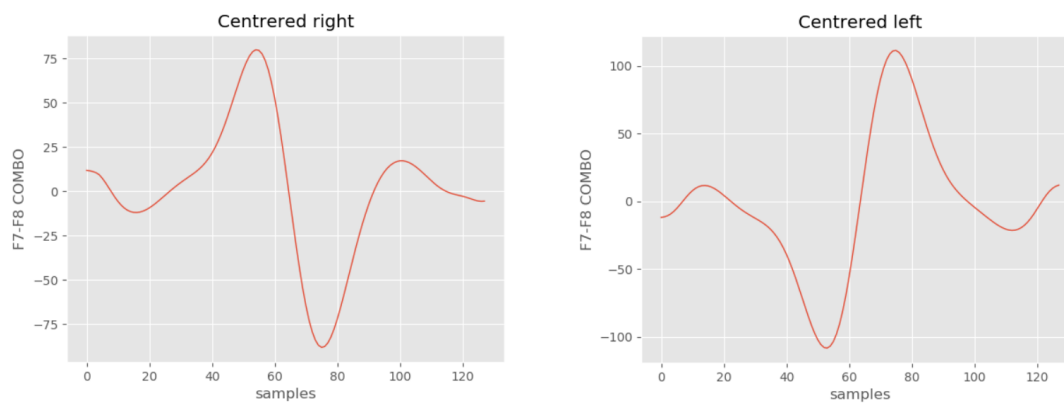
Εικόνα 5.5: Χαρακτηριστικά των δεδομένων εκπαίδευσης του 2ου SVM

5.1.3 3^ο SVM

Το τελευταίο SVM είναι υπεύθυνο για την ταξινόμηση των δεδομένων που αντιστοιχούν σε μία από τις δύο οριζόντιες κινήσεις του ματιού. Έτσι στην **class0** έχουν αντιστοιχηθεί τα δεδομένα που αφορούν την **αριστερή κίνηση** και ως **class1** αυτά της **δεξιάς**. Ουσιαστικά αυτό το SVM διαχωρίζει την κλάση1 της κίνησης (movement) του προηγούμενου SVM σε δύο υποκλάσεις. Συνολικά πάρθηκαν αρχικά 101 παραδείγματα και για τις 2 κλάσεις, 51 για την κλάση1 και 52 για την κλάση2 αντίστοιχα, διάρκειας 4 δευτερολέπτων. Στην συνέχεια, από αυτά τα 4 δευτερόλεπτα εξάχθηκαν τα επιθυμητά χαρακτηριστικά του σήματος, χρησιμοποιώντας τη μέθοδο sliding window που περιεγράφηκε παραπάνω, με αποτέλεσμα το τελικό μήκος κάθε παραδείγματος δεδομένων να είναι 1 sec, δηλαδή 128 samples. Οι ακόλουθες γραφικές παραστάσεις απεικονίζουν την μέθοδο εύρεσης του παραθύρου με τα ζητούμενα χαρακτηριστικά του σήματος (2 πρώτες), για δεξιά και αριστερή κίνηση αντίστοιχα. Οι τελευταίες έχουν δημιουργηθεί μετά από αποκοπή και κεντράρισμα των δεδομένων στο νέο παράθυρο και απεικονίζουν σύμφωνα με τα προαναφερθέντα την μορφή των τελικών χαρακτηριστικών που εισάγονται στο σύνολο δεδομένων εκπαίδευσης.



Εικόνα 5.6: Αρχικά σήματα F7-F8 καναλιών (δεξιά-αριστερή κίνηση)



Εικόνα 5.7: Τελικά σήματα των F7-F8 (128 δείγματα)

Παρατηρείται ότι τα δεδομένα που εισάγονται τελικά στο σύνολο δεδομένων εκπαίδευσης είναι αντίθετα ως προς την φάση τους και επομένως εύκολα γραμμικά διαχωρίσιμα.

5.2 Εύρεση βέλτιστου SVM

Συμπερασματικά έχοντας κάνει grid search με τα συγκεκριμένα δεδομένα εκπαίδευσης που παρουσιάστηκαν παραπάνω, βρίσκουμε τις βέλτιστες παραμέτρους για κάθε ταξινομητή. Όσον αφορά την εκπαίδευση των SVM αυτή έγινε με την μέθοδο cross validation για k=10. Τα τελικά αποτελέσματα του συστήματος παρατίθενται στον ακόλουθο πίνακα.

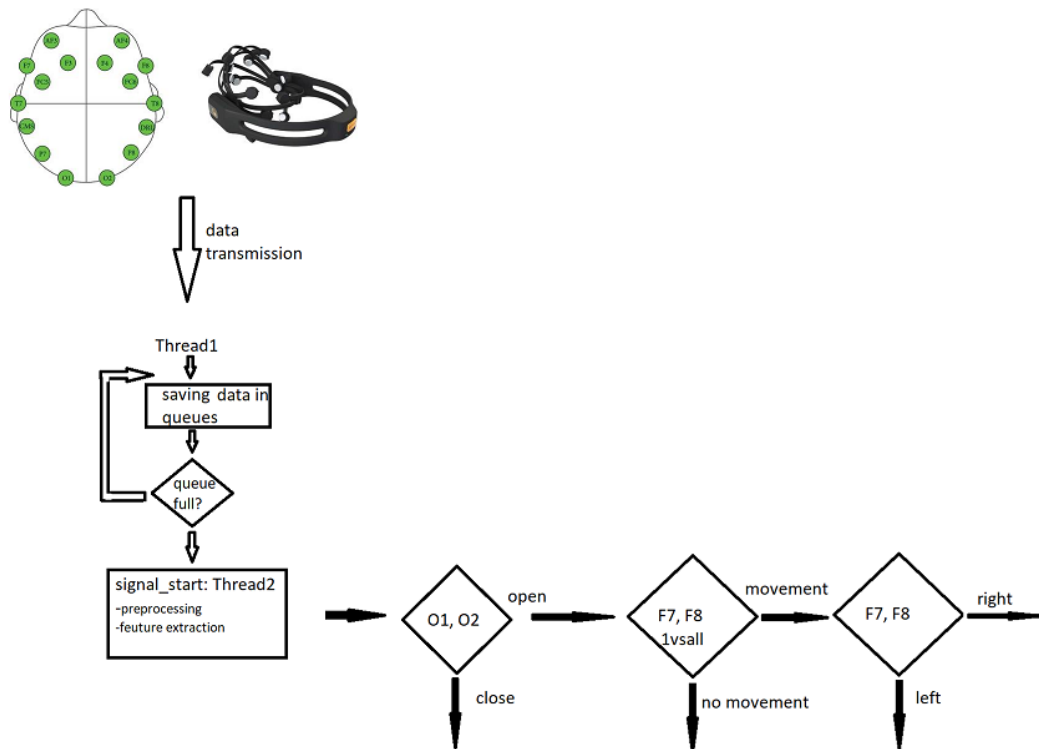
	1st SVM	2nd SVM	3rd SVM
C	0.5	0.5	0.5
kernel	linear	rbf	linear
coef0	0	0	0
max_iter	200	-1	-1
decision function shape	ovo	ovo	ovo
gamma	0.01	0.0001	0.01
accuracy	94%	99%	100%
deviation	0.11%	0.07%	0%

