



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Χρονική και συχνοτική ανάλυση ακουστικών σημάτων ροχαλητού

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αναστάσης Π. Κελίρης

Επιβλέπουσα: Κωνσταντίνα Σ. Νικήτα

Καθηγήτρια Ε.Μ.Π.

Αθήνα, Ιούνιος 2012



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Χρονική και συχνοτική ανάλυση ακουστικών σημάτων ροχαλητού

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αναστάσης Π. Κελίρης

Επιβλέπουσα: Κωνσταντίνα Σ. Νικήτα

Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την Ιουνίου 2012.

.....
Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια Ε.Μ.Π.

.....
Δ. Δ. Κουτσούρης
Καθηγητής Ε.Μ.Π.

.....
Γ. Ματσόπουλος
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2012

.....
Αναστάσης Π. Κελίρης

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Αναστάσης Π. Κελίρης
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, 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 απευθύνονται προς τον συγγραφέα.

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

Περίληψη

Το ροχαλητό είναι το κύριο σύμπτωμα του συνδρόμου της αποφρακτικής άπνοιας - υπόπνοιας (Obstructive sleep apnea hypopnea syndrome - OSAHS). Η μέθοδος που χρησιμοποιείται αποκλειστικά μέχρι σήμερα, για διάγνωση του OSAHS, είναι το πολυσωματογράφημα (*polysomnogram* - PSG) το οποίο είναι χρονοβόρο και οικονομικά ασύμφορο. Σύμφωνα με μελέτες που έχουν γίνει, υπάρχουν ενδείξεις ότι τα ηχητικά σήματα ροχαλητού περιέχουν αρκετή πληροφορία για τη διάγνωση του συνδρόμου.

Σε αυτή την εργασία αναπτύχθηκε στο MATLAB® το πακέτο γραφικού διαδραστικού περιβάλλοντος Snor-Lux, το οποίο περιλαμβάνει αλγόριθμους για αποθρομβοποίηση σημάτων ροχαλητού, κατάτμησή τους σε επιμέρους ροχαλητά και ακολούθως επεξεργασία και εξαγωγή διαφόρων χαρακτηριστικών των σημάτων τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας. Σκοπός του πακέτου είναι η χρήση του στην διπλωματική εργασία καθώς και στην υποβοήθηση της αυτόματης διάγνωσης του OSAHS. Στη συνέχεια χρησιμοποιήθηκαν οι αλγόριθμοι σε δύο επιμέρους μελέτες σε σήματα ροχαλητού που ηχογραφήθηκαν στο Κέντρο Μελέτης Ύπνου του νοσοκομείου Ευαγγελισμός. Η πρώτη μελέτη διεξάγεται από το Κέντρο Μελέτης Ύπνου και αφορά την επίδραση του συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών σε ασθενείς με OSAHS. Έγινε αποθρομβοποίηση των σημάτων, κατάτμησή τους και εξαγωγή της μέγιστης και μέσης έντασης καθώς και της συχνότητας ροχαλητών ως κριτήρια αξιολόγησης της μελέτης. Η δεύτερη μελέτη αφορά την ικανότητα διάγνωσης του OSAHS μέσω χρονοσυχνοτικών χαρακτηριστικών των σημάτων ροχαλητού. Συγκεκριμένα μελετήθηκε η διαγνωστική ικανότητα των συχνοτήτων φωνοσυντονισμού μεμονωμένων ροχαλητών (formant frequencies $F1$, $F2$, $F3$) και τριών ενδεικτικών συχνοτήτων του φάσματος (μέση συχνότητα $f_{50\%}$, μέγιστη συχνότητα: $f_{90\%}$, και συχνότητα κορυφής f_{peak}), με τη βοήθεια καμπυλών ROC. Η ομαδοποίηση των εξεταζόμενων σε υγιείς - απνοϊκούς έγινε χρησιμοποιώντας δύο διαφορετικές τιμές κατωφλίου AHI ($AHI \geq 10$ ε/ώ και $AHI \geq 15$ ε/ώ).

Τα αποτελέσματα έδειξαν ότι για την κατηγοριοποίηση $AHI \geq 10$ ε/ώ οι συχνότητες $F1$, $f_{50\%}$ και $f_{90\%}$ είναι ικανές να διαγνώσουν το OSAHS. Συγκεκριμένα στην ανάλυση ROC για την $F1$ το κατώφλι διαχωρισμού είναι $F1 = 470$ Hz με ευαισθησία 97.7% και ειδικότητα 99.5%, για την $f_{50\%}$ το κατώφλι διαχωρισμού είναι $f_{50\%} = 113$ Hz με ευαισθησία 96% και ειδικότητα 98% και για την $f_{90\%}$ το κατώφλι διαχωρισμού είναι $f_{90\%} = 344$ Hz με ευαισθησία 94.5% και ειδικότητα 97%. Οι άλλες τρεις αναλύσεις με κριτήριο διαχωρισμού $AHI \geq 10$ ε/ώ, καθώς και οι έξι αναλύσεις με κριτήριο διαχωρισμού $AHI \geq 15$ ε/ώ κρίθηκαν αναξιόπιστες και δεν καταλήγουν σε αξιοποιήσιμο συμπέρασμα.

Μελετώντας τα αποτελέσματα καταλήγουμε στο συμπέρασμα ότι η διάγνωση του OSAHS μέσω χρονοσυχνοτικών χαρακτηριστικών των ακουστικών σημάτων ροχαλητού είναι εφικτή, απαιτεί όμως περαιτέρω μελέτη και ανάλυση μεγαλύτερου όγκου δεδομένων με περισσότερους υγιείς εξεταζόμενους ώστε να γίνει μπορεί να γίνει η σύγκριση των χαρακτηριστικών υγιών και απνοϊκών εξεταζόμενων με περισσότερη αξιοπιστία.

Λέξεις Κλειδιά

ακουστικά σήματα ροχαλητού, σύνδρομο αποφρακτικής άπνοιας - υπόπνοιας, συχνότητες φωνοσυντονισμού, καμπύλες ROC, λογισμικό επεξεργασίας σημάτων ροχαλητού, μέση - μέγιστη συχνότητα και συχνότητα κορυφής, Matlab GUI

Abstract

Snoring is the primary symptom of obstructive sleep apnea hypopnea syndrome (OSAHS). The gold standard used so far for diagnosing OSAHS is an overnight polysomnogram, which is time consuming and economically expensive. According to studies, there is evidence that snoring sound signals contain enough information and are able to diagnose the syndrome. In this thesis a Graphical User Interface package was developed in MATLAB®, called Snor-Lux. Snor-Lux includes algorithms for denoising snoring sound signals, partitioning large signals into individual snores, processing them and extracting information concerning both the time and frequency domain. The aim of this package is to be used for the needs of the thesis and to assist the automatic diagnosis of OSAHS.

The algorithms developed were then used in two separate clinic trials, with data from snoring sound signals recorded at the Sleep Center of "Evangelismos" hospital. The first study is conducted by the Sleep Center and examines the effect of the combination of nasal decongestants and corticosteroids in patients with OSAHS. The signals were initially denoised and segmented. The maximum and average intensity was extracted as well as the snoring frequency as evaluation criteria for the study. The second study examines the diagnostic ability of six time-frequency domain features: the formant frequencies (F1, F2, F3) and three indicative frequencies of the snoring signals' spectrum (mean frequency $f_{50\%}$, maximal frequency $f_{90\%}$ and peak frequency f_{peak}), using ROC curve analysis. The clustering of the examinees into a healthy and an apneic group was performed using two different thresholds for AHI (AHI ≥ 10 e/h and AHI ≥ 15 e/h).

The results for the first clustering (AHI ≥ 10 e/h) were encouraging as frequencies F1, $f_{50\%}$ and $f_{90\%}$ were found able to diagnose OSAHS. Specifically, the results of the ROC curve analysis for these frequencies are: F1 - Threshold 470 Hz, sensitivity 97.7%, specificity 99.5%, $f_{50\%}$ - Threshold 113 Hz, sensitivity 96%, specificity 98%, $f_{90\%}$ - Threshold 344 Hz, sensitivity 94.5%, specificity 97%. The ROC curve analysis of the rest three frequencies in the clustering with AHI ≥ 10 e/h as well as all six frequencies in the clustering with AHI ≥ 15 e/h were inconclusive and thus we cannot extract useful information for their diagnostic ability of OSAHS.

The conclusion is that the diagnosis of OSAHS using time-frequency features of snoring sound signals is feasible but requires further study. A larger dataset is required, containing more examinees, especially healthy, so that the comparison of healthy and apneic examinees is more reliable.

Keywords

snoring sound signals, obstructive sleep apnea hypopnea syndrome, formant frequencies, ROC curve analysis, snoring signals processing software, average - maximal - peak frequency, Matlab GUI

Ευχαριστίες

Στα πλαίσια της διπλωματικής μου εργασίας θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες σε όλους όσους βοήθησαν στην περάτωση της.

Αρχικά ευχαριστώ θερμά την επιβλέπουσα Δρ. Κωνσταντίνα Νικήτα που μου εμπιστεύτηκε την ανάθεση αυτής της εργασίας και μου παρείχε τις πολύτιμες συμβουλές της καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Ειλικρινείς ευχαριστίες απευθύνω και στην Δρ. Σπυρέττα Γολεμάτη για την καθοδήγηση και τη βοήθεια που μου προσέφερε, ειδικά στο ερευνητικό κομμάτι. Ευχαριστώ επίσης το ιατρικό προσωπικό του Κέντρου Μελέτης Ύπνου για τις συμβουλές τους στα ιατρικά θέματα και την άψογη συνεργασία τους.

Ένα ευχαριστώ οφείλω σε όλους τους καθηγητές μου κατά τη διάρκεια των σπουδών μου για τις αστείρευτες γνώσεις που μου μετάδωσαν όλα αυτά τα χρόνια, το καλύτερο εφόδιο στις δύσκολες εποχές που ζούμε.

Επιπλέον θέλω να εκφράσω την ευγνωμοσύνη μου προς τους γονείς μου για την διαρκή τους υποστήριξη, τόσο ηθική όσο και οικονομική, που μου επέτρεψε να διεκπεραιώσω τις σπουδές μου σε ένα άνετο περιβάλλον και την αδερφή μου για την υπομονή που έδειξε.

Τέλος ευχαριστώ τους φίλους μου για τη συμπαράσταση τους στις δύσκολες στιγμές και τη βοήθεια τους.

"Once we accept our limits, we go beyond them."

- Albert Einstein

Περιεχόμενα

Περίληψη.....	i
Λέξεις Κλειδιά.....	i
Abstract	ii
Keywords	ii
Ευχαριστίες.....	iii
Περιεχόμενα.....	v
Ευρετήριο Πινάκων	viii
Ευρετήριο Εικόνων	ix
Λίστα Συντομογραφιών.....	xi
Κεφάλαιο 1: Εισαγωγή.....	1
1.1: Το ροχαλητό και η αποφρακτική άπνοια - υπόπνοια κατά τον ύπνο	2
1.2: Εργαστηριακή μελέτη ύπνου - Πολυσωματογράφημα	5
1.3: Σύντομη επισκόπηση υπολογιστικών μεθόδων ανάλυσης ακουστικών σημάτων ροχαλητού για τη διάγνωση του OSAHS.....	7
1.4: Σκοπός	11
Κεφάλαιο 2: Θεωρητικό Υπόβαθρο.....	13
2.1: Μέθοδοι φασματικής αφαίρεσης για αποθορυβοποίηση ηχητικών σημάτων.....	14
2.1.1: Ο αλγόριθμος φασματικής αφαίρεσης Boll.....	15
2.1.2: Ο αλγόριθμος φασματικής αφαίρεσης Berouti.....	15
2.1.3: Ο αλγόριθμος πολυζωνικής φασματικής αφαίρεσης Kamath	15
2.1.4: Ο αλγόριθμος φασματικής αφαίρεσης με φίλτρα Wiener.....	16
2.1.5: Επισκόπηση μεθόδων αποθορυβοποίησης.....	16
2.2: Κατάτμηση ηχητικών σημάτων	17
2.2.1: Ενέργεια Σήματος.....	18
2.2.2: Φασματικό Κεντροειδές.....	18
2.3: Χρονοσυχνοτικά χαρακτηριστικά ηχητικών σημάτων.....	18
2.3.1: Αναπαράσταση σήματος στο πεδίο του χρόνου	19
2.3.2: Αναπαράσταση σήματος στο πεδίο της συχνότητας	19
2.3.3: Ενέργεια Βραχέως Χρόνου	21
2.3.4: Ρυθμός διελεύσεων γύρω από το μηδέν.....	22
2.3.5: Ένταση σήματος	23
2.3.6: Διφάσμα σήματος	24
2.3.7: Σάφμα.....	25

2.3.8: Σάφμα Mel συχνοτήτων	26
2.3.9: Αυτοσυσχέτιση σήματος	27
2.3.10: Συνεχής μετασχηματισμός κυματιδίων	28
2.3.11: Φασματογράφημα	29
2.3.12: Συχνότητες φωνοσυντονισμού	30
Κεφάλαιο 3: Ανάπτυξη λογισμικού Snor-Lux	33
3.1: Επιλογή πλατφόρμας για την ανάπτυξη γραφικού διαδραστικού περιβάλλοντος ..	34
3.2: Ανάπτυξη GUI στο MATLAB®	34
3.3: Μετατροπή GUI σε αυτόνομη εφαρμογή	35
3.4: Χαρακτηριστικά του λογισμικού Snor-Lux	35
Κεφάλαιο 4: Υλικό - Μέθοδοι	37
4.1: Σύστημα καταγραφής	38
4.2: Προεπεξεργασία δεδομένων	38
4.3: Μελέτη επίδρασης ρινικού αποσυμφορητικού σε συνδυασμό με κορτικοστεροειδή σε ασθενείς με OSAHS	39
4.3.1: Πρωτόκολλο μελέτης επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή	39
4.3.2: Υλικό μελέτης επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή	41
4.3.3: Μέθοδος υπολογισμού συχνότητας και μέσης - μέγιστης έντασης ροχαλητών	42
4.4: Μελέτη διαγνωστικής ικανότητας χρονοσυχνοτικών χαρακτηριστικών ως προς το OSAHS	43
4.4.1: Επιλογή χρονοσυχνοτικών χαρακτηριστικών για εξέταση της διαγνωστικής τους ικανότητας	44
4.4.2: Υλικό μελέτης διαγνωστικής ικανότητας των FF , $f_{50\%}$, $f_{90\%}$ και f_{peak}	44
4.4.3: Μέθοδος υπολογισμού συχνοτήτων φωνοσυντονισμού ροχαλητών	46
4.4.4: Μέθοδος υπολογισμού $f_{50\%}$, $f_{90\%}$ και f_{peak}	47
4.4.5: Καμπύλες λειτουργικού χαρακτηριστικού δέκτη	47
Κεφάλαιο 5: Αποτελέσματα	51
5.1: Αποτελέσματα μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών	52
5.2: Αποτελέσματα εξαγωγής FF , $f_{50\%}$, $f_{90\%}$ και f_{peak}	52
5.3: Ανάλυση με καμπύλες ROC (Κριτήριο $AHI \geq 10$ ε/ώ)	54
5.3.1: Στατιστική ανάλυση F1 (Κριτήριο $AHI \geq 10$ ε/ώ)	54
5.3.2: Στατιστική ανάλυση F2 (Κριτήριο $AHI \geq 10$ ε/ώ)	56
5.3.3: Στατιστική ανάλυση F3 (Κριτήριο $AHI \geq 10$ ε/ώ)	57