Τα αποτελέσματα όσον αφορά το πρώτο SVM είναι ικανοποιητικά, με μόνη αστοχία την σχετικά μεγάλη απόκλιση που παρουσιάζεται, το οποίο μπορεί να οφείλεται στο σύνολο δεδομένων που χρησιμοποιήθηκε. Σχετικά με την απόδοση του δεύτερου παρατηρούμε ότι ο διαχωρισμός είναι κατά 99% επιτυχής και η απόκλιση μηδενική, ενώ για το 3^ο SVM έχουμε 99% ακρίβεια και 0.07% απόκλιση, γεγονός που αποδεικνύει ότι οι δύο ταξινομητές έχουν βελτιστοποιηθεί σχεδόν στο έπακρον. Πρέπει να σημειωθεί ότι σε σχέση με τα αποτελέσματα που συναντώνται στην βιβλιογραφία, τα αποτελέσματα που βγάλαμε είναι καλύτερα από πολλές παρόμοιες μελέτες.

5.3 Σύστημα πραγματικού χρόνου (Online system)

Στην προηγούμενη ενότητα περιεγράφηκε η αρχιτεκτονική και η εκπαίδευση του συστήματος που υλοποιήθηκε. Εκτός από την offline δοκιμή του συστήματος, το σύστημα αυτό υλοποιήθηκε και δοκιμάστηκε και **σε πραγματικό χρόνο (real time)**. Πιο αναλυτικά, προκειμένου το ήδη υλοποιημένο σύστημα να μπορεί να ανταποκριθεί σε πραγματικό χρόνο, χρειάστηκε η χρήση multithreading. Έτσι δημιουργήθηκαν **2 Threads**, ένα για το διάβασμα και την αποθήκευση των τιμών που καταγράφονται, και ένα δεύτερο για την επεξεργασία τους και την λήψη της απόφασης. Ειδικότερα, το **πρώτο thread** διαβάζει τα δεδομένα των 4 καναλιών (O1, O2, F7, F8) που μεταδίδονται από την συσκευή και τα αποθηκεύει σε 4 ουρές (queues) αντίστοιχα. Η χωρητικότητα των ουρών εξαρτάται από τα πόσα δευτερόλεπτα επιθυμούμε να επεξεργαζόμαστε κάθε φορά, και είναι κάτι που θα μας απασχολήσει στην συνέχεια. Μόλις τελειώσει η καταγραφή των τιμών από το πρώτο Thread, ξεκινά το **δεύτερο thread** να λαμβάνει τις τιμές από τις ουρές και προχωρά στην κατάλληλη προεπεξεργασία των δεδομένων του κάθε

καναλιού όπως προαναφέρθηκε. Θα ακολουθήσει η είσοδος των δεδομένων στην αλυσίδα των SVM προκειμένου να παρθεί η απόφαση για το ποια κίνηση εκτελέστηκε. Σε αυτό το σημείο είναι σημαντικό να σημειωθεί ότι τα SVM έχουν ήδη εκπαιδευτεί με τα training sets που έχουν δημιουργηθεί και βελτιστοποιηθεί. Σημειώνεται ότι το thread1 τρέχει συνεχώς και ενημερώνει ανά κάποια δευτερόλεπτα, ανάλογα με την χωρητικότητα των ουρών, το thread2 για να αρχίσει να εκτελείται. Σύμφωνα με τα παραπάνω φαίνεται ότι το σύστημα μας λαμβάνει αποφάσεις για το ποια κίνηση έχει εκτελεστεί, κάθε περίπου N-1 δευτερόλεπτα, όπου N είναι το μέγεθος της ουράς.



Εικόνα 5.8: Αρχιτεκτονική online συστήματος

5.3.1 Βασικός περιορισμός

Το κυρίαρχο πρόβλημα που παρουσιάζεται στην real time χρήση του συστήματος είναι ότι δεν μπορεί να προβλεφθεί εκ των προτέρων πότε θα γίνει μία κίνηση από τον χρήστη. Ο ταξινομητής επομένως δεν γνωρίζει ποια είναι η χρονική σχέση του σήματος που καταγράφεται με κάποια κίνηση, ή ακόμα αν υπάρχει κίνηση. Θεωρητικά οι ταξινομητές θα μπορούσαν να λειτουργούν όπως πριν με την ίδια ακρίβεια, αλλά στην πράξη επειδή τα SVM έχουν εκπαιδευτεί σε συγκεκριμένα παραδείγματα δεδομένων τα οποία είναι άμεσα χρονικά συνδεδεμένα με την κίνηση μπορεί αυτό να οδηγήσει σε αναξιόπιστα αποτελέσματα. Για αυτό τον λόγο αναφέρονται παρακάτω οι παράγοντες οι οποίοι δοκιμάστηκαν και μπορούν να βελτιώσουν όσο το δυνατό περισσότερο την ακρίβεια της υλοποίησης του συστήματος σε πραγματικό χρόνο.

➤ Χρήση 2^{ου} SVM

Είναι σημαντικό λοιπόν να μπορεί να γίνει κάποιου είδους πρόβλεψη για το αν υπάρχει κάποια κίνηση ή όχι στα δεδομένα εισόδου. Το τελευταίο είναι απαραίτητο επειδή μπορεί ο χρήστης στο συγκεκριμένο παράθυρο καταγραφής να μην έχει προλάβει ή να μην επιθυμεί να εκτελέσει κάποια κίνηση. Έτσι, η χρήση του δεύτερου ταξινομητή συμβάλει στην απαίτηση της πρόβλεψης καθώς κάνει αυτόν ακριβώς τον διαχωρισμό.

➤ Πειραματισμός με παράθυρα

Ένα βασικό κομμάτι της υλοποίησης που χρειάζεται προσοχή και επηρεάζει την ακρίβεια του συστήματος είναι η **επιλογή του χρόνου καταγραφής** των δεδομένων, δηλαδή στην πράξη η χωρητικότητα των ουρών. Στην περίπτωση της offline υλοποίησης ήταν σίγουρο ότι το κάθε παράθυρο που καταγράφεται περιέχει τα επιθυμητά δεδομένα για την κατηγοριοποίηση της εκάστοτε κίνησης, κάτι το οποίο δεν είναι σίγουρο στην online υλοποίηση. Αρχικά, για απλότητα σημειώνεται ότι οι ουρές των 4 καναλιών είναι ίδιας χωρητικότητας. Με την προϋπόθεση αυτή δοκιμάστηκαν τα παρακάτω μεγέθη παραθύρων, ανάλογα με τον χρόνο καταγραφής. Συγκεκριμένα δοκιμάσαμε για χρόνους καταγραφής 4, 3 και 2.5 δευτερόλεπτα, με τα αποτελέσματα να παρουσιάζονται στον ακόλουθο πίνακα.

Action	Time window (sec)		
	4	3	2.5
Open (no movement)	80%	80%	40%
Close	80%	80%	50%
Right	90%	80%	50%
Left	90%	80%	50%

Πίνακας 5: Αποτελέσματα πειραματισμού διαφορετικών παραθύρων

Από τον παραπάνω πίνακα βλέπουμε ότι η ακρίβεια του παραθύρου των 2.5 δευτερολέπτων είναι πολύ κακή. Ειδικά όσον αφορά τις καταστάσεις των open/ close ματιών, το σύστημα δεν είναι σε θέση να εκπαιδευτεί σωστά, διότι το διάστημα χαρακτηριστικών να είναι πολύ μικρού μεγέθους (160 samples). Παράλληλα, βλέπουμε ότι για την ίδιο μέγεθος παραθύρου η επίδοση των 2 άλλων SVM είναι πολύ χειρότερη, κάτι το οποίο μπορεί να δικαιολογηθεί λόγω του ότι το παράθυρο καταγραφής είναι σχετικά μικρό και οι πιθανότητες να κόβει την απότομη μεταβολή της οριζόντιας κίνησης στην μέση είναι πολύ μεγαλύτερη. Παρατηρείται λοιπόν ότι καλύτερα αποτελέσματα παρουσιάζει το παράθυρο μεγέθους 4 δευτερολέπτων, παρόλο που και το παράθυρο μήκους 3 δευτερολέπτων παρουσιάζει αρκετά ικανοποιητικά αποτελέσματα. Προκειμένου να επιλεγεί το βέλτιστο μέγεθος του παραθύρου καταγραφής, οφείλουμε να λάβουμε υπόψιν το tradeoff μεταξύ της απόδοσης και του χρόνου απόφασης του συστήματος,

το οποίο εξαρτάται άμεσα από την επιλογή του παραθύρου. Είναι προφανές ότι όσο μεγαλύτερο παράθυρο χρησιμοποιείται τόσο μεγαλύτερη είναι η πιθανότητα να εντοπιστεί μία κίνηση, εάν έχει εκτελεστεί, και να γίνει σωστά η ταξινόμηση. Παρόλα αυτά όσο μεγαλώνει το μέγεθος του παραθύρου τόσο μεγαλώνει και η χρονική απόσταση μεταξύ των αποφάσεων που λαμβάνει το σύστημα, αφού όπως προαναφέρθηκε ο χρόνος αυτός ισούται με $N-1$, όπου N : μέγεθος του παραθύρου. Στην συγκεκριμένη περίπτωση επιλέξαμε να χρησιμοποιήσουμε το παράθυρο των 4 δευτερολέπτων, με το οποίο συνεχίζονται οι πειραματικές δοκιμές του συστήματος παρακάτω, για να έχουμε πιο σίγουρα αποτελέσματα.

5.4 Πειραματικές δοκιμές

Με την προϋπόθεση ότι έχουν αποφασιστεί τα χαρακτηριστικά του συνολικού συστήματος, με τα οποία επιτυγχάνεται η καλύτερη δυνατή απόδοση του συστήματος, έγιναν δοκιμές σε συνολικά 5 άτομα (3 άντρες και 2 γυναίκες) έτσι ώστε να παρθεί μία πιο συνολική εικόνα της ακρίβειας του συστήματος σε δοκιμές πραγματικού χρόνου. Όλοι οι εθελοντές αφού ενημερώθηκαν για τον τρόπο διεξαγωγής του πειράματος, εκτέλεσαν 10 επαναλήψεις για κάθε κίνηση, δηλαδή για open/close και right/left. Κατά την διάρκεια κάθε πειράματος, καταγράφονταν τα ποσοστά επιτυχίας της κάθε κίνησης του κάθε εθελοντή, προκειμένου να ελέγξουμε την απόδοση του συστήματος σε on-λάιν εφαρμογή. Είναι φανερό ότι το ποσοστό επιτυχίας της κάθε κίνησης μετριόνταν κατηγοριοποιώντας μία κίνηση ως σωστή εάν το σύστημα την ταξινομούσε στην σωστή κατηγορία. Παρακάτω παρουσιάζεται ο πίνακας με τα ποσοστά επιτυχίας για κάθε κίνηση του εκάστοτε εθελοντή (Subject).