5.3.4: Στατιστική ανάλυση $f_{50\%}$ (Κριτήριο $AHI \geq 10$ ε/ώ).....	58
5.3.5: Στατιστική ανάλυση $f_{90\%}$ (Κριτήριο $AHI \geq 10$ ε/ώ).....	60
5.3.6: Στατιστική ανάλυση f_{peak} (Κριτήριο $AHI \geq 10$ ε/ώ).....	61
5.4: Ανάλυση με καμπύλες ROC (Κριτήριο $AHI \geq 15$ ε/ώ)	63
5.4.1: Στατιστική ανάλυση F1 (Κριτήριο $AHI \geq 15$ ε/ώ)	63
5.4.2: Στατιστική ανάλυση F2 (Κριτήριο $AHI \geq 15$ ε/ώ)	65
5.4.3: Στατιστική ανάλυση F3 (Κριτήριο $AHI \geq 15$ ε/ώ)	66
5.4.4: Στατιστική ανάλυση $f_{50\%}$ (Κριτήριο $AHI \geq 15$ ε/ώ).....	67
5.4.5: Στατιστική ανάλυση $f_{90\%}$ (Κριτήριο $AHI \geq 15$ ε/ώ).....	68
5.4.6: Στατιστική ανάλυση f_{peak} (Κριτήριο $AHI \geq 15$ ε/ώ).....	69
5.5: Σύνοψη αποτελεσμάτων	71
Κεφάλαιο 6: Συζήτηση - Συμπεράσματα.....	75
6.1: Γενικά σχόλια μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού με κορτικοστεροειδή.....	76
6.2: Γενικά σχόλια μελέτης διαγνωστικής ικανότητας FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}	76
6.3: Μελλοντικές προοπτικές.....	78
6.4: Συμπέρασμα.....	79
Βιβλιογραφία	81
Παράρτημα A: Εγχειρίδιο Λειτουργίας του πακέτου Snor-Lux	A-1
A.1: Εισαγωγή.....	A-2
A.2: Εκκίνηση του Snor-Lux	A-2
A.3: Επιφάνεια αρχικής οθόνης	A-3
A.4: Επιφάνεια φόρτωσης αρχείου και αποθρομβοποίησης.....	A-4
A.5: Επιφάνεια κατάτμησης σήματος	A-8
A.6: Επιφάνεια επεξεργασίας σήματος	A-9
A.7: Επιφάνεια επικοινωνίας	A-12

Ευρετήριο Πινάκων

Πίνακας 1.1: Επισκόπηση σημαντικότερης βιβλιογραφίας στο θέμα ανάλυσης ακουστικών σημάτων ροχαλητών	10
Πίνακας 4.1: Χαρακτηριστικά εξεταζομένων (Μελέτη επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή)	42
Πίνακας 4.2: Χαρακτηριστικά εξεταζομένων (Μελέτη διαγνωστικής ικανότητας των FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}).....	45
Πίνακας 5.1: Αποτελέσματα μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών	52
Πίνακας 5.2: Μέση τιμή \pm Τυπική απόκλιση (STD) των FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}	53
Πίνακας 5.3: Αποτελέσματα ανάλυσης ROC για F1 (AHI ≥ 10 ε/ώ).....	55
Πίνακας 5.4: Αποτελέσματα ανάλυσης ROC για F2 (AHI ≥ 10 ε/ώ).....	57
Πίνακας 5.5: Αποτελέσματα ανάλυσης ROC για F3 (AHI ≥ 10 ε/ώ).....	58
Πίνακας 5.6: Αποτελέσματα ανάλυσης ROC για $f_{50\%}$ (AHI ≥ 10 ε/ώ)	59
Πίνακας 5.7: Αποτελέσματα ανάλυσης ROC για $f_{90\%}$ (AHI ≥ 10 ε/ώ)	61
Πίνακας 5.8: Αποτελέσματα ανάλυσης ROC για f_{peak} (AHI ≥ 10 ε/ώ).....	62
Πίνακας 5.9: Αποτελέσματα ανάλυσης ROC για F1 (AHI ≥ 15 ε/ώ).....	64
Πίνακας 5.10: Αποτελέσματα ανάλυσης ROC για F2 (AHI ≥ 10 ε/ώ).....	65
Πίνακας 5.11: Αποτελέσματα ανάλυσης ROC για F3 (AHI ≥ 15 ε/ώ).....	66
Πίνακας 5.12: Αποτελέσματα ανάλυσης ROC για $f_{50\%}$ (AHI ≥ 15 ε/ώ)	67
Πίνακας 5.13: Αποτελέσματα ανάλυσης ROC για $f_{90\%}$ (AHI ≥ 15 ε/ώ)	68
Πίνακας 5.14: Αποτελέσματα ανάλυσης ROC για f_{peak} (AHI ≥ 15 ε/ώ).....	69
Πίνακας 5.15: Συγκεντρωτικός πίνακας αποτελεσμάτων ανάλυσης με καμπύλες ROC.....	71
Πίνακας 5.16: Υπολογιστικό κόστος αλγορίθμων.....	72
Πίνακας A.1: Μέθοδοι επεξεργασίας που προσφέρει το Snor-Lux.....	A-10

Ευρετήριο Εικόνων

Εικόνα 1.1: Τοποθέτηση ηλεκτροδίων κατά τη διάρκεια του PSG.....	6
Εικόνα 1.2: Παράδειγμα πλήρους καταγραφής PSG	7
Εικόνα 2.1: Διάγραμμα αποθρομβοποίησης με φασματική αφαίρεση.....	14
Εικόνα 2.2: Αρχικό σήμα και οι αποθρομβοποιημένες εκδόσεις του	17
Εικόνα 2.3: Αναπαράσταση σήματος ροχαλητού στο πεδίο του χρόνου	19
Εικόνα 2.4: Φάσμα σήματος ροχαλητού σε dB	21
Εικόνα 2.5: Ενέργεια βραχέως χρόνου σήματος ροχαλητού.....	22
Εικόνα 2.6: ZCR σήματος ροχαλητού	23
Εικόνα 2.7: Ένταση σε dB σήματος ροχαλητού	24
Εικόνα 2.8: Διφάσμα σήματος ροχαλητού	25
Εικόνα 2.9: Σάφμα σήματος ροχαλητού	26
Εικόνα 2.10: Σάφμα μεl συχνοτήτων σήματος ροχαλητού	27
Εικόνα 2.11: Αυτοσυσχέτιση σήματος ροχαλητού	28
Εικόνα 2.12: Συντελεστές συνεχούς μετασχηματισμού κυματιδίων σήματος ροχαλητού..	29
Εικόνα 2.13: Φασματογράφημα σήματος ροχαλητού	30
Εικόνα 2.14: Αντιστοιχία FF με τις τρεις πρώτες αρμονικές σήματος σε κλειστό σωλήνα ..	30
Εικόνα 2.15: Φίλτρο υπολογισμού FF σήματος ροχαλητού	31
Εικόνα 4.1: Σχεδιασμός μελέτης επίδρασης ρινικού αποσυμφορητικού σε συνδυασμό με κορτικοστεροειδή.....	41
Εικόνα 4.2: Παράδειγμα παραθύρου Kaiser μήκους 512 δειγμάτων, με $\beta=1.5$	43
Εικόνα 4.3: Παράδειγμα καμπύλης ROC.....	49
Εικόνα 5.1: Οι τρεις πρώτες συχνότητες φωνοσυντονισμού για σήμα ροχαλητού.....	53
Εικόνα 5.2: Οι συχνότητες $f_{50\%}$, $f_{90\%}$ και f_{peak} στο φάσμα σήματος ροχαλητού.....	54
Εικόνα 5.3: Καμπύλη ROC για F1 (AHI ≥ 10 ε/ώ)	54
Εικόνα 5.4: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (F1).....	55
Εικόνα 5.5: Καμπύλη ROC για F2 (AHI ≥ 10 ε/ώ)	56
Εικόνα 5.6: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (F2).....	56
Εικόνα 5.7: Καμπύλη ROC για F3 (AHI ≥ 10 ε/ώ)	57
Εικόνα 5.8: Καμπύλη ROC για $f_{50\%}$ (AHI ≥ 10 ε/ώ).....	58
Εικόνα 5.9: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN ($f_{50\%}$)	59

Εικόνα 5.10: Καμπύλη ROC για $f_{90\%}$ (AHI ≥ 10 ε/ώ).....	60
Εικόνα 5.11: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN ($f_{90\%}$)	60
Εικόνα 5.12: Καμπύλη ROC για f_{peak} (AHI ≥ 10 ε/ώ).....	61
Εικόνα 5.13: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (f_{peak})	62
Εικόνα 5.14: Καμπύλη ROC για F1 (AHI ≥ 15 ε/ώ)	63
Εικόνα 5.15: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (F1)	64
Εικόνα 5.16: Καμπύλη ROC για F2 (AHI ≥ 15 ε/ώ)	65
Εικόνα 5.17: Καμπύλη ROC για F3 (AHI ≥ 15 ε/ώ)	66
Εικόνα 5.18: Καμπύλη ROC για $f_{50\%}$ (AHI ≥ 15 ε/ώ).....	67
Εικόνα 5.19: Καμπύλη ROC για $f_{90\%}$ (AHI ≥ 15 ε/ώ).....	68
Εικόνα 5.20: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN ($f_{90\%}$)	69
Εικόνα 5.21: Καμπύλη ROC για f_{peak} (AHI ≥ 15 ε/ώ).....	70
Εικόνα 5.22: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (f_{peak})	70
Εικόνα A.1: Αρχική οθόνη περιβάλλοντος εργασίας στο MATLAB®	A-2
Εικόνα A.2: Αρχική επιφάνεια του Snor-Lux	A-3
Εικόνα A.3: Μήνυμα επιβεβαίωσης εξόδου από το Snor-Lux.....	A-3
Εικόνα A.4: Επιφάνεια φόρτωσης αρχείου και αποθρομβοποίησης.....	A-4
Εικόνα A.5: Μενού πλοήγησης για την επιλογή αρχείου	A-5
Εικόνα A.6: Παράθυρο σημάτων που προκύπτει από το κουμπί "All Methods"	A-6
Εικόνα A.7: Συμπληρωμένη επιφάνεια φόρτωσης και αποθρομβοποίησης αρχείου	A-7
Εικόνα A.8: Επιφάνεια κατάτμησης σήματος	A-8
Εικόνα A.9: Επιφάνεια επεξεργασίας σήματος	A-9
Εικόνα A.10: Επιφάνεια επικοινωνίας	A-12

Λίστα Συντομογραφιών

Συντομογραφία	Επεξήγηση
AHI	Apnea Hypopnea Index - Δείκτης άπνοιας υπόπνοιας
AUC	Area Under the Curve - Εμβαδόν κάτω από την καμπύλη ROC
BMI	Body Mass Index - Δείκτης μάζας σώματος
CI	Confidence Interval - Διάστημα εμπιστοσύνης
CWT	Continuous Wavelet Transform - Συνεχής Μετασχηματισμός Κυματιδίων
DFT	Discrete Fourier Transform - Διακριτός Μετασχηματισμός Fourier
ECG	Electrocardiography - Ηλεκτροκαρδιογράφημα
EEG	Electroencephalography - Ηλεκτροεγκεφαλογράφημα
EOG	Electrooculography - Ηλεκτροοφθαλμογράφημα
f_s	Sampling Frequency - Συχνότητα δειγματοληψίας
FF	Formant Frequencies - Συχνότητες φωνοσυντονισμού
FFT	Fast Fourier Transform - Ταχύς μετασχηματισμός Fourier
FN	False Negative - Ψευδώς αρνητικά αποτελέσματα
FP	False Positive - Ψευδώς θετικά αποτελέσματα
FT	Fourier Transform - Μετασχηματισμός Fourier
GUI	Graphical User Interface - Γραφικό διαδραστικό περιβάλλον
HMM	Hidden Markov Models - Κρυμμένα μαρκοβιανά μοντέλα
HRNR	Harmonic Regeneration Noise Reduction - Μείωση θορύβου με αρμονική αναπαραγωγή
IDFT	Inverse Discrete Fourier Transform - Αντίστροφος διακριτός μετασχηματισμός Fourier
IFT	Inverse Fourier Transform - Αντίστροφος μετασχηματισμός Fourier
LPC	Linear Predictive Coding - Κωδικοποίηση γραμμικής πρόβλεψης
MCR	Matlab component runtime
MFC	Mel-Frequency Cepstrum - Σάφμα mel συχνοτήτων
MFCC's	Mel-Frequency Cepstrum Coefficients - Συντελεστές σάφματος mel συχνοτήτων
OSAHS	Obstructive sleep apnea hypopnea syndrome - Σύνδρομο αποφρακτικής άπνοιας υπόπνοιας
PSG	Polysomnogram - Πολυσωματογράφημα
ROC	Receiver Operating Characteristic - Λειτουργική χαρακτηριστικού δέκτη
SE	Sensibility - Ευαισθησία
SNR	Signal to Noise Ratio - Λόγος σήματος προς θόρυβο
SPC	Specificity - Ειδικότητα
STD	Standard Deviation - Τυπική απόκλιση
STE	Short Time Energy - Ενέργεια βραχέως χρόνου
TN	True Negative - Αληθώς αρνητικά αποτελέσματα
TNR	True Negative Rate - Ποσοστό αληθών αρνητικών αποτελεσμάτων
TP	True Positive - Αληθώς θετικά αποτελέσματα
TPR	True Positive Rate - Ποσοστό αληθών θετικών αποτελεσμάτων
TSNR	Two-Step Noise Reduction - Μείωση θορύβου σε δύο βήματα
ZCR	Zero Crossings - Διελεύσεις γύρω από το μηδέν
ε/ώ	Επεισόδια ανά ώρα (Μονάδα μέτρησης AHI)

Κεφάλαιο 1: Εισαγωγή

Στο κεφάλαιο αυτό παρέχονται οι βασικές πληροφορίες που αφορούν το ροχαλητό και τις διαταραχές κατά τη διάρκεια του ύπνου που σχετίζονται με το ροχαλητό, όπως είναι η άπνοια, η υπόπνοια και το σύνδρομο της αποφρακτικής άπνοιας - υπόπνοιας (*Obstructive sleep apnea hypopnea syndrome* - OSAHS). Επίσης περιγράφεται το πιο ευρέως διαδεδομένο εργαλείο διάγνωσης του OSAHS, το πολυυπνογράφημα (PSG). Ακολούθως παρουσιάζεται σύντομη επισκόπηση των μελετών που έχουν γίνει μέχρι σήμερα και αφορούν την προσπάθεια διάγνωσης και εκτίμησης σοβαρότητας του OSAHS μέσω εξαγωγής χρονοσυχνοτικών χαρακτηριστικών των ηχητικών σημάτων ροχαλητού. Γίνεται ιδιαίτερη αναφορά στα ηχητικά σήματα ροχαλητού και τα χαρακτηριστικά αυτών, αφού αποτελούν το κεντρικό αντικείμενο έρευνας και επεξεργασίας στην διπλωματική εργασία. Ο αναλυτικός σκοπός της διπλωματικής εργασίας παρουσιάζεται στο τέλος του κεφαλαίου.