Action	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
Open (no movement)	90%	80%	70%	70%	90%
Close	70%	80%	60%	60%	100%
Right	70%	90%	80%	70%	80%
Left	80%	90%	80%	70%	80%

Πίνακας 6: Αποτελέσματα δοκιμών των 4 κινήσεων

Παρατηρείται ότι όλοι οι εθελοντές πέτυχαν καλή απόδοση για τις οριζόντιες κινήσεις ματιών, με πολλούς από αυτούς να πετυχαίνουν επιτυχία από 80% και πάνω. Η μεγαλύτερη αστοχία παρουσιάστηκε στις κινήσεις open (no movement) και close. Όσον αφορά την πρώτη το ποσοστό αποτυχίας θα μπορούσε να δικαιολογηθεί από το γεγονός ότι είναι αρκετά δύσκολο ο χρήστης να είναι ακίνητος (χωρίς να κινεί άλλα μέλη του σώματός του) και κυρίως να κρατήσει ακίνητο το βλέμμα του. Σχετικά με την κατηγορία close, διαφαίνεται ότι σε κάθε άνθρωπο δεν είναι το ίδιο εμφανή τα άλφα κύματα, πράγμα το οποίο είναι λογικό, και ως εκ τούτου έχουμε και μεγάλη διακύμανση στην ακρίβεια για αυτήν την κατηγορία. Παρόλα αυτά δεδομένων των περιορισμών του συστήματος και της συσκευής, τα αποτελέσματα κρίνονται ικανοποιητικά.

Στην συνέχεια, εκτός από την εξέταση των 4 κινήσεων δοκιμάστηκε και μία σειρά από συνδυασμούς, η οποία εκτελέστηκε από την εθελόντρια (Subject 5) που πέτυχε την καλύτερη ακρίβεια στο παραπάνω πίνακα.

Actions	Ποσοστό επιτυχίας
close, right	100%
open, left	100%
open, left, close	100%
right, open, left	66,67%
close, open, right	100%
right, left, close	66,67%
close, open, right, left	75%
left, right, left, right	100%

Πίνακας 7: Πίνακας συνδυαστικών αποτελεσμάτων

Παρατηρώντας τα αποτελέσματα στους παραπάνω πίνακες (Πίνακας 6 και Πίνακας 7) βλέπουμε ότι η απόδοση των συνδυαστικών κινήσεων κυμαίνεται από 66,67% έως 100%, ανάλογα με τον αριθμό των κινήσεων και το είδος τους. Έτσι φαίνεται ότι το σύστημα που υλοποιήθηκε επιτυγχάνει αρκετά καλή απόδοση πράγμα το οποίο το καθιστά εφαρμόσιμο και στην πράξη για πραγματική οδήγηση ενός demo car. Την υλοποίηση αυτή θα την μελετήσουμε στο επόμενο κεφάλαιο.

6 Κεφάλαιο: Υλοποίηση σε RC car

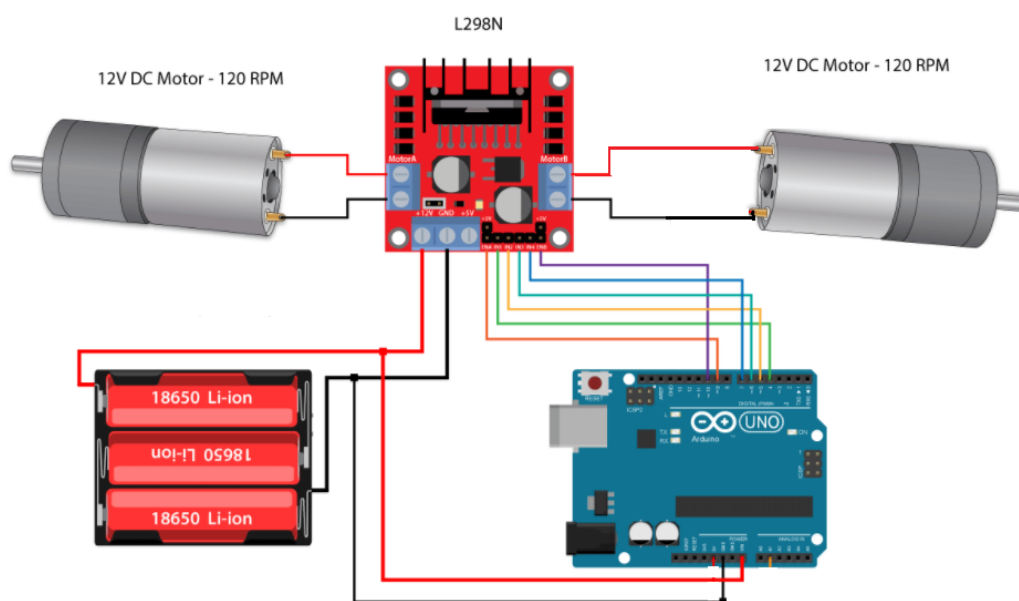
6.1 Σχεδιασμός demo car

Όπως έχει αναφερθεί και προηγουμένως σκοπός μας είναι η αναγνώριση και ο διαχωρισμός των 4 κινήσεων, έτσι ώστε να μπορεί ο χρήστης δυναμικά να ελέγχει με την κίνηση των ματιών του ένα demo car. Για αυτό τον λόγο κατασκευάστηκε ένα μικρό demo car, το οποίο συνδέσαμε κατάλληλα με το υπόλοιπο σύστημά μας. Πιο αναλυτικά για την υλοποίηση του συνολικού συστήματος πλοήγησης του μικρού αμαξιδίου (demo car) χρειάστηκαν :

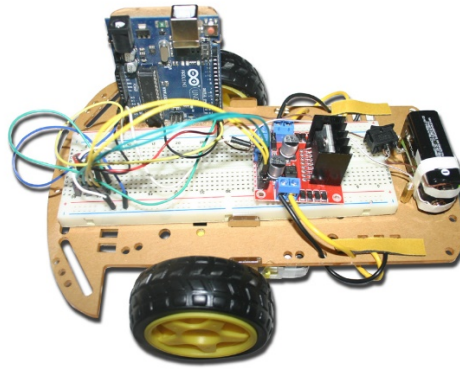
- Μία πλακέτα Arduino UNO
- Ένας Driver L298N 2A
- Καλώδια (Jumpers)
- Demo car
- Μπαταρίες
- Breadboard

Τα παραπάνω εξαρτήματα συνδέθηκαν όπως φαίνεται στην παρακάτω Εικόνα 6.1. Σημειώνεται ότι οι δύο DC κινητήρες ενσωματώθηκαν πάνω στο demo car που κατασκευάστηκε το οποίο φαίνεται στην Εικόνα 6.2 μαζί με όλα τα επιμέρους εξαρτήματα.

Εικόνα 6.1: Σύνδεση εξαρτημάτων



Εικόνα 6.1: Σύνδεση εξαρτημάτων



Εικόνα 6.2: Demo RC car

6.1.1 Περιγραφή λειτουργίας

Η λειτουργία του συστήματος είναι σχετικά απλή. Όπως φαίνεται το L298N είναι υπεύθυνο στο να δίνει την απαραίτητη ισχύ στους κινητήρες, μέσω της τροφοδότησής του. Το συγκεκριμένο εξάρτημα είναι απαραίτητο καθώς δεν είναι δυνατό να συνδεθούν κατευθείαν οι κινητήρες του αμαξιού με το Arduino, διότι δεν μπορεί να παράξει την απαραίτητη τάση για την οδήγησή τους. Στην συνέχεια, ο driver συνδέεται με το Arduino, το οποίο μεταφέρει τις εντολές που λαμβάνει από την έξοδο USB του υπολογιστή μέσω κατάλληλου περιβάλλοντος. Όπως φαίνεται από το παραπάνω σχήμα κάθε κινητήρας κατευθύνεται μέσω 2 pins του Arduino. Έτσι προκειμένου να επιτευχθεί η επιθυμητή κίνηση κάθε φορά πρέπει να δοθεί ο κατάλληλος συνδυασμός σε αυτά τα pins.

Action	Motor 1 (in1, in2)	Motor 2 (in3, in4)
forward	(HIGH, LOW)	(HIGH, LOW)
stop	(LOW, LOW)	(LOW, LOW)
right	(HIGH, LOW)	(LOW, HIGH)
left	(LOW, HIGH)	(HIGH, LOW)

Εικόνα 6.3: Τιμές των pins για την κίνηση του αμαξιού

Δίπλα από το όνομα του κάθε κινητήρα σημειώνονται τα αντίστοιχα pins με τα οποία ελέγχεται. Παρατηρείται επίσης ότι οι στροφές δεξιά και αριστερά, επιτυγχάνονται με αντίθετη κίνηση των τροχών.

6.2 Αντιστοιχία κινήσεων ματιών - demo car

Η αντιστοίχιση των κινήσεων για την οδήγηση του αμαξιδίου επιτεύχθηκε στην πράξη μέσω της σύνδεσης του προγράμματος που δημιουργήθηκε στο περιβάλλον του Arduino (σε C) και του προγράμματος της Python, το οποίο υλοποιεί το σύστημα απόφασης που περιεγράφηκε στο Κεφάλαιο 5. Η αντιστοιχία των κινήσεων φαίνεται στον παρακάτω πίνακα.

Κίνηση ματιών	Κίνηση demo car
close	εκκίνηση/ σταμάτημα
open	συνέχιση πορείας
right	δεξιά στροφή
left	αριστερή στροφή

Πίνακας 8: Αντιστοιχία κινήσεων ματιών-demo car

Αναλυτικότερα, η σύνδεση των δύο προγραμμάτων έγινε μέσω της σειριακής πόρτας COM3 του υπολογιστή, με την οποία επικοινωνεί το Arduino. Πάνω σε αυτή συνδέσαμε το υπάρχων σύστημα απόφασης που είχαμε υλοποιήσει, και με αυτό τον τρόπο έγινε εφικτή η ανταλλαγή δεδομένων. Πιο συγκεκριμένα, κάθε φορά που το σύστημα λαμβάνει μία απόφαση για την κλάση των δεδομένων εισόδου (Open/Close, Right/Left), στέλνει κατάλληλο μήνυμα στο Arduino μέσω της συγκεκριμένης πόρτας, το οποίο μεταφράζεται σε αντίστοιχη κίνηση των τροχών.

6.3 Πειραματικές δοκιμές

Έχοντας υλοποιήσει την σύνδεση του συστήματος απόφασης με το σύστημα που είναι υπεύθυνο για την αντίστοιχη κίνηση του demo car, επόμενο στάδιο είναι η δοκιμή του σε εισερχόμενα δεδομένα εισόδου, με το σύστημα να λαμβάνει αποφάσεις για το ποια κίνηση έχει εκτελεί σε πραγματικό χρόνο. Συγκεκριμένα για τον σκοπό αυτό χρησιμοποιήθηκαν έτοιμα καταγεγραμμένα δεδομένα κινήσεων, τα οποία έχουν παρθεί από το EPOC. Δοκιμάστηκαν 6 κινήσεις εν σειρά: Close, Open, Right, Open, Left, Close. Το αμαξάκι ξεκινά την πορεία του αφού πάρει την πρώτη εντολή Close, συνεχίζει ευθεία (Open), στην συνέχεια στρίβει δεξιά (Right), έπειτα συνεχίζει πάλι ευθεία (Open), στρίβει αριστερά (Left) και τέλος σταματά με την τελευταία εντολή Close. Παρατηρείται ότι το demo car ανταποκρίνεται άμεσα στις αλλαγές της κατεύθυνσης, κάθε φορά που λαμβάνεται η απόφαση από το σύστημά μας για το ποια κίνηση έχει εκτελέσει ο χρήστης. Στην συγκεκριμένη δοκιμή επειδή τα δεδομένα μας είναι ήδη καταγεγραμμένα και διαβάζονται σειριακά χωρίς να υπάρχει η καθυστέρηση της μετάδοσης του on-λάνιν συστήματος, έχουμε εισάγει εμείς μία καθυστέρηση περίπου 2 δευτερολέπτων, έτσι ώστε να προσομοιαστεί η εφαρμογή πραγματικού χρόνου.