1.1: Το ροχαλητό και η αποφρακτική άπνοια - υπόπνοια κατά τον ύπνο

Το ροχαλητό αποτελεί μια επικρατούσα διαταραχή, η οποία επηρεάζει το 20 - 40% του γενικού πληθυσμού, μόνιμα ή περιστασιακά [1]. Ένα από τα κύρια προβλήματα για τους ερευνητές του τομέα είναι ότι δεν υπάρχει ακριβής ορισμός του ροχαλητού. Σύμφωνα με το λεξικό αγγλικής γλώσσας American Heritage ροχαλητό θεωρείται "η αναπνοή κατά τη διάρκεια του ύπνου με σκληρούς ήχους ρουθουνίσματος που προκαλούνται από δονήσεις της μαλθακής υπερώας" ("*to breathe during sleep with harsh, snorting noises caused by vibration of the soft palate*") [2].

Η αιτία πρόκλησης του ροχαλητού είναι η παρεμπόδιση της ομαλής διέλευσης του αέρα κάτι το οποίο μπορεί να οφείλεται σε πολλούς παράγοντες. Το ροχαλητό προέρχεται από κραδασμούς των φαρυγγικών τοιχωμάτων και κυρίως της μαλθακής υπερώας των αμυγδαλών και της σταφυλής. Άλλα ανατομικά εμπόδια που παρεμποδίζουν τη διέλευση του αέρα μπορεί να είναι το στραβό ρινικό διάφραγμα, οι υπερτροφικές ρινικές κόγχες καθώς και οι αδενοειδείς εκβλαστήσεις στα παιδιά. Υπάρχουν αρκετοί προδιαθεσικοί παράγοντες που θεωρείται ότι συμβάλουν στο ροχαλητό όπως είναι η παχυσαρκία σε άτομα με κοντό λαιμό, το κάπνισμα, η χρήση ηρεμιστικών και υπνωτικών ουσιών, η ύπτια θέση του σώματος κατά τη διάρκεια του ύπνου καθώς και η λήψη βαριών γευμάτων πριν την κατάκλιση [3].

Το ροχαλητό μπορεί να εμφανιστεί σε συνδυασμό με ένα διαταραγμένο πρότυπο ύπνου ενώ επίσης μπορεί να συνδέεται με πολλά συμπτώματα όπως είναι η άπνοια και η αποφρακτική υπόπνοια. Η άπνοια και η υπόπνοια αποτελούν παθήσεις που αφορούν τη διαταραχή του ύπνου και χαρακτηρίζονται από επαναλαμβανόμενα, παρατεταμένα επεισόδια πλήρους ή μερικής διακοπής, αντίστοιχα, της αναπνοής κατά τη διάρκεια του ύπνου. Σαν άπνοια ορίζεται η πλήρης διακοπή της αναπνοής για τουλάχιστον δέκα δευτερόλεπτα. Κατά τη διάρκεια της άπνοιας δεν υπάρχει καθόλου ήχος αναπνοής ενώ η συνέχιση της αναπνοής συνοδεύεται με μια σειρά ροχαλητών. Υπάρχουν τρία είδη άπνοιών:

- Κεντρική άπνοια: Απουσία της αναπνευστικής προσπάθειας καθ' όλη τη διάρκεια του επεισοδίου

- Αποφρακτική άπνοια: Συνοδεύεται από αναπνευστική προσπάθεια καθ' όλη τη διάρκεια του επεισοδίου
- Μικτή άπνοια: Συνδυάζει την απουσία της αναπνευστικής προσπάθειας κατά το αρχικό τμήμα του επεισοδίου και την εμφάνιση της αναπνευστικής προσπάθειας κατά το δεύτερο μέρος [4].

Κατά την υπόπνοια η διακοπή της αναπνοής είναι μερική και παρατηρείται ελάττωση του αέρα σε ποσοστό 50% ή μεγαλύτερο. Η υπόπνοια μπορεί να συνοδεύεται από ελάττωση στα ίδια ποσοστά των αναπνευστικών κινήσεων θώρακος - κοιλιάς και πτώσης του κορεσμού αιμοσφαιρίνης κατά 4% ή μεγαλύτερο.

Το σύνδρομο της αποφρακτικής άπνοιας - υπόπνοιας είναι ένας τύπος υπνικής άπνοιας που οφείλεται σε απόφραξη του άνω αεραγωγού, κατά τη διάρκεια επίμονων αναπνευστικών προσπαθειών. Η σοβαρότητα του OSAHS καθορίζεται από το δείκτη απνοιών - υποπνοιών (*Apnea/Hypopnea Index* - AHI) ο οποίος ορίζεται ως ο αριθμός απνοιών και υποπνοιών ανά ώρα ύπνου. Η μονάδα μέτρησης του AHI ορίζεται ως τα επεισόδια απνοιών και υποπνοιών ανά ώρα (ε/ώ - επεισόδια ανά ώρα) Υπάρχουν τρεις κατηγορίες σοβαρότητας του συνδρόμου με βάση το AHI: [5], [6]

- 5-15 άπνοιες - υπόπνοιες ανά ώρα: Ήπια σοβαρότητα
- 15-30 άπνοιες - υπόπνοιες ανά ώρα: Μέτρια σοβαρότητα
- Περισσότερες από 30 άπνοιες - υπόπνοιες ανά ώρα: Μεγάλη σοβαρότητα

Το σύνδρομο αποφρακτικής άπνοιας - υπόπνοιας στον ύπνο αποτελεί την πιο συχνή από τις διαταραχές της αναπνοής. Αποτελεί σοβαρό πρόβλημα ιδίως για τις χώρες της Ευρώπης, της Β. Αμερικής και της Ασίας, λόγω της υψηλής της επίπτωσης. Εκτιμάται ότι το 6% των ενηλίκων πάσχει από αποφρακτική άπνοια, με το 75% αυτών να παραμένει αδιάγνωστο. Επιδημιολογικά δεδομένα σε πληθυσμό της Β. Αμερικής δείχνουν ότι το 20% περίπου του ενήλικου πληθυσμού πάσχει τουλάχιστον από ήπιο σύνδρομο ($5 \leq \text{AHI} \leq 15$ ε/ώ) ενώ το 7% περίπου των ενηλίκων πάσχει από μέτριο έως σοβαρό σύνδρομο ($\text{AHI} > 15$ ε/ώ). Το σύνδρομο παρουσιάζει μεγαλύτερη επίπτωση στον ανδρικό πληθυσμό [7].

Μελέτη στους Έλληνες, και συγκεκριμένα σε 50 οδηγούς σιδηροδρόμων, έδειξε: μέση τιμή AHI 11 ± 14 επεισόδια/ώρα. Ανάλογα με την τιμή του AHI, κατετάγησαν σε τρεις

Κεφάλαιο 1

κατηγορίες: Ομάδα πρώτη: φυσιολογική αναπνοή στον ύπνο ($n = 19$, $AHI < 5$ ε/ώ). Ομάδα δεύτερη: ήπιο OSA ($n = 20$, $5 \leq AHI \leq 15$ ε/ώ). Ομάδα τρίτη: μέτριο - σοβαρό OSA ($n = 11$, $AHI > 15$ ε/ώ) [8].

Η υποκείμενη παθοφυσιολογία του OSAHS είναι πολύπλοκη και δεν είναι πλήρως κατανοητή. Ωστόσο είναι γενικά αποδεκτό ότι η σταθερότητα και η βατότητα των άνω αεραγωγών εξαρτάται από τη δράση του στοματοφαρυγγικού αεραγωγού και των απαγωγέων μυών, οι οποίοι ενεργοποιούνται με ρυθμικό τρόπο σε κάθε εισπνοή. Όταν η δύναμη που ασκείται από τους απαγωγείς μύες, για μια δεδομένη διατομή του ανώτερου αεραγωγού, δεν ξεπερνά την αρνητική πίεση που παράγει η αναπνευστική δραστηριότητα στο διάφραγμα και τους μεσοπλεύριους μύες, ο άνω αεραγωγός υπόκειται σε κατάρρευση. Η παρεμπόδιση της σωστής λειτουργίας του άνω αεραγωγού μπορεί να συμβεί και όταν η πίεση αναρρόφησης είναι πολύ ψηλή, ή εάν οι δυνάμεις που ασκούν οι διαστολικοί μύες του άνω αεραγωγού είναι πολύ μικρές για κάθε δεδομένη πίεση αναρρόφησης. Παράγοντες που συμβάλλουν στην απόφραξη του άνω αεραγωγού είναι η υπερβολική απώλεια μυϊκού τόνου, η ανατομική στένωση του άνω αεραγωγού και ελαττωματικά προστατευτικά αντανακλαστικά του άνω αεραγωγού [9].

Η σοβαρότητα του OSAHS είναι πολύ μεγάλη καθώς οι άνθρωποι που πάσχουν από το σύνδρομο έχουν σημαντικά μειωμένη ποιότητα ύπνου. Οι συνέπειες εκτείνονται στο σωματικό, ψυχολογικό αλλά και κοινωνικό τομέα και περιλαμβάνουν κούραση, μειωμένη επαγγελματική απόδοση, ατυχήματα κάθε μορφής, διαταραχές συγκέντρωσης και προσοχής, γαστροοισοφαγική παλινδρόμηση, μαθησιακές δυσκολίες, προβλήματα ύπνου του συντρόφου στο κρεβάτι και συζυγικές δυσκολίες. Μελέτη που έγινε στον Καναδά, σε 1600 άτομα, 800 με OSAHS και 800 υγιή, σε διάρκεια τριών ετών έδειξε ότι τα άτομα στην ομάδα με OSAHS είχαν διπλάσιες πιθανότητες να εμπλακούν σε αυτοκινητιστικό δυστύχημα και τριπλάσιες έως και πενταπλάσιες πιθανότητες να εμπλακούν σε σοβαρό δυστύχημα με τραυματισμό προσώπων [10]. Το σύνηθες ροχαλητό, σύμφωνα με άλλη μελέτη, αποτελεί σε αρκετές περιπτώσεις πέρα από τα προβλήματα στον ύπνο, συζυγική δυσαρμονία, διαζύγιο, οικογενειακή επιθετικότητα και μέχρι και αυτοκτονία [11]. Στο πλαίσιο διερεύνησης του μεγάλου αυτού προβλήματος έχουν γίνει πολλές μακρόχρονες επιδημιολογικές μελέτες, όπως η Wisconsin Sleep Cohort και η Busselton Health Study. Έχει αποδειχτεί ότι η σοβαρή αποφρακτική άπνοια συνδυάζεται με τριπλάσιο κίνδυνο

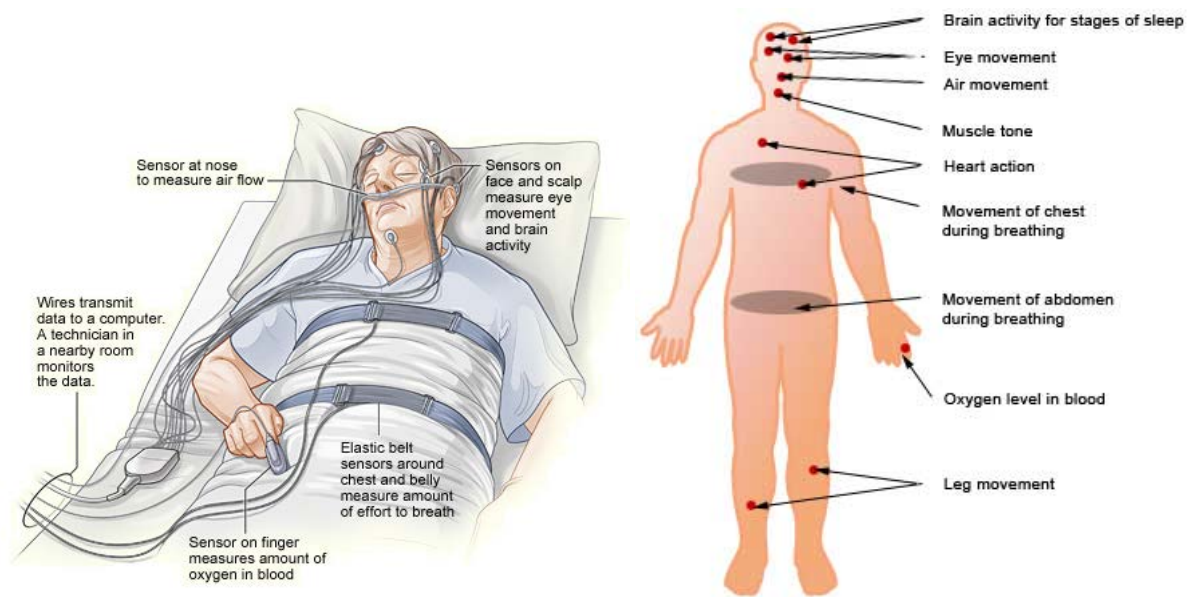
θνησιμότητας από κάθε άλλη αιτία και με υψηλότερο κίνδυνο για καρδιαγγειακά νοσήματα σε διάρκεια παρακολούθησης 18 ετών. Αποτελεί, επίσης, ανεξάρτητο παράγοντα κινδύνου ανάπτυξης νοσημάτων όπως η υπέρταση, τα αγγειακά εγκεφαλικά επεισόδια, οι αρρυθμίες και τα στεφανιαία επεισόδια [12].

1.2: Εργαστηριακή μελέτη ύπνου - Πολυσωματογράφημα

Για την ορθή διάγνωση του OSAHS απαιτείται η καταγραφή, μελέτη και αξιολόγηση πολλών παραμέτρων που σχετίζονται με την αναπνοή και τη σταδιοποίηση του ύπνου. Ο ακρογωνιαίος λίθος για τη διάγνωση του OSAHS σήμερα είναι η πολυσωματογραφική μελέτη ή πολυσωματογράφημα (*polysomnogram* - PSG). Το PSG πραγματοποιείται κατά τη διάρκεια νυκτερινού ύπνου, στο εργαστήριο ύπνου. Το εργαστήριο ύπνου περιλαμβάνει ένα ήσυχο, ηχομονωμένο εάν είναι δυνατό, μονόκλινο δωμάτιο στο οποίο κοιμάται ο εξεταζόμενος και ένα δεύτερο, ξεχωριστό χώρο όπου βρίσκονται τα καταγραφικά μηχανήματα και το προσωπικό που παρακολουθεί την καταγραφή. Είναι απαραίτητο οι δύο χώροι να διαχωρίζονται αφού τα καταγραφικά όργανα, ο εξεταζόμενος και το ιατρικό προσωπικό δεν μπορούν να συνυπάρξουν λόγω πιθανότητας αλλοίωσης της καταγραφής και κατά συνέπεια των αποτελεσμάτων της εξέτασης.