7 Κεφάλαιο : Συμπεράσματα και μελλοντική εργασία

7.1 Σύνοψη

Στην παρούσα διπλωματική εργασία μελετήθηκαν τα EEG και EOG σήματα του ανθρωπίνου εγκεφάλου με στόχο την δημιουργία μίας διεπαφής εγκεφάλου – υπολογιστή (βλ. Κεφάλαιο 1). Συγκεκριμένα αφού έγινε πρώτα μία ανασκόπηση των βιοσημάτων που χρησιμοποιούνται σε τέτοιες εφαρμογές (βλ. Κεφάλαιο 2), επικεντρωθήκαμε στα σήματα που σχετίζονται με τις κινήσεις των ματιών. Η μελέτη του θεωρητικού υποβάθρου για την δημιουργία του κάθε σταδίου της εφαρμογής BCI και οι τεχνικές που χρησιμοποιήθηκαν παρατίθενται στο Κεφάλαιο 3. Στην συνέχεια, στο Κεφάλαιο 4 παρουσιάζονται τα πρώτα βήματα για την δημιουργία της προτεινόμενης BCI εφαρμογής, τα οποία είναι: η καταγραφή, η προεπεξεργασία και η εξαγωγή των κατάλληλων χαρακτηριστικών από τα σήματα, στάδια ιδιαίτερα σημαντικά για την μετέπειτα ταξινόμηση των σημάτων. Ειδικότερα, ασχοληθήκαμε με 4 κινήσεις των ματιών ανοιχτά/ κλειστά μάτια και οριζόντια κίνηση δεξιά και αριστερή. Μέσα από αυτές τις μεθόδους επεξεργασίας των σημάτων εξήχθησαν σημαντικά αποτελέσματα και συμπεράσματα, τα οποία βοήθησαν στον διαχωρισμό τους. Συγκεκριμένα, για τα κανάλια O1, O2 τα οποία συσχετίζονται με το άνοιγμα/ κλείσιμο των ματιών, παρατηρήσαμε ότι υπάρχει μία μεγάλη διαφορά στην τιμή της ισχύος στο πεδίο συχνοτήτων όταν ο χρήστης έχει κλειστά τα μάτια τους, σε σχέση με όταν τα έχει ανοιχτά. Αυτή η εμφανής διαφορά των δύο καταστάσεων αποτέλεσε τον τρόπο με τον οποίο τα διαχωρίσαμε στην συνέχεια. Παράλληλα, όσον αφορά τα κανάλια F7, F8 τα οποία επιλέχθηκαν για την οριζόντια κίνηση των ματιών μετά από ποικίλες δοκιμές προεπεξεργασίας παρατηρήσαμε ότι η οριζόντια κίνηση των ματιών επιφέρει μία μεγάλη μεταβολή αντίθετου πρόσημου στα δύο κανάλια. Βασιζόμενοι στην παραπάνω παρατήρηση, αφαιρέσαμε τα δύο κανάλια με σκοπό να γίνει ακόμα εντονότερη η μεταβολή των τιμών, όταν ο χρήστης κινεί δεξιά ή αριστερά το βλέμμα του. Έχοντας επεξεργαστεί κατάλληλα τα σήματα μας, επόμενο στάδιο είναι η ταξινόμηση τους. Για τον λόγο αυτό εξετάστηκαν 3 τρόποι ταξινόμησης και εν τέλει στο Κεφάλαιο 5, παρουσιάζεται το συνολικό σύστημα με βάση τα SVM, τα οποία κρίθηκαν ως ο καλύτερος τρόπος διαχωρισμού των κινήσεων. Σε αυτό το σημείο είναι σημαντικό να σημειωθεί ότι το σύστημα μας εκπαιδεύτηκε βάση ενός συνόλου δεδομένων το οποίο δημιουργήθηκε μέσω συγκεκριμένου πρωτοκόλλου. Τα συμπεράσματα του συστήματος ταξινόμησης ήταν ιδιαιτέρως ικανοποιητικά, αφού συγκεκριμένα για το 1^ο SVM πετύχαμε ακρίβεια 94%, για το 2^ο SVM 99% και για το 3^ο SVM η ακρίβεια είναι 100%. Παράλληλα, τροποποιήσαμε το προαναφερθέν σύστημα έτσι ώστε να ανταποκρίνεται και σε μετρήσεις πραγματικού χρόνου, με τα αποτελέσματα των πειραματικών δοκιμών να κυμαίνονται σε απόδοση από το 70 - 100%, ανάλογα με την κίνηση. Καταλήγοντας, αφού επιτεύχθηκε η ταξινόμηση των σημάτων, τελευταίο κομμάτι της εργασίας αποτελεί η σύνδεση του προαναφερόμενου συστήματος με το demo car που κατασκευάστηκε, πράγμα που περιγράφεται στο Κεφάλαιο 6.

Συμπεραίνουμε λοιπόν ότι η διεπαφή εγκεφάλου - υπολογιστή που δημιουργήσαμε για τον απλό έλεγχο ενός demo car πετυχαίνει αρκετά καλή ακρίβεια. Είναι σημαντικό να σημειωθεί ότι η ακρίβεια των ταξινομητών που εκπαιδεύσαμε είναι πολύ ικανοποιητική και σε σχέση με πολλές μελέτες της βιβλιογραφίας καλύτερη για τον διαχωρισμό των συγκεκριμένων κινήσεων.

7.2 Μελλοντική εργασία

Είναι γεγονός ότι η ανάπτυξη εφαρμογών που βασίζονται στο Brain Computer Interface, είναι ένα πολύ νέο ερευνητικό πεδίο που αναπτύσσεται κυρίως τα τελευταία 20 χρόνια. Είναι φανερό επομένως ότι υπάρχει ακόμα μεγάλο πεδίο βελτίωσης τόσο όσο αφορά την εφαρμοσιμότητα και αποδοτικότητα των υπαρχουσών μεθόδων όσο και στην εύρεση νέων μεθόδων και εφαρμογών. Έτσι και στην προκειμένη περίπτωση παρότι το συνολικό σύστημα της εφαρμογής που προτάθηκε μπορεί να επιτύχει καλά αποτελέσματα υπό τους όρους και τους περιορισμούς που έχουμε αναφέρει, μπορεί να γίνει περαιτέρω έρευνα προκειμένου να βελτιώσει ακόμα περισσότερο την αποδοτικότητα και εφαρμοσιμότητα του στην πράξη. Η έρευνα αυτή θα μπορούσε να επικεντρωθεί στα παρακάτω ζητήματα:

1. Έρευνα περισσότερων κινήσεων

Όπως περιεγράφηκε στην παρούσα διπλωματική μελετήθηκαν μόνο 4 κινήσεις των ματιών. Παρόλα αυτά κάτι τέτοιο είναι περιοριστικό διότι οι κινήσεις του ματιού είναι πολύ περισσότερες για παράδειγμα κίνηση πάνω – δεξιά. Για αυτό το λόγο είναι χρήσιμο να μελετηθούν και άλλες κινήσεις για να υπάρχει μεγαλύτερη ποικιλία στο σύστημά μας. Μία άλλη σκέψη είναι να ερευνηθεί η συνεχής αποκωδικοποίηση της κίνησης του ματιού, και όχι μεμονωμένων κινήσεων του.