Οι παράμετροι που καταγράφονται τυπικά σε ένα PSG είναι το ηλεκτροεγκεφαλογράφημα (EEG) δύο έως τεσσάρων απαγωγών, το ηλεκτροοφθαλμογράφημα (EOG), η ροή αέρα από και προς τη μύτη και το στόμα, ο μυϊκός τόνος της σιαγώνας, η κίνηση των κάτω άκρων, η κίνηση του θώρακα και της κοιλιακής χώρας, ο κορεσμός της αιμοσφαιρίνης με τη χρήση οξυμέτρου, η θέση του σώματος κατά τον ύπνο (ύπτια, αριστερή ή δεξιά πλάγια) και ο καρδιακός ρυθμός μέσω ηλεκτροκαρδιογραφήματος (ECG). Η καταγραφή των παραμέτρων που σχετίζονται με την σταδιοποίηση του ύπνου είναι ιδιαίτερα σημαντική λόγω του ότι απαιτείται να είναι γνωστός ο ακριβής χρόνος ύπνου του εξεταζόμενου ώστε να είναι εφικτός ο υπολογισμός του AHI. Η τοποθέτηση των απαιτούμενων ηλεκτροδίων και οργάνων μέτρησης φαίνεται στην Εικόνα1. 1 [13].

Κεφάλαιο 1

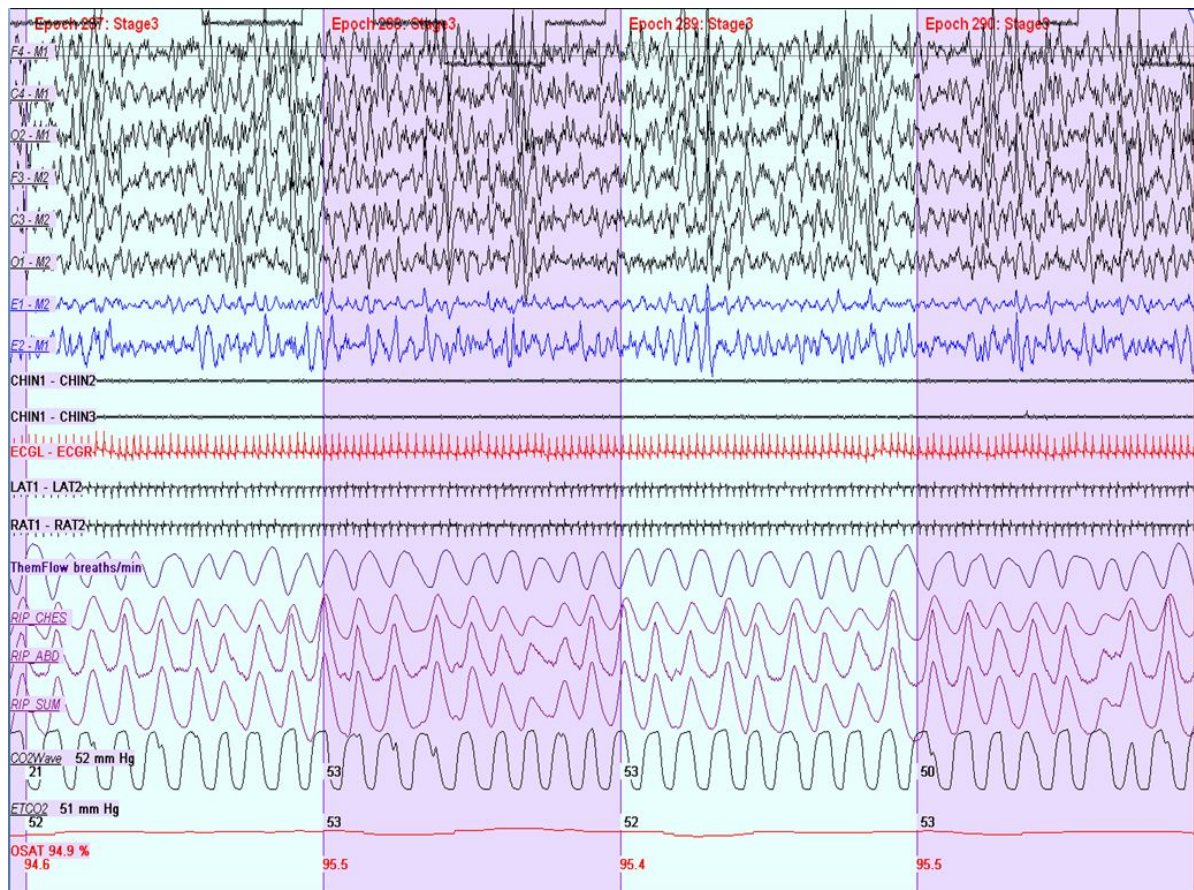


Εικόνα 1.1: Τοποθέτηση ηλεκτροδίων κατά τη διάρκεια του PSG

[Πηγές: <http://www.osrmedical.com> (Αριστερά), <http://mediconweb.com> (Δεξιά)]

Το κόστος μιας μελέτης PSG ανέρχεται στην Ελλάδα σήμερα στα 147 ευρώ ενώ η μέση τιμή στην Αμερική ανέρχεται στα 2030 ευρώ. [Πηγή: <http://www.newchoicehealth.com>].

Στην Εικόνα 1.2 μπορούμε να δούμε την πλήρη καταγραφή ενός PSG και τα διάφορα κανάλια που την αποτελούν. Όπως μπορεί να παρατηρηθεί το αποτέλεσμα μιας μελέτης ύπνου είναι αρκετά σύνθετο και απαιτεί εμπειρία για την σωστή ανάγνωση και ανάλυση των προς εξαγωγή συμπερασμάτων. Η επεξεργασία των δεδομένων για τη σταδιοποίηση του ύπνου, τον καθορισμό μικροαφυπνίσεων και αναπνευστικών επεισοδίων, την εκτίμηση του ECG και της κίνησης των κάτω άκρων γίνεται από εξειδικευμένο ιατρό ή από εξειδικευμένο τεχνικό προσωπικό του εργαστηρίου ύπνου υπό την επίβλεψη και ευθύνη του εξειδικευμένου ιατρού του κέντρου. Δεν θεωρείται αποδεκτή ακόμα η αυτόματη, μέσω λογισμικού, διάγνωση σταδίων ύπνου, μικροαφυπνίσεων και αναπνευστικών επεισοδίων λόγω των αυξημένων σφαλμάτων που προκύπτουν με τη χρήση των τώρα υπαρχόντων αλγορίθμων [14].



Εικόνα 1.2: Παράδειγμα πλήρους καταγραφής PSG
[Πηγή: <http://www.thoracic.org>]

Τελευταίως έχει αυξηθεί η σημασία των ακουστικών σημάτων ροχαλητού αφού πολλοί ερευνητές προσπαθούν να ανακαλύψουν σύνδεση αυτών με το OSAHS ώστε να αποφευχθεί η χρήση PSG λόγω του χρόνου και κόστους που απαιτεί. Αυτό οδήγησε στην ενσωμάτωση μικροφώνων στα εργαστήρια ύπνου που ηχογραφούν καθ' όλη τη διάρκεια της εξέτασης τον εξεταζόμενο. Τα ακουστικά σήματα καταγράφονται σε υπολογιστή στο δωμάτιο ελέγχου και χρησιμοποιούνται μέχρι σήμερα καθαρά για ερευνητικούς σκοπούς.

1.3: Σύντομη επισκόπηση υπολογιστικών μεθόδων ανάλυσης ακουστικών σημάτων ροχαλητού για τη διάγνωση του OSAHS

Με την ανάπτυξη της επιστήμης της ψηφιακής επεξεργασίας σήματος τα τελευταία χρόνια καθώς και με την εξέλιξη της τεχνολογίας στους τομείς της διαθέσιμης επεξεργαστικής ισχύος και αποθηκευτικού όγκου δεδομένων, οι ερευνητές σε κάθε τομέα ανέπτυξαν νέες μεθόδους προσέγγισης για την επίλυση προβλημάτων. Έτσι και στον

Κεφάλαιο 1

τομέα της διάγνωσης του OSAHS έγινε προσπάθεια παράκαμψης της ανάγκης χρήσης PSG και χρήση μόνον ηχητικών σημάτων ροχαλητού για διάγνωση του συνδρόμου.

Η αρχή έγινε από τους Fiz et al. το 1996 [15] όπου αναλύθηκαν τα φάσματα ισχύος ροχαλητών 17 ατόμων, 10 απνοιικών και 7 υγιών, που ηχογραφήθηκαν σε εργαστήριο ύπνου. Τα πρώτα συμπεράσματα ήταν θετικά για την διαγνωστική ικανότητα του OSAHS αφού παρατηρήθηκε διαφορά στα φάσματα ισχύος των υγιών ατόμων σε σχέση με τους απνοιικούς. Συγκεκριμένα, η επεξεργασία των δεδομένων έδειξε ότι το φάσμα των ροχαλητών των υγιών ατόμων είχε μια θεμελιώδη συχνότητα και αρκετές αρμονικές ενώ στο φάσμα ροχαλητού απνοιικών η ενέργεια ήταν διασκορπισμένη σε ένα στενότερο εύρος συχνοτήτων, χωρίς αρμονικές. Μετά από αυτή την έρευνα πολλοί μελετητές ασχολήθηκαν με το θέμα της ηχητικής ανάλυσης των σημάτων ροχαλητού και εξετάστηκε η διαγνωστική ικανότητα πολλών και αρκετές φορές ιδιαίτερα σύνθετων χαρακτηριστικών των ακουστικών σημάτων ροχαλητού.

Σε πιο πρόσφατες μελέτες, στην [16] γίνεται προσπάθεια ορισμού του ροχαλητού αφού δεν υπάρχει ακριβής ορισμός από κάποιο παγκόσμιο ιατρικό οργανισμό και μελέτη της διαγνωστικής ικανότητας ενός νέου μεγέθους που ορίζουν οι συγγραφείς, την πιθανότητα άλματος τόνου (pitch jump probability). Τα αναπνευστικά επεισόδια κατά τη διάρκεια του ύπνου χωρίστηκαν σε τρεις κατηγορίες: επεισόδιο αναπνοής (breathing episode), επεισόδια ροχαλητού (snoring episode) και καθαρό επεισόδιο αναπνοής (pure-breath episode). Όσο αφορά την πιθανότητα άλματος τόνου, τα δεδομένα προέρχονταν από 35 άτομα και το συμπέρασμα που εξήχθη ήταν ότι είναι δυνατή η διάγνωση του OSAHS με το συγκεκριμένο μέγεθος αλλά απαιτεί τεράστιες βάσεις δεδομένων για να πετύχει την απαιτούμενη ακρίβεια.

Στην [17] εξετάζεται ποιοτικά η ικανότητα διάγνωσης του OSAHS με τη βοήθεια του διφάσματος (bispectrum). Το υλικό που αναλύθηκε ήταν δέκα ροχαλητά από κάθε άτομο για 16 άτομα, 9 απνοιικούς και 7 υγιείς. Τα αποτελέσματα της διφασματικής ανάλυσης δεν ήταν απόλυτα όμως έδειξαν ότι ίσως είναι εφικτή η διάγνωση του συνδρόμου αφού τα ροχαλητά απνοιικών επιδείκνυαν σε μεγαλύτερο βαθμό φαινόμενα σύζευξης φάσης από αυτά των υγιών. Παρόμοια μελέτη ήταν η [18] όπου αναλύθηκαν 80 επεισόδια ροχαλητού μόνον απνοιικών ατόμων και αναλύθηκαν ποιοτικά τα αποτελέσματα του συνεχούς μετασχηματισμού κυματιδίων. Η ανάλυση των δεδομένων έδειξε ότι για κάθε ασθενή

υπάρχει ένα συγκεκριμένο εύρος συχνοτήτων που διεγείρεται για κάθε επεισόδιο ροχαλητού αλλά με διαφορετικό πλάτος κάθε φορά ενώ η κυματομορφή για κάθε σειρά ροχαλητών ακολουθούν συγκεκριμένο μοτίβο.

Στην [19] μελετήθηκε η ικανότητα διάγνωσης OSAHS μέσω των συχνοτήτων φωνοσυντονισμού. Ως υλικό, αναλύθηκαν ροχαλητά από 40 άτομα, 30 απνοϊκούς και 10 υγιείς, με σαφείς ακουστικές συνιστώσες δόνησης. Τα ροχαλητά αρχικά αποθορυβοποιήθηκαν και ακολούθως έγινε εξαγωγή των τριών συχνοτήτων φωνοσυντονισμού (F1, F2, F3) με τη βοήθεια του αλγορίθμου γραμμικής πρόβλεψης (LPC - Linear predictive coding). Αναλύθηκαν τα αποτελέσματα μόνο για την πρώτη συχνότητα (F1) σύμφωνα με τα οποία η συχνότητα F1 είναι υψηλότερη σε ροχαλητά απνοϊκών από ότι σε υγιών, με βέλτιστο κατώφλι διαχωρισμού υγιών - απνοϊκών, τα 470 Hz.

Εκτός από τις αναφερθείσες, έχουν γίνει αρκετές προσπάθειες για την εξέταση της διαγνωστικής ικανότητας με χρήση διαφόρων μεγεθών: των γραμμικών αλληλεπιδράσεων σε μη στατικά σήματα ροχαλητού [20], ανάλυση τόνου και ολική απόκριση αεραγωγού [21], ακουστική ανάλυση βασισμένη στη θεωρία του χάους [22], ανάλυση φασματικής περιβάλλουσας [23] και πολλών άλλων. Στις περισσότερες μελέτες τα αποτελέσματα είναι θετικά όσο αφορά την ικανότητα διάγνωσης του συνδρόμου αποφρακτικής άπνοιας - υπόπνοιας χρησιμοποιώντας μόνο τις ακουστικές ηχογραφήσεις ροχαλητών, όμως δεν έχει φτάσει ακόμα κάποιο μέγεθος την απαιτούμενη ακρίβεια διάγνωσης ώστε να αντικατασταθεί η ανάγκη πολυσωματογραφήματος για τη διάγνωση. Οι μέθοδοι, τα μεγέθη που μελετήθηκαν και τα αποτελέσματα και συμπεράσματα στα οποία κατέληξαν οι σημαντικότερες έρευνες στο θέμα των διαταραχών ύπνου και ιδιαίτερα του OSAHS συνοψίζονται σε ένα άρθρο από τους Pevernagie et al [1].

Σε αρκετές μελέτες γίνεται προσπάθεια υποβοήθησης των υπόλοιπων ερευνητών στο χώρο, με ανάπτυξη αλγορίθμων και τεχνικών για την αυτόματη αναγνώριση και κατάτμηση ροχαλητών, για την κατηγοριοποίησή τους και για την αποθορυβοποίησή τους. Προτάθηκε στην [24] ένας αλγόριθμος που χρησιμοποιεί τη μέθοδο των κρυμμένων Μαρκοβιανών μοντέλων (HMM - Hidden Markov Models) για την αυτόματη αναγνώριση και κατάτμηση ροχαλητών από πολύωρες ηχογραφήσεις. Στην [25] χρησιμοποιούνται οι δια- και ενδο- ατομικές φασματικές κατανομές ενέργειας των σημάτων ροχαλητού για την κατηγοριοποίηση ροχαλητού / μη ροχαλητού. Και οι δύο μελέτες καταλήγουν στο

Κεφάλαιο 1

συμπέρασμα ότι τα μεγέθη που προτείνουν είναι ικανά να χρησιμοποιηθούν στην αυτόματη κατηγοριοποίηση όμως τα αποτελέσματά τους, αν και πολύ θετικά δεν παρέχουν την απαιτούμενη ακρίβεια ώστε να μπορούν να υιοθετηθούν. Η [26] ασχολείται με την αποθоруβοποίηση των ηχογραφήσεων που περιέχουν ροχαλητά και χρησιμοποιεί το μετασχηματισμό κυματιδίων. Προτείνει ένα κατώφλι για την αποθоруβοποίηση στο πεδίο των κυματιδίων το οποίο εξαρτάται από τη συσχέτιση της κλίμακας του μετασχηματισμού (level correlation dependent threshold), με το οποίο επιτυγχάνονται καλύτερα αποτελέσματα από τους ήδη υπάρχοντες αλγορίθμους.

Στον Πίνακα 1.1 παραθέτονται τα σημαντικότερα άρθρα της παγκόσμιας βιβλιογραφίας με τη μεθοδολογία που ακολουθήθηκε και τα κυριότερα αποτελέσματα τους.

Πίνακας 1.1: Επισκόπηση σημαντικότερης βιβλιογραφίας στο θέμα ανάλυσης ακουστικών σημάτων ροχαλητών

A/A	B.A.	Συγγραφείς Έτος	Περιπτώσεις	Μεθοδολογία	Αποτελέσματα
1	[15]	Fiz et al 1996	A: 10 Y: 7	Συχνότητα Κορυφής f_{peak} Συχνότητα Μέσου f_{mean} Συχνότητα Μεγίστου f_{max}	Υγιείς (Hz): $f_{peak} = 264 \pm 107$ $f_{mean} = 325 \pm 58$ $f_{max} = 462 \pm 85$ Απνοιικοί (Hz): $f_{peak} = 157 \pm 136$ $f_{mean} = 223 \pm 147$ $f_{max} = 455 \pm 199$
2	[16]	Abeyratne et al 2005	A: 12 Y:4	Αυτόματη κατάτμηση ροχαλητών Άλλατα τόνου σε ενδο-ροχαλητά (Intra snore pitch jump - ISPJ)	Απόδοση Κατάτμησης 99.6% Απόδοση ISPJ Ευαισθησία: 83% Ειδικότητα: 67% Κατώφλι AHI: 10 ε/ώ
3	[24]	Duckitt et al 2006	6 ηχογρ/σεις	Ταξινόμηση με χρήση Κρυμμένων Μαρκοβιανών Μοντέλων (HMM)	Ακρίβεια ταξινόμησης: 82-89%
4	[25]	Cavusoglu et al 2007	A: 12 Y: 18	Ταξινόμηση με χρήση της κατανομής φασματικής ενέργειας των ροχαλητών	Ακρίβεια ταξινόμησης: 97.3% για υγιείς 90.2% για συνδυασμό υγιών - απνοιικών
5	[17]	Keong et al 2007	A: 9 Y: 7	Διφασματική ανάλυση	Υψηλότερος βαθμός φαινομένων σύζευξης φάσης για τους απνοιικούς εξεταζόμενους
6	[18]	Matsiki et al 2007	A: 7 Y: 0	Συνεχής Μετασχηματισμός Κυματιδίων με χρήση του κυματιδίου Morlet	Το φάσμα συχνοτήτων είναι χαρακτηριστικό για κάθε ασθενή

7	[19]	Keong et al 2008	A: 30 Y: 10	Συχνότητες Φωνοσυντονισμού	Υγιείς (Hz): $f_1 = 360 \pm 187$ Απνοιικοί (Hz): $f_1 = 724 \pm 201$ Βέλτιστο κατώφλι διαχωρισμού: $f_1 = 470$ Hz
8	[26]	Keong et al 2008	A: 30 Y: 10	Βελτίωση σημάτων ροχαλητού με αναλλοίωτο στη μετατόπιση μετασηματισμό κυματιδίων	Βελτίωση στο σηματοθορυβικό λόγο σε σχέση με τις άλλες μεθόδους κατά 1-8 dB
9	[20]	Keong et al 2008	A: 30 Y: 10	Τετραγωνική ζεύξη φάσεως μέσω δισυνοχικής ανάλυσης κυματιδίων PF1: συχνότητα με μεγαλύτερη ζεύξη PFS: άθροισμα συχνοτήτων κορυφών	Ευαισθησία: 85% Ειδικότητα: 90.7% Βέλτιστο κατώφλι PF1: 285 Hz Βέλτιστο κατώφλι PSF: 492 Hz
10	[21]	Karunajeewa et al 2010	A: 28 Y: 13	Τόνος (Pitch) και Ολική απόκριση αεραγωγού (TAR)	Ευαισθησία: 89.3% Ειδικότητα: 92.3%

B.A.: Βιβλιογραφική αναφορά, A: Απνοιικοί, Y: Υγιείς

1.4: Σκοπός

Η προσπάθεια διάγνωσης του OSAHS μέσω επεξεργασίας των ακουστικών σημάτων ροχαλητού τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας γίνεται εδώ και περίπου 15 χρόνια, κατά τη διάρκεια των οποίων έχουν μελετηθεί ποικίλα χαρακτηριστικά των σημάτων αυτών. Τα αποτελέσματα αρκετών από αυτές τις μελέτες είναι ιδιαίτερα ενθαρρυντικά, ενώ πολλοί ερευνητές δείχνουν πεπεισμένοι ότι είναι εφικτή η διάγνωση ύπαρξης αλλά και σοβαρότητας του OSAHS χρησιμοποιώντας μόνο τα ηχητικά σήματα ροχαλητού.

Σκοπός της διπλωματικής εργασίας ήταν η ανάπτυξη προγράμματος εξαγωγής χρονοσυχνοτικών χαρακτηριστικών ακουστικών σημάτων ροχαλητού για την υποβοήθηση της διάγνωσης του OSAHS και ο έλεγχος της διαγνωστικής ικανότητας μερικών από αυτά με τη χρήση ανάλυσης καμπυλών ROC.

Στα πλαίσια της διπλωματικής εργασίας αναπτύχθηκε αρχικά ένα φιλικό προς το χρήστη πακέτο λογισμικού στο MATLAB®, το Snor-Lux, το οποίο παρέχει τη δυνατότητα αποθορυβοποίησης, κατάτμησης και εν συνεχεία εξαγωγής πολλών χρονοσυχνοτικών χαρακτηριστικών σημάτων ροχαλητού, με απώτερο σκοπό υποβοήθηση της

Κεφάλαιο 1

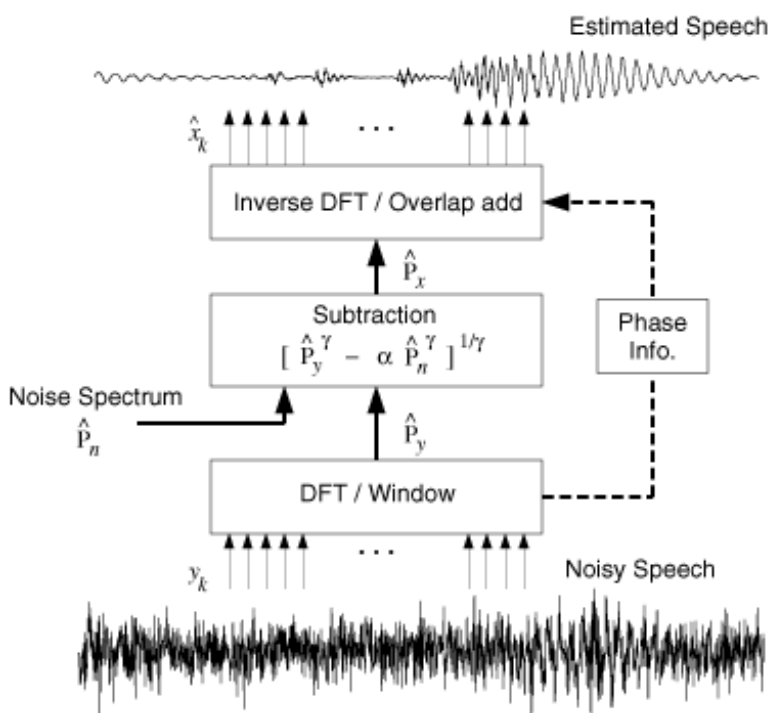
αυτοματοποιημένης διάγνωσης του OSAHS. Ακολούθως το Snor-Lux και οι αλγόριθμοι που περιέχει αναπτύχθηκαν κατάλληλα για την αξιοποίηση σε δύο επιμέρους μελέτες. Η πρώτη αφορά την αξιολόγηση της επίδρασης του συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών σε ασθενείς με OSAHS. Εξήχθησαν η συχνότητα ροχαλητού και η μέση και μέγιστη ένταση ροχαλητών των συμμετεχόντων στην έρευνα. Τα μεγέθη αυτά χρησιμοποιήθηκαν ως μέτρο σύγκρισης της επίδρασης του συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών σε απνοϊκούς ασθενείς, για τον καθορισμό της αποδοτικότητας και αποτελεσματικότητάς της ως μέτρο αντιμετώπισης του OSAHS. Η δεύτερη μελέτη αφορά την ικανότητα διάγνωσης του OSAHS μέσω των τριών πρώτων συχνοτήτων φωνοσυντονισμού των ροχαλητών (formant frequencies - FF) και τριών ενδεικτικών συχνοτήτων του φάσματος (μέση συχνότητα $f_{50\%}$, μέγιστη συχνότητα: $f_{90\%}$, και συχνότητα κορυφής f_{peak}). Αντικείμενο επεξεργασίας για την πρώτη μελέτη αποτέλεσαν τρεις ολονύκτιες ηχογραφήσεις εξεταζομένων, που ηχογραφήθηκαν στο Κέντρο Μελέτης Ύπνου του νοσοκομείου Ευαγγελισμός. Οι εξεταζόμενοι οποίοι υποβλήθηκαν παράλληλα σε πολυσωματογραφική εξέταση ώστε να είναι γνωστά μεγέθη ενδιαφέροντος (π.χ. AHI και συνολικός χρόνος ύπνου). Για τη δεύτερη μελέτη χρησιμοποιήθηκαν ηχητικά σήματα ροχαλητού από 11 εξεταζόμενους που και πάλι ηχογραφήθηκαν στο Κέντρο Μελέτης Ύπνου του Ευαγγελισμού, με παράλληλη πολυσωματογραφική εξέταση.

Κεφάλαιο 2: Θεωρητικό Υπόβαθρο

Στο κεφάλαιο αυτό γίνεται μια εισαγωγή στα χρονοσυχνοτικά χαρακτηριστικά των ηχητικών σημάτων ροχαλητού που εξετάζονται στην διπλωματική εργασία καθώς και στις μεθόδους αποθορυβοποίησης και κατάτμησης των σημάτων αυτών. Αναλύεται αρχικά η αποθορυβοποίηση ακουστικών σημάτων με τεχνικές αφαίρεσης φάσματος (spectral subtraction). Ακολούθως αναλύεται μια μέθοδος κατάτμησης των πολύωρων ηχογραφήσεων σε επιμέρους ροχαλητά και τέλος αναλύονται ένα-ένα τα χρονοσυχνοτικά χαρακτηριστικά που εξάγει το πακέτο Snor-Lux.

2.1: Μέθοδοι φασματικής αφαίρεσης για αποθρομβοποίηση ηχητικών σημάτων

Οι αλγόριθμοι φασματικής αφαίρεσης είναι ιστορικά από τους πρώτους αλγορίθμους που προτάθηκαν για την μείωση του θορύβου σε ηχητικά σήματα. Η μέθοδος της φασματικής αφαίρεσης βασίζεται σε μια απλή αρχή. Υιοθετώντας προσθετικό θόρυβο, μπορούμε να λάβουμε την εκτίμηση του φάσματος θορύβου από το φάσμα του ενθόρυβου σήματος. Το φάσμα του θορύβου μπορεί να εκτιμηθεί, και να αναπροσαρμοστεί, κατά την διάρκεια όπου το σήμα απουσιάζει. Η υπόθεση που κάνουμε είναι ότι ο θόρυβος είναι στατικός ή μια πολύ αργά μεταβαλλόμενη διαδικασία και ότι το φάσμα του θορύβου δεν αλλάζει σημαντικά μεταξύ των περιόδων αναπροσαρμογής. Το βελτιωμένο σήμα προκύπτει από τον υπολογισμό του αντίστροφου διακριτού μετασχηματισμού Fourier (Inverse discrete Fourier transform - IDFT) του εκτιμώμενου φάσματος του σήματος χρησιμοποιώντας την φάση του ενθόρυβου σήματος (Εικόνα 2.1). Ο αλγόριθμος είναι υπολογιστικά απλός αφού περιλαμβάνει μόνον ένα κανονικό και ένα αντίστροφο μετασχηματισμό Fourier. Η απλή όμως διαδικασία αφαίρεσης έχει και τα μειονεκτήματά της. Η αφαιρετική διαδικασία πρέπει να γίνει προσεκτικά για να αποφευχθεί αλλοίωση της ομιλίας. Αν γίνει πολύ μεγάλη αφαίρεση, τότε μπορεί να χαθεί κάποια χρήσιμη πληροφορία του σήματος, ενώ αν γίνει μικρή αφαίρεση τότε μπορεί να μην γίνει καλή αποθρομβοποίηση.



Εικόνα 2.1: Διάγραμμα αποθρομβοποίησης με φασματική αφαίρεση. Το εκτιμώμενο φάσμα θορύβου αφαιρείται από το φάσμα ενός σήματος θορυβώδους ομιλίας. Το αποτέλεσμα συνδυάζεται με τη θορυβώδη φάση πριν τον υπολογισμό του IDFT για την παραγωγή της βελτιωμένης κυματομορφής ομιλίας.