2. Καταγραφή δεδομένων

Η καταγραφή δεδομένων στην εργασία αυτή έγινε με την συσκευή EPOC της εταιρείας EMOTIV. Αυτή είναι μία εμπορική συσκευή κάτι το οποίο συνεπάγεται ότι η ποιότητα καταγραφής των σημάτων δεν είναι τόσο καλή όσο άλλων συσκευών αποκλειστικής εργαστηριακής χρήσης. Παρόλο που τα αποτελέσματα που πάρθηκαν ήταν ικανοποιητικά για το συγκεκριμένο πρόβλημα, πρέπει να σημειωθεί ότι λόγω των σφαλμάτων της συσκευής αρκετά δεδομένα που καταγράφηκαν δεν συμπεριλήφθηκαν καθόλου. Το τελευταίο επηρεάζει άμεσα την απόδοση ιδίως της απόκρισης του συστήματος σε καταγραφές πραγματικού χρόνου. Είναι χρήσιμο επομένως να μελετηθούν και άλλες φορητές συσκευές καταγραφής ηλεκτροεγκεφαλογραφημάτων.

3. Online εφαρμογή

Στο κομμάτι που υλοποιήσαμε την εφαρμογή έτσι ώστε να ανταποκρίνεται σε πραγματικό χρόνο, σημειώσαμε τον βασικό περιορισμό του συστήματος ο οποίος είναι η πρόβλεψη του αν υπάρχει κάποια κίνηση στα εισερχόμενα δεδομένα ή όχι. Τον περιορισμό αυτό προσπαθήσαμε να τον αντιμετωπίσουμε με την χρήση 2^{ου} SVM και την επιλογή του βέλτιστου παραθύρου καταγραφής. Στην συγκεκριμένη εφαρμογή επιλέξαμε ένα λίγο μεγαλύτερο παράθυρο καταγραφής, κάτι το οποίο συνεπάγεται αύξηση της καθυστέρησης του συστήματος, προκειμένου να έχουμε καλύτερη ακρίβεια στην ταξινόμηση των κινήσεων. Αυτή δεν είναι προφανώς μία βέλτιστη λύση του

προβλήματος, για αυτό τον λόγο είναι σημαντικό να μελετηθούν και άλλοι τρόποι πρόβλεψης ενός συμβάντος, για παράδειγμα με την μεταβολή της διασποράς των τιμών, έτσι ώστε να αυξηθεί ο χρόνος ανταπόκρισης της εφαρμογής.

4. Βελτίωση συστήματος απόφασης

Όπως βλέπουμε στην παρούσα διπλωματική το σύστημα απόφασης αποτελεί τον πυρήνα της εφαρμογής BCI που προτείνεται. Παρόλο που η ακρίβεια του συστήματος είναι ικανοποιητική για το συγκεκριμένο πρόβλημα, ειδικά για το πρώτο SVM υπάρχουν περιθώρια βελτίωσης της ακρίβειας αλλά και της απόκλισης της. Το τελευταίο θα μπορούσε να βελτιωθεί με την χρήση διαφορετικών ή περισσότερων δεδομένων εκπαίδευσης. Κάτι ακόμα που θα μπορούσε να διερευνηθεί είναι η επιλογή διαφορετικών χαρακτηριστικών για την εκπαίδευση των SVM.

5. Βελτίωση κινηματικής demo car

Τελευταίο κομμάτι της εφαρμογής ήταν η σύνδεση ενός demo car με το προαναφερθέν σύστημα απόφασης, και στην συνέχεια η αντιστοίχιση των 4 κινήσεων των ματιών με 4 κινήσεις του αμαξιδίου. Όπως περιεγράφηκε η αντιστοίχιση αυτή είναι αρκετά θεμελιώδης, πράγμα το οποίο δεν προσομοιώνει την πραγματική κινηματική ενός αμαξιδίου. Για να μπορέσω να οδηγηθεί επαρκώς το αμαξάκι πρέπει να μελετηθεί πιο λεπτομερώς το κινηματικό του μοντέλο, έτσι ώστε να έχουμε καλύτερη ακρίβεια στην κίνησή του. Παραδείγματος χάρη οι στροφές του αμαξιδίου είναι σημαντικό να καλιμπραριστούν κατάλληλα προκειμένου το αμαξάκι να στρίβει κάθε φορά όσο επιθυμεί ο χρήστης. Βελτιώνοντας το κινηματικό μοντέλο, η εφαρμογή αυτή μπορεί μελλοντικά να δοκιμαστεί και σε αναπηρικό αμαξίδιο, κάτι ιδιαίτερα ενδιαφέρον.