Πηγή: <http://www.cslu.ogi.edu>

2.1.1: Ο αλγόριθμος φασματικής αφαίρεσης Boll

Ο αλγόριθμος Boll είναι ένας αλγόριθμος γραμμικής φασματικής αφαίρεσης ο οποίος μειώνει τον στατικό θόρυβο από σήματα ήχου αφαιρώντας τον φασματικό θόρυβο που υπολογίζεται κατά την διάρκεια απουσίας ομιλίας. Μπορούν να εφαρμοστούν συμπληρωματικές διεργασίες για τη μείωση του απομένοντα θορύβου μετά την αρχική αφαίρεση. Στη συνέχεια ο αλγόριθμος επανασυνθέτει μια κυματομορφή στο πεδίο του χρόνου και μπορεί να χρησιμοποιηθεί σαν προεπεξεργαστής στενής ζώνης σε συστήματα επικοινωνίας φωνής, συστήματα αναγνώρισης ομιλίας, συστήματα επικύρωσης ηχείων και συστήματα αποθορυβοποίησης [27].

2.1.2: Ο αλγόριθμος φασματικής αφαίρεσης Berouti

Στον αλγόριθμο του Berouti, ο οποίος είναι ένας μη γραμμικός αλγόριθμος φασματικής αφαίρεσης που χρησιμοποιεί τη μέθοδο της υπεραφαίρεσης, πραγματοποιείται αφαίρεση του φάσματος του θορύβου κατά έναν παράγοντα α , όπου ο α είναι ένας αριθμός μεγαλύτερος της μονάδας και ποικίλει ανάλογα με το τρέχον παράθυρο επεξεργασίας. Η δυσκολία αυτής της μεθόδου έγκειται στο γεγονός ότι πρέπει να δοθεί προσοχή ώστε τα φασματικά στοιχεία του σήματος ενδιαφέροντος να υπερβαίνουν ένα κατώτατο όριο, το φασματικό δάπεδο (spectral floor). Ο αλγόριθμος στη συνέχεια εκφράζει το φασματικό δάπεδο ως ένα ποσοστό β του αρχικού φάσματος ισχύος [28].

2.1.3: Ο αλγόριθμος πολυζωνικής φασματικής αφαίρεσης Kamath

Ο αλγόριθμος πολυζωνικής (multiband) φασματικής αφαίρεσης του Kamath διαφέρει από το μη γραμμικό αλγόριθμο στον τρόπο εκτίμησης των παραγόντων υπεραφαίρεσης. Ο υπολογισμός του παράγοντα α γίνεται για κάθε ζώνη συχνοτήτων (frequency band) σε αντίθεση με το μη γραμμικό αλγόριθμο όπου ο υπολογισμός γίνεται για κάθε ομάδα συχνοτήτων. Στην πολυζωνική φασματική αφαίρεση αρχικά γίνεται διαίρεση του φάσματος του σήματος προς επεξεργασία σε N μη επικαλυπτόμενες ζώνες και ακολούθως εκτελείται φασματική αφαίρεση σε κάθε ζώνη ανεξάρτητα [29].

2.1.4: Ο αλγόριθμος φασματικής αφαίρεσης με φίλτρα Wiener

Στον αλγόριθμο αυτό χρησιμοποιούνται φίλτρα Wiener για την αποθρομβοποίηση του σήματος ενδιαφέροντος. Το μοντέλο του αλγορίθμου αυτού βασίζεται στη θεώρηση του Scarlat ότι:

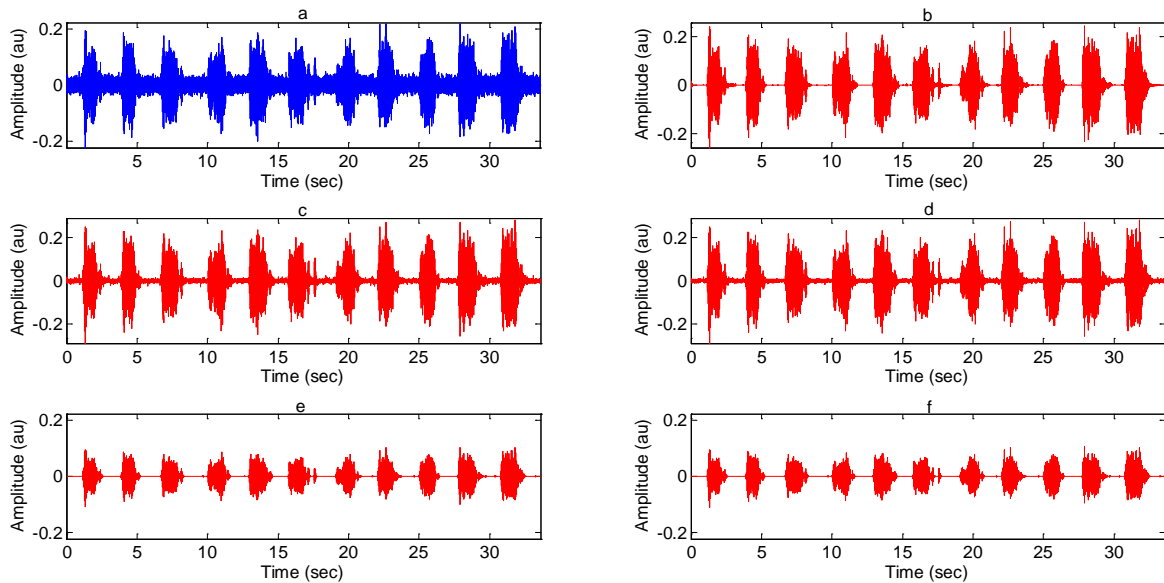
$$A \text{ Posteriori SNR} = A \text{ Priori SNR} + 1$$

Ο υπολογισμός του A Priori SNR γίνεται με τη μέθοδο της "Καθοδηγούμενης Απόφασης" για την εύρεση της συνάρτησης κέρδους που απαιτεί ο αλγόριθμος για την παραγωγή του αποθρομβοποιημένου σήματος [30].

Στην [31] προτείνονται δύο τεχνικές για περαιτέρω βελτίωση του σηματοθρομβικού λόγου. Η κλασική αποθρομβοποίηση με φίλτρα Wiener απαιτεί την μη αμερόληπτη εκτίμηση του a-priori SNR με αποτέλεσμα η μείωση του θορύβου να μη γίνεται αποδοτικά και να παρουσιάζονται ανεπιθύμητα φαινόμενα αντήχησης. Για να ξεπεραστεί το πρόβλημα αυτό προτείνεται η τεχνική μείωσης θορύβου σε δύο βήματα (Two-Step Noise Reduction - TSNR). Στο πρώτο βήμα γίνεται η εκτίμηση του a-priori SNR ενώ το δεύτερο βήμα απομακρύνει την μεροληψία της κλασικής προσέγγισης. Η δεύτερη τεχνική αφορά την βελτίωση της αρμονικής παραμόρφωσης στο αποθρομβοποιημένο σήμα, που οφείλεται στην αβεβαιότητα των εκτιμητών μικρών σηματοθρομβικών λόγων. Η αρμονική παραμόρφωση και η αβεβαιότητα που την προκαλεί είναι κυρίως αποτέλεσμα της δυσκολίας εκτίμησης της πυκνότητας φάσματος σε ηχογραφήσεις από μόνο ένα μικρόφωνο. Η τεχνική, που ονομάζεται μείωση θορύβου με αρμονική αναπαραγωγή (Harmonic Regeneration Noise Reduction - HRNR), χρησιμοποιεί μη γραμμικές μεθόδους για την αναπαραγωγή των υποβαθμισμένων αρμονικών του σήματος. Το προκύπτον σήμα χρησιμοποιείται στη συνέχεια για την βελτίωση της εκτίμησης του a-priori SNR, διατηρώντας τις αρμονικές του σήματος.

2.1.5: Επισκόπηση μεθόδων αποθρομβοποίησης

Στην Εικόνα 2.2 παραθέτονται γραφικές παραστάσεις των μεθόδων αποθρομβοποίησης με φασματική αφαίρεση που αναφέρονται στα κεφάλαια 2.1.1 -2.2.4.



Εικόνα 2.2: Αρχικό σήμα και οι αποθουροποιημένες εκδόσεις του

a: Αρχικό Σήμα προς αποθουροποίηση (σήμα ροχαλητών μήκους 33 δευτερολέπτων)

b: Αποτέλεσμα αποθουροποίησης με τη μέθοδο BOLL

c: Αποτέλεσμα αποθουροποίησης με τη μέθοδο Berouti

d: Αποτέλεσμα αποθουροποίησης με τη μέθοδο Kamath

e: Αποτέλεσμα αποθουροποίησης με τη μέθοδο φίλτρων Wiener - TSNR

f: Αποτέλεσμα αποθουροποίησης με τη μέθοδο φίλτρων Wiener - HRNR

2.2: Κατάτμηση ηχητικών σημάτων

Η αυτόματη κατάτμηση ενός συνεχούς σήματος ήχου σύμφωνα με την περιέχουσα πληροφορία του είναι μια βασική διεργασία σε κάθε σύστημα επεξεργασίας ήχου. Η χειροκίνητη κατάτμηση ενός σήματος είναι εξαιρετικά χρονοβόρα και δύσκολη διαδικασία ενώ είναι υποκειμενική αφού εξαρτάται από την οπτικοακουστική αντίληψη του ατόμου που την εκτελεί. Η αυτόματη κατάτμηση δεν είναι αλάνθαστη, όμως είναι εγγενώς συνεπής και έχει το πλεονέκτημα ότι τα αποτελέσματα θα είναι ακριβώς τα ίδια εάν ξαναχρησιμοποιηθεί ο αλγόριθμος κατάτμησης στα ίδια δεδομένα. Ο αλγόριθμος αυτόματης κατάτμησης που χρησιμοποιείται στην διπλωματική εργασία χρησιμοποιεί τα μεγέθη της ενέργειας σήματος (Signal Energy) και του φασματικού κεντροειδούς (Spectral Centroid) για την κατάτμηση των αποθουροποιημένων σημάτων ροχαλητού που δίνονται ως είσοδος [32].

2.2.1: Ενέργεια Σήματος

Η ενέργεια σήματος ορίζεται ως εξής: Έστω ότι έχουμε ένα διακριτό σήμα ήχου χωρισμένο σε i παράθυρα μήκους N δειγμάτων το καθένα:

$$x_i(n), \quad n = 1, \dots, N$$

Για κάθε παράθυρο i υπολογίζεται η ενέργεια σύμφωνα με την εξίσωση:

$$E(i) = \frac{1}{N} \sum_{n=1}^N |x_i(n)|^2$$

Το απλό αυτό μέγεθος μπορεί να χρησιμοποιηθεί για την ανίχνευση τμημάτων ησυχίας σε σήματα ήχου αλλά και για την διάκριση μεταξύ ηχητικών κατηγοριών [32].

2.2.2: Φασματικό Κεντροειδές

Το φασματικό κεντροειδές ορίζεται ως το κέντρο βάρους του φάσματος ενός σήματος. Έστω ότι έχουμε και πάλι ένα διακριτό σήμα χωρισμένο σε i παράθυρα μήκους N δειγμάτων το καθένα. Το φασματικό κεντροειδές του είναι:

$$C_i = \frac{\sum_{k=1}^N (k+1) X_i(k)}{\sum_{k=1}^N X_i(k)}$$

όπου $X_i(k)$ είναι οι συντελεστές του διακριτού μετασχηματισμού Fourier (DFT - Discrete Fourier Transform) του παραθύρου επεξεργασίας i .

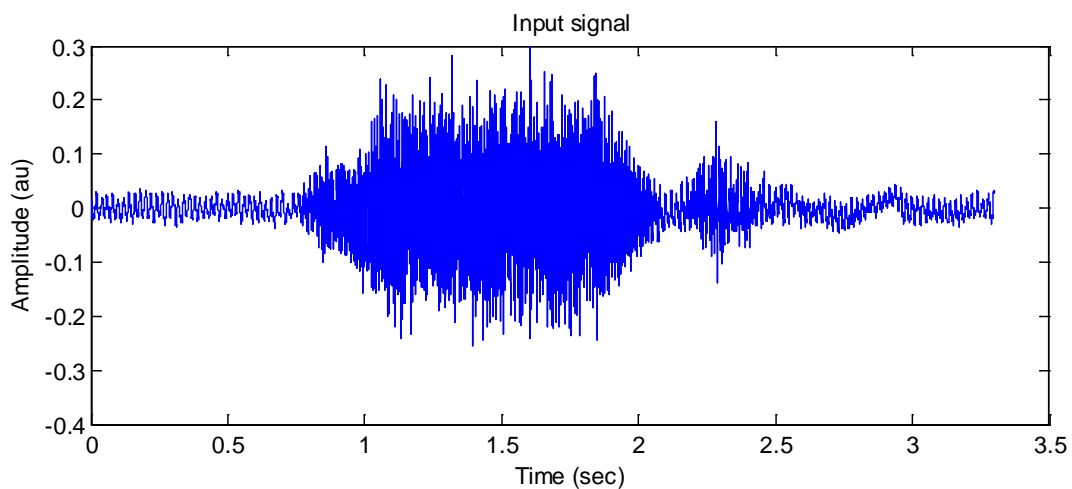
Το φασματικό κεντροειδές αποτελεί μέτρηση της φασματικής θέσης, με ψηλές τιμές να αντιστοιχούν σε "λαμπρότερους" ήχους. Υπάρχουν ενδείξεις ότι η αλληλουχία των φασματικών κέντρων βάρους μεταβάλλεται έντονα για διαφορετικά τμήματα ήχου [32].

2.3: Χρονοσυχνοτικά χαρακτηριστικά ηχητικών σημάτων

Στην ενότητα αυτή αναλύονται τα χρονικά και συχνοτικά χαρακτηριστικά σημάτων ροχαλητού και γενικά ήχου, που υπολογίζονται στο πακέτο Snor-Lux. Ιδιαίτερη έμφαση δίνεται στα μεγέθη που χρησιμοποιήθηκαν στο ερευνητικό μέρος. Στο Παράρτημα Β φαίνονται γραφικές παραστάσεις που λήφθηκαν από το Snor-Lux, όλων των μεγεθών που αναλύονται.

2.3.1: Αναπαράσταση σήματος στο πεδίο του χρόνου

Η πιο συνηθισμένη και απλούστερη μέθοδος αναπαράστασης ενός μονοδιάστατου σήματος, όπως είναι και τα σήματα ήχου, είναι ένα γράφημα που απεικονίζει τη μεταβολή του σήματος στο πεδίο του χρόνου. Το γράφημα μπορεί να απεικονίζει την μεταβολή της έντασης, πυκνότητας ή οποιουδήποτε άλλου μετρούμενου μεγέθους του σήματος με το χρόνο (Εικόνα 2.3). Στα σήματα ήχου απεικονίζεται η μεταβολή της πυκνότητας του αέρα στο χρόνο, ενώ το πλάτος της κυματομορφής αποτελεί την διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής πίεσης ενός ήχου [33].



Εικόνα 2.3: Αναπαράσταση σήματος ροχαλητού στο πεδίο του χρόνου

2.3.2: Αναπαράσταση σήματος στο πεδίο της συχνότητας

Μια άλλη συνηθισμένη μορφή αναπαράστασης ενός σήματος είναι η αναπαράσταση στο πεδίο της συχνότητας - φάσμα (frequency domain ή spectrum) και απεικονίζει το πλάτος των συχνοτήτων που αποτελούν ένα σήμα. Το φάσμα ενός σήματος μπορεί να εξαχθεί μέσω του μετασχηματισμού Fourier, με τις τιμές που προκύπτουν να απεικονίζονται σε γράφημα πλάτους και φάσης έναντι της συχνότητας. Κάθε σήμα που μπορεί να αναπαρασταθεί ως μεταβολή του πλάτους σε σχέση με το χρόνο, έχει αντίστοιχο (δυϊκό) φάσμα. Η αναπαράσταση στο πεδίο της συχνότητας χρησιμοποιείται εκτενώς αφού για αρκετά σήματα γίνονται ευκολότερα αντιληπτές οι εσωτερικές διεργασίες που χαρακτηρίζουν το σήμα στο πεδίο αυτό [33].

Ο μετασχηματισμός Fourier (FT - *Fourier Transform*) για μη περιοδικές συναρτήσεις στο συνεχές χρόνο περιέχει άπειρο εύρος συχνοτήτων και για θεωρούμενο μονοδιάστατο σήμα $x(t)$ ορίζεται ως εξής:

Κεφάλαιο 2

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\omega) e^{j\omega t} d\omega$$

Αντίστοιχα, ο αντίστροφος μετασχηματισμός ορίζεται:

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt$$

όπου το ω θεωρείται συνεχής μεταβλητή συχνότητας και ισχύει

$$\omega = 2\pi t$$

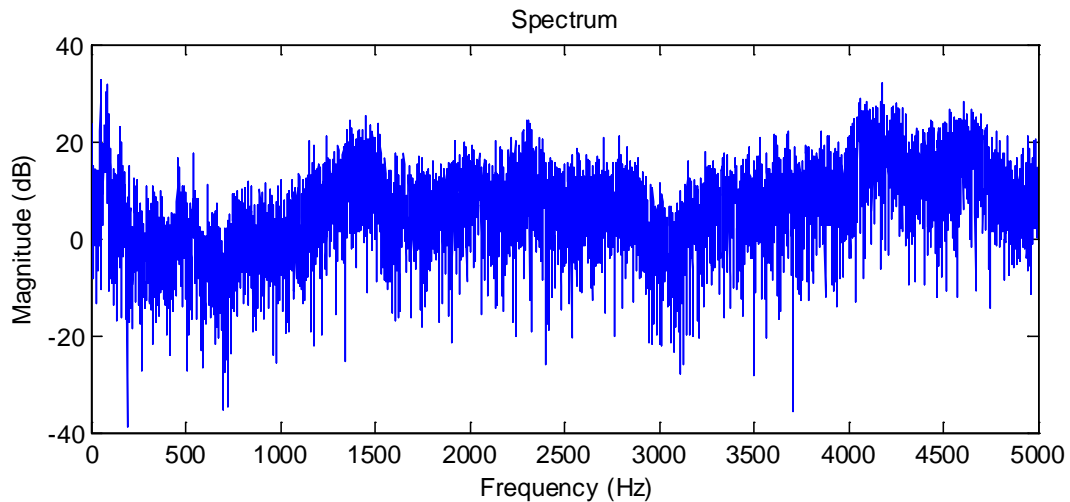
Στην διπλωματική εργασία γίνεται επεξεργασία διακριτών σημάτων που προκύπτουν από τη δειγματοληψία αναλογικού σήματος που ηχογραφείται στο εργαστήριο ύπνου. Ο ευθύς και αντίστροφος διακριτός μετασχηματισμός Fourier εφαρμόζονται σε πεπερασμένα διακριτά σήματα και ορίζονται αντίστοιχα:

$$x[n] = \frac{1}{2\pi} \int_{2\pi} X(e^{j\omega}) e^{j\omega n} d\omega \text{ και}$$

$$X(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} x[n] e^{-j\omega n}$$

όπου $x[n]$ το πεπερασμένο, διακριτό σήμα στο πεδίο του χρόνου, $X(e^{j\omega})$ ο διακριτός μετασχηματισμός Fourier του σήματος $x[n]$, ω η συνεχής μεταβλητή συχνότητας και n οι διακριτές μονάδες χρόνου. Σημειώνεται ότι η ολοκλήρωση στον ευθύ μετασχηματισμό γίνεται σε οποιοδήποτε διάστημα μήκους 2π λόγω του ότι το σήμα $X(e^{j\omega})$ είναι περιοδικό με περίοδο 2π [33].

Η μετάβαση στο πεδίο της συχνότητας γίνεται με τη χρήση του γρήγορου μετασχηματισμού Fourier (FFT - fast Fourier transform) ο οποίος είναι ένας αποδοτικός αλγόριθμος υπολογισμού του διακριτού μετασχηματισμού Fourier [34]. Παράδειγμα του φάσματος ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.4.



Εικόνα 2.4: Φάσμα σήματος ροχαλητού σε dB

2.3.3: Ενέργεια Βραχέως Χρόνου

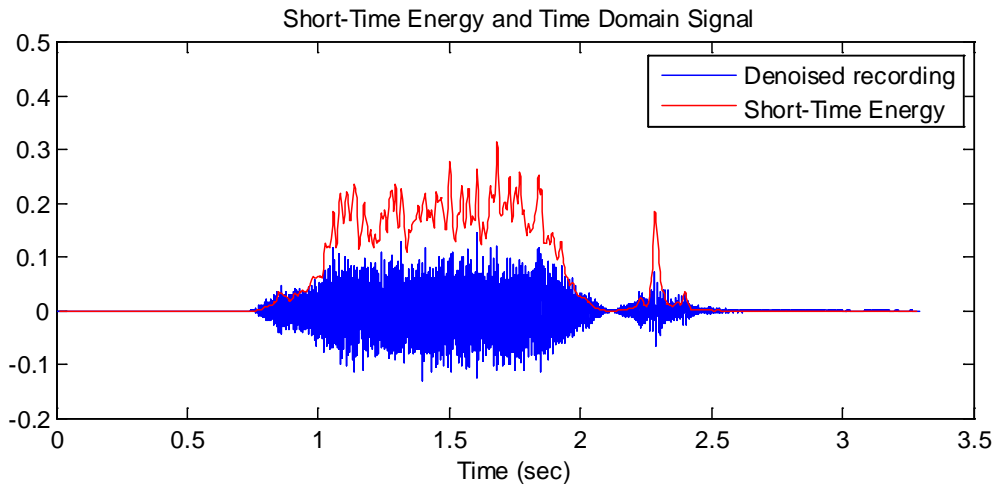
Η ενέργεια βραχέως χρόνου (STE - Short Time Energy) είναι η ενέργεια που περιέχει ένα σήμα, υπολογισμένη ανά μικρά, παραθυρομένα τμήματα του αρχικού σήματος πληροφορίας. Η μέτρηση της STE ενός ηχητικού σήματος μπορεί να χρησιμοποιηθεί για την διάκριση μεταξύ έμφωνης και άφωνης ομιλίας καθώς και για την ανίχνευση της μετάβασης από τη μία στην άλλη κατηγορία, αφού η ενέργεια της έμφωνης ομιλίας είναι πολύ μεγαλύτερη από αυτή της άφωνης. Η STE ενός δειγματοληπτημένου σήματος ορίζεται:

$$E_{\hat{n}} = \sum_{m=-\infty}^{+\infty} (x[m]w[\hat{n} - m])^2$$

όπου $x[m]$ το διακριτό σήμα ήχου και $w[m]$ το παράθυρο που χρησιμοποιείται.

Η επιλογή του μήκους και του τύπου παραθύρου που θα χρησιμοποιηθεί εξαρτάται από τον ήχο που θα μελετηθεί και απαιτεί συμβιβασμούς. Τα συνηθισμένα μήκη παραθύρου που χρησιμοποιούνται για σήματα ήχου είναι μήκους 20-30 ms [35].

Ένα παράδειγμα ενέργειας βραχέως χρόνου ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.5.



Εικόνα 2.5: Ενέργεια βραχέως χρόνου σήματος ροχαλητού

2.3.4: Ρυθμός διελεύσεων γύρω από το μηδέν

Ο ρυθμός διελεύσεων γύρω από το μηδέν (ZCR - Zero Crossing Rate) υποδεικνύει τη συχνότητα εναλλαγής πρόσημου του πλάτους ενός σήματος. Το ZCR είναι ένδειξη της μέσης συχνότητας ενός σήματος, και ορίζεται:

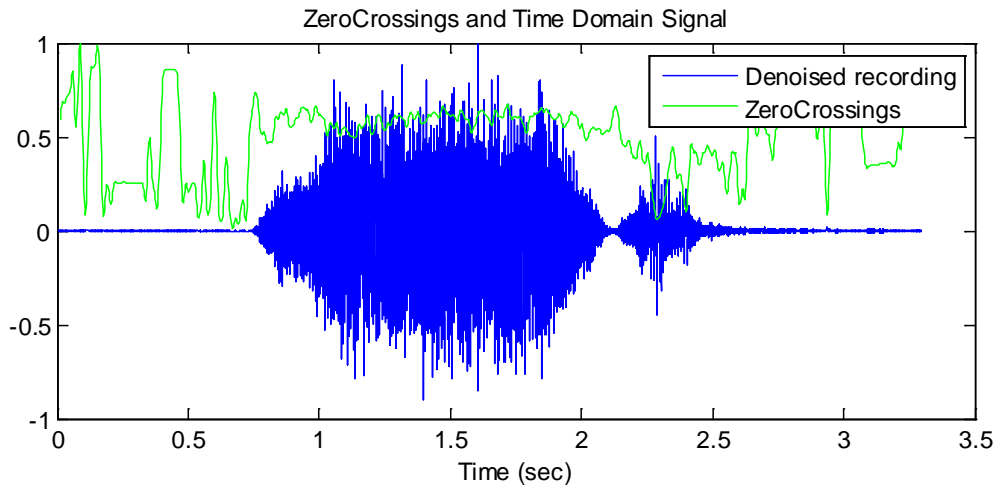
$$ZCR_{\hat{n}} = \sum_{m=-\infty}^{+\infty} \frac{1}{2} |sgn\{x[m]\} - sgn\{x[m-1]\}| w[\hat{n} - m]$$

όπου $sgn\{\}$ η συνάρτηση πρόσημου:

$$sgn\{x\} = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

$x[m]$ το διακριτό σήμα ήχου και $w[m]$ το παράθυρο που επιλέγεται για επεξεργασία.

Σε μαθηματικούς όρους ένα σημείο διέλευσης από το μηδέν είναι ένα σημείο όπου το πρόσημο μιας συνάρτησης μεταβάλλεται (από θετικό σε αρνητικό ή αντίστροφα) και αναπαριστάται με την διέλευση από τον άξονα στη γραφική παράσταση της συνάρτησης. Η μέτρηση των σημείων διέλευσης από το μηδέν χρησιμοποιείται ευρέως σε συστήματα επεξεργασίας ήχου για την εκτίμηση της θεμελιώδους συχνότητας ενός σήματος [35]. Ένα παράδειγμα του ρυθμού διελεύσεων γύρω από το μηδέν ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.6.



Εικόνα 2.6: ZCR σήματος ροχαλητού

2.3.5: Ένταση σήματος

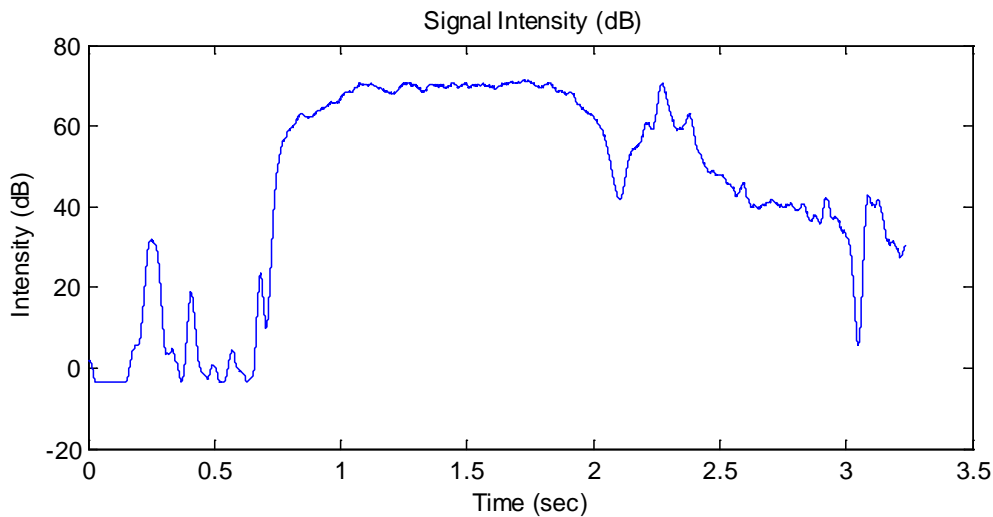
Η ένταση ενός ηχητικού σήματος σε decibel (dB) εκφράζει τη διαφορά στάθμης της πυκνότητας του αέρα. Η κλίμακα dB που χρησιμοποιείται αποτελεί μια λογαριθμική κλίμακα με δεκαδική υποδιαίρεση. Η ένταση ενός σήματος σε dB ορίζεται ως το δεκαπλάσιο του δεκαδικού λογαρίθμου του λόγου της έντασης του ήχου προς την ένταση αναφοράς του κατωφλίου ακουστότητας. Η ένταση του κατωφλίου ακουστότητας έχει οριστεί [36]:

$$I_0 = 10^{-12} \text{ W/m}^2$$

ενώ η έκφραση του ήχου έντασης σε dB δίνεται από τη σχέση:

$$I_{dB} = 10 \log_{10} \left(\frac{I_1}{I_0} \right)$$

Ένα παράδειγμα της έντασης σε dB ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.7.



Εικόνα 2.7: Ένταση σε dB σήματος ροχαλητού

2.3.6: Διφάσμα σήματος

Το διφάσμα (bispectrum) εμπίπτει στον τομέα της στατιστικής ανάλυσης στα μαθηματικά και αποτελεί ένα στατιστικό εργαλείο που χρησιμοποιείται για την εύρεση μη γραμμικών αλληλεπιδράσεων. Ο μετασχηματισμός Fourier της ροπής δεύτερης τάξης, της συνάρτησης αυτοσυσχέτισης, αποτελεί το γνωστό φάσμα ισχύος (power spectrum). Αντίστοιχα ο μετασχηματισμός της ροπής τρίτης τάξης $C_3(t_1, t_2)$, ονομάζεται διφάσμα (bispectrum) ή πυκνότητα διφάσματος (bispectral densities). Το θεώρημα συνέλιξης:

$$\mathcal{F}[f * g] = \mathcal{F}[f] * \mathcal{F}[g]$$

όπου ο τελεστής \mathcal{F} εφαρμόζει το μετασχηματισμό Fourier, επιτρέπει το γρήγορο υπολογισμό του διφάσματος ενός σήματος [37].