Βιβλιογραφία

- Bekdash, M. A. (2015, May). Visual evoked potentials response to different colors and intensities. *In BioSignal Analysis, Processing and Systems (ICBAPS), 2015 International Conference on* (pp. 104-107). IEEE.
- Belkacem, A. N. (2014). Classification of four eye directions from EEG signals for eye-movement-based communication systems. *life*, 1, 3.
- Belkacem, A. N.-a. (2015). Real-time control of a video game using eye movements and two temporal EEG sensors. *Computational intelligence and neuroscience*, 2015, 1.
- Bell, C. J. (2008). Control of a humanoid robot by a noninvasive brain–computer interface in humans. *. Journal of neural engineering*, 5(2), 214.
- Bonnet, L. L. (2013). Two brains, one game: design and evaluation of a multiuser BCI video game based on motor imagery. *. IEEE Transactions on Computational Intelligence and AI in games*, 5(2), 185-198.
- Boser, B. E. (1992, July). A training algorithm for optimal margin classifiers. *. In Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152). ACM.
- Bulling, A. R. (2009). Wearable EOG goggles: eye-based interaction in everyday environments (pp. 3259-3264). ACM.
- by Alan V. Oppenheim (Author), . R. (χ.χ.). *Discrete-Time Signal Processing (3rd Edition) (Prentice-Hall Signal Processing Series)*.
- Chen, Y. J. (2016). Fuzzy tracking and control algorithm for an SSVEP-based BCI system. *Applied Sciences*, 6(10), 270.
- Choi, B. &. (2013). A low-cost EEG system-based hybrid brain-computer interface for humanoid robot navigation and recognition. *PloS one*, 8(9), e74583.
- Choi, I. R. (2017). A systematic review of hybrid brain-computer interfaces: Taxonomy and usability perspectives. *. PloS one*, 12(4), e0176674.
- Elstob, D. &. (2016). A low cost eeg based BCI prosthetic using motor imagery. *arXiv preprint arXiv:1603.02869*.
- Farwell, L. A. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6), 510-523.
- Gao, X. Y. (2015, April). Evaluating driving fatigue detection algorithms using eye tracking glasses. *. In Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on* (pp. 767-770). IEEE.
- Guger, C. D. (2009). How many people are able to control a P300-based brain–computer interface (BCI)? *. Neuroscience letters*, 462(1), 94-98.

- Güneysu, A. &. (2013). An SSVEP based BCI to control a humanoid robot by using portable EEG device. *In Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE (pp. 6905-6908). IEEE.*
- Haykin, S. S. (2009). *Neural networks and learning machines (Vol. 3)*. S. S., & Haykin Upper Saddle River, NJ, USA:: Pearson.
- Hoffmann, U. V. (2008). An efficient P300-based brain–computer interface for disabled subjects. *Journal of Neuroscience methods, 167(1), 115-125.*
- <http://www.bem.fi/book/28/28.htm>.
- <http://www.bem.fi/book/28/28.html>.
- <http://www.instructables.com/id/L298n-Arduino-Library/>.
- <https://howtomechatronics.com/tutorials/arduino/arduino-and-hc-05-bluetooth-module-tutorial/>.
- <https://www.arduino.cc/>.
- https://www.medicine.mcgill.ca/physio/vlab/Other_exps/EOG/eogintro_n.htm.
- Huo, X. Q. (2016, July). Driving fatigue detection with fusion of EEG and forehead EOG. . *In Neural Networks (IJCNN), 2016 International Joint Conference on (pp. 897-904). IEEE.*
- Hwang, H. J. (2012). Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. . *Journal of neuroscience methods, 208(1), 59-65.*
- Iturrate, I. A. (2009). A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation. *IEEE Transactions on Robotics, 25(3), 614-627.*
- Kartsch, V. B. (2017, May). A wearable EEG-based drowsiness detection system with blink duration and alpha waves analysis. . *In Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on (pp. 251-254). IEEE.*
- Kasim, M. A. (2017). User-Friendly LabVIEW GUI for Prosthetic Hand Control Using Emotiv EEG Headset. *Procedia Computer Science, 105, 276-281.*
- Kim, B. H. (2014). Computers in biology and medicine, 51, 82-92. *Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking. .*
- LaFleur, K. C. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface. *Journal of neural engineering, 10(4), 046003.*
- Li, M. Z. (2013). An EEG based control system for intelligent wheelchair. . *In Applied Mechanics and Materials (Vol. 300, pp. 1540-1545). Trans Tech Publications.*
- Li, W. J. (2012, January). Development of mind control system for humanoid robot through a brain computer interface. . *In Intelligent System Design and Engineering Application (ISDEA), 2012 Second International Conference on (pp. 679-6).*
- Lin, J. S. (2015). Implementing remote presence using quadcopter control by a non-invasive BCI device. *Computer Science and Information Technology, 3(4), 122-126.*

- López-Gil, J. M.-G.-G. (2016). Method for improving EEG based emotion recognition by combining it with synchronized biometric and eye tracking technologies in a non-invasive .
- Mao, X. L. (2017). Progress in EEG-based brain robot interaction systems. . *Computational intelligence and neuroscience*.
- Martínez-Cagigal, V. G.-P. (2017). An asynchronous P300-based brain-computer interface web browser for severely disabled people. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(8), 1332-1342.
- Meriño, L. N. (2017). Asynchronous control of unmanned aerial vehicles using a steady-state visual evoked potential-based brain computer interface. . *Brain-Computer Interfaces*, 4(1-2), 122-135.
- Miralles, F. V.-P. (2015). Brain computer interface on track to home. . *The Scientific World Journal*.
- Mondada, L. K. (2016). Electroencephalography as implicit communication channel for proximal interaction between humans and robot swarms. *Swarm Intelligence*, 10(4), 247-265.
- Mouli, S. P. (2013, November). Performance analysis of multi-frequency SSVEP-BCI using clear and frosted colour LED stimuli. . In *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference* .
- Nijholt, A. B. (2009). Turning shortcomings into challenges: Brain–computer interfaces for games. . *Entertainment computing*, 1(2), 85-94.
- Ramli, R. A. (2015). Classification of eyelid position and eyeball movement using EEG signals. . *Malaysian Journal of Computer Science*, 28(1), 28-45.
- Rösler, O. &. (2013). A first step towards eye state prediction using eeg. Proc. of the AIHLS.
- Samadi, M. R. (2014, September). EEG signal processing for eye tracking. . In *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European (pp. 2030-2034)*. IEEE.
- Signals and Systems (2nd Edition)* . (χ.χ.). Alan V. Oppenheim (Author), Alan S. Willsky (Author), with S. Hamid (Author).
- Yu, Y. H. (2012, September). FlyingBuddy2: a brain-controlled assistant for the handicapped. In *UbiComp (pp. 669-670)*.
- Zhao, J. M. (2014, June). SSVEP-based hierarchical architecture for control of a humanoid robot with mind. . In *Intelligent Control and Automation (WCICA), 2014 11th World Congress on (pp. 2401-2406)*. IEEE.
- Χρήστος, Μ. (2017). Ηλεκτροεγκεφαλογραφικά Και Ηλεκτρομυογραφικά Σήματα Στα Νευρορομποτικά Συστήματα.