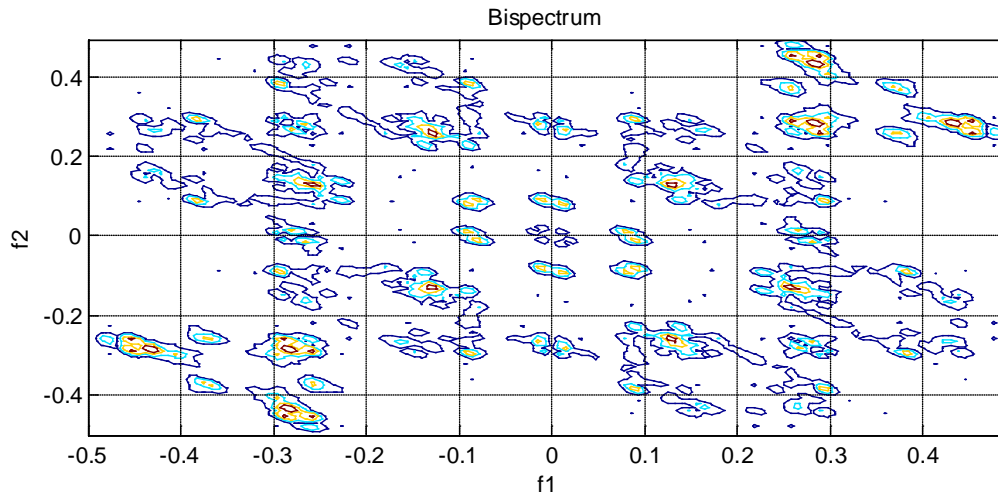
Το διφάσμα εμπίπτει στην κατηγορία των φασμάτων υψηλότερης τάξης ή αλλιώς πολυφασμάτων και παρέχει συμπληρωματικές πληροφορίες για το σήμα αφού περιέχει πληροφορία και για τη φάση του σήματος. Το πολυφάσμα τρίτης τάξης (διφάσμα) είναι το ευκολότερο στον υπολογισμό και ως εκ τούτου το πιο δημοφιλές. Το διφάσμα ορίζεται ως εξής:

$$B(\omega_1, \omega_2) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} C_3(m, n) e^{-j(\omega_1 m + \omega_2 n)}$$

όπου $C_3(m, n)$ είναι η ροπή τρίτης τάξης που ορίζεται:

$$C_3(m, n) = E\{x(k)x(k+m)x(k+n)\} \quad [17]$$

Ένα παράδειγμα του διφάσματος ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.8.



Εικόνα 2.8: Διφάσμα σήματος ροχαλητού

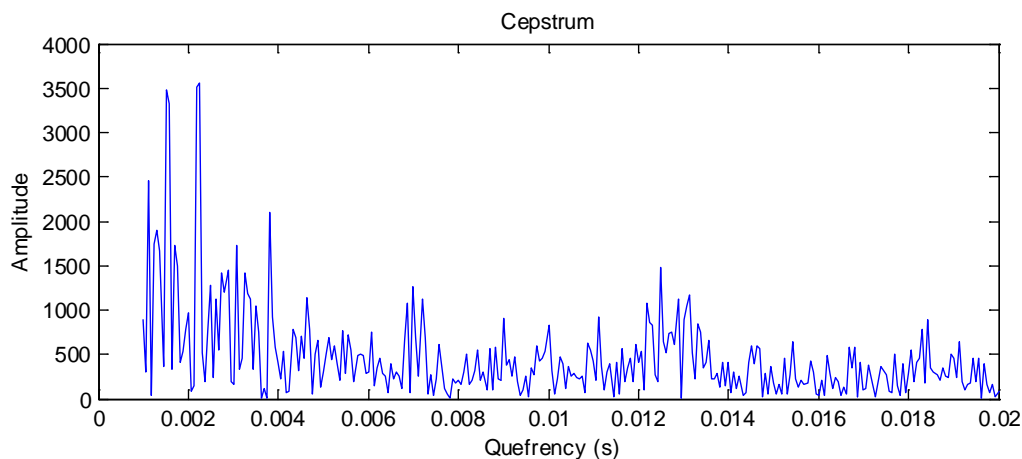
2.3.7: Σάφμα

Το σάφμα (cepstrum) προκύπτει από τον αντίστροφο μετασχηματισμό Fourier (IFT - Inverse Fourier Transform) του λογαρίθμου ενός σήματος. Υπάρχουν πολλά είδη σάφματος όπως το μιγαδικό, το πραγματικό σάφμα και το σάφμα φάσης, όμως το πιο δημοφιλές είναι το σάφμα ισχύος (*Power Cepstrum*) το οποίο βρίσκει εφαρμογές στην ανάλυση ανθρώπινης ομιλίας. Ο όρος σάφμα προκύπτει από την αναστροφή των τριών πρώτων γραμμάτων της λέξης φάσμα (αντίστοιχα spectrum - cepstrum αναστροφή των τεσσάρων πρώτων γραμμάτων) [38].

Το σάφμα δίνει πληροφορίες σχετικά με το ρυθμό αλλαγής σε διαφορετικές ζώνες συχνοτήτων και χρησιμοποιείται ευρέως στην ομομορφική επεξεργασία σήματος. Η ανεξάρτητη μεταβλητή ενός γραφήματος σάφματος ονομάζεται quefrequency. Η quefrequency είναι ένα μέγεθος χρόνου έχει όμως διαφορετική σημασία από την έννοια σε σήματα στο πεδίο του χρόνου. Για παράδειγμα εάν ένα δειγματοληπτημένο ηχητικό σήμα, με συχνότητα δειγματοληψίας 44100 Hz , έχει μια μεγάλη κορυφή στο σάφμα, με quefrequency 100 δείγματα, τότε η κορυφή δείχνει την ύπαρξη ενός τόνου με συχνότητα $44100/100 = 441 \text{ Hz}$ [39]. Ο μαθηματικός ορισμός του σάφματος ισχύος φαίνεται παρακάτω:

$$C_p = |\mathcal{F}\{\log(|\mathcal{F}\{f(t)\}|^2)\}|^2$$

Ένα παράδειγμα του σάφματος ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.9.



Εικόνα 2.9: Σάφμα σήματος ροχαλητού

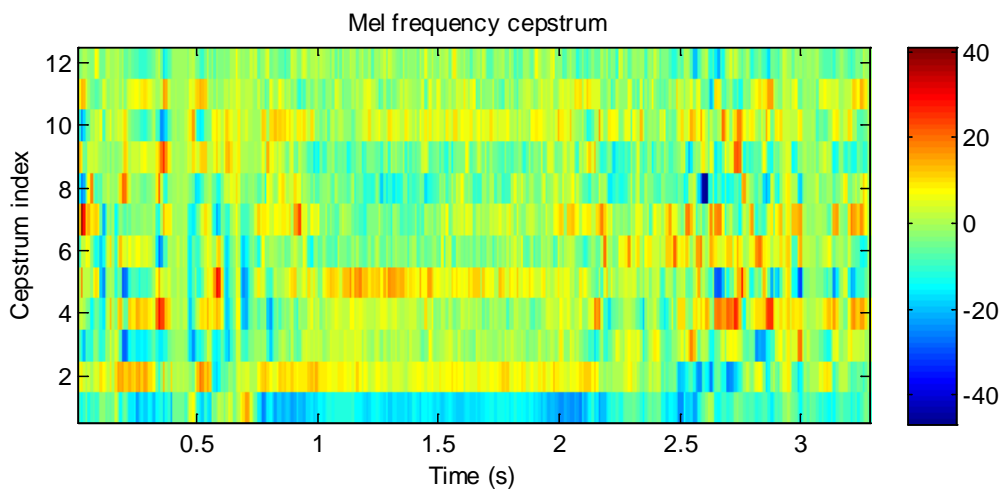
2.3.8: Σάφμα Mel συχνοτήτων

Το σάφμα συχνοτήτων Mel (MFC - Mel-Frequency Cepstrum) είναι μια αναπαράσταση του βραχέως φάσματος ισχύος ενός σήματος ήχου. Βασίζεται σε ένα γραμμικό μετασχηματισμό συνημίτονου του λογαριθμικού φάσματος ισχύος του σήματος σε μια μη γραμμική κλίμακα συχνότητας, την κλίμακα mel. Οι σαφματικοί συντελεστές mel συχνότητας (MFCC's Mel-Frequency cepstral coefficients) είναι συντελεστές οι οποίοι αθροιστικά δημιουργούν ένα MFC και εξάγονται από μια μορφή σαφματικής αναπαράστασης ενός σήματος ήχου, ένα μη γραμμικό "φάσμα φάσματος". Η διαφορά που έχει από το σάφμα είναι ότι στο MFC οι ζώνες συχνότητας είναι ομοιόμορφα κατανομημένες στην κλίμακα mel, η οποία προσεγγίζει καλύτερα την απόκριση του ανθρώπινου ακουστικού συστήματος [40].

Η διαδικασία λήψης των MFCC's ακολουθεί 5 βήματα:

- Λήψη του FT του παραθυρομένου σήματος
- Χαρτογράφηση των τιμών του ληφθέντος φάσματος στην κλίμακα mel με τη χρήση τριγωνικών επικαλυπτόμενων παραθύρων
- Λογαρίθμηση των τιμών σε κάθε συχνότητα mel
- Εφαρμογή του διακριτού μετασχηματισμού συνημίτονου σε κάθε λογαριθμημένη τιμή mel
- Οι MFCC's είναι τα πλάτη του προκύπτοντος φάσματος

Οι MFCC's χρησιμοποιούνται ως χαρακτηριστικά σε συνδυασμό συνήθως με αλγόριθμους κρυμμένων μοντέλων Markov (HMM - *Hidden Markov Models*) σε συστήματα αναγνώρισης ομιλίας που μπορούν να αναγνωρίζουν αυτόματα αριθμούς και σε συστήματα αναγνώρισης ομιλητή. Παράδειγμα του σάφματος mel συχνοτήτων φαίνεται στην Εικόνα 2.10.



Εικόνα 2.10: Σάφμα mel συχνοτήτων σήματος ροχαλητού

2.3.9: Αυτοσυσχέτιση σήματος

Η αυτοσυσχέτιση (autocorrelation) ενός σήματος είναι η συσχέτιση ενός σήματος με τον εαυτό του και δείχνει την ομοιότητα μεταξύ παρατηρήσεων ως μια συνάρτηση των χρονικών διαφορών μεταξύ τους. Χρησιμοποιείται στα μαθηματικά και την επεξεργασία σήματος για εύρεση επαναλαμβανόμενων προτύπων, όπως την εύρεση κάποιου περιοδικού σήματος που περιέχει θόρυβο και την εύρεση της θεμελιώδους συχνότητας ενός σήματος [33].

Δεδομένου ενός σήματος $x(t)$ η συνεχής αυτοσυσχέτιση του $R_{ff}(\tau)$, ορίζεται ως το συνεχές ολοκλήρωμα ετεροσυσχέτισης του $x(t)$ με τον εαυτό του, στη χρονική στιγμή τ :

$$R_{ff}(\tau) = (x(t) * \bar{x}(-\tau))(\tau) = \int_{-\infty}^{+\infty} x(t + \tau) \bar{x}(t) dt = \int_{-\infty}^{+\infty} x(t) \bar{x}(t - \tau) dt$$

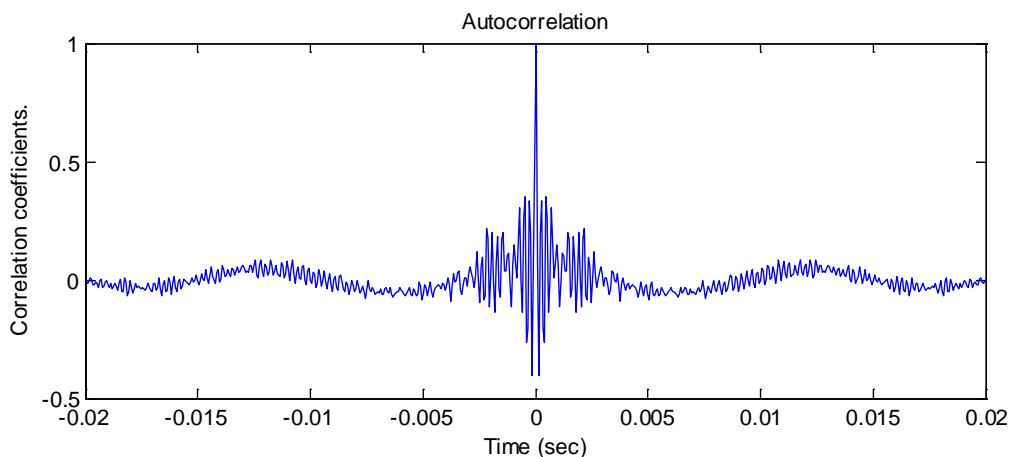
όπου \bar{x} το συζυγές σήμα του x και ο τελεστής $*$ αναπαριστά συνέλιξη.

Κεφάλαιο 2

Για διακριτά σήματα ορίζεται:

$$R_{ff}(j) = \sum_n x_n \bar{x}_{n-j}$$

Η αυτοσυσχέτιση ενός σήματος ροχαλητού φαίνεται στην Εικόνα 2.11.



Εικόνα 2.11: Αυτοσυσχέτιση σήματος ροχαλητού

2.3.10: Συνεχής μετασχηματισμός κυματιδίων

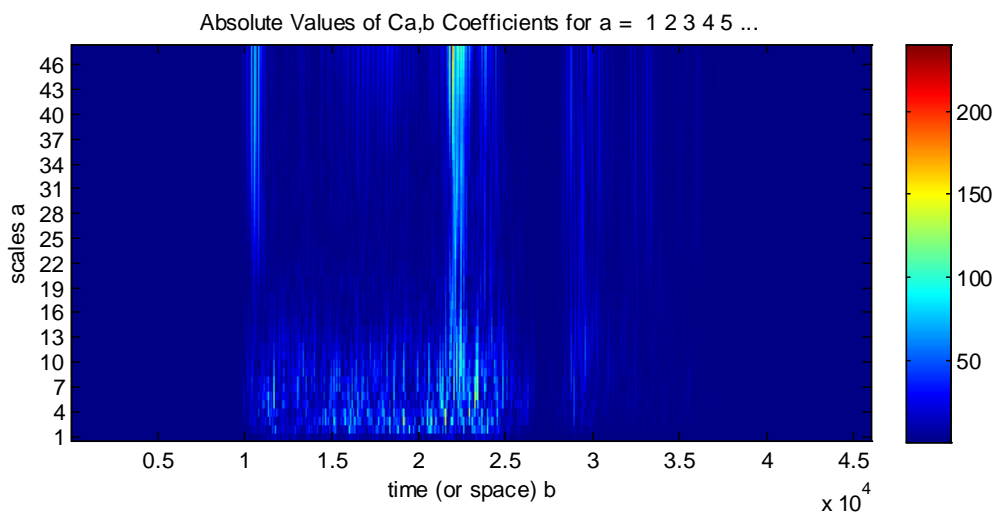
Ο συνεχής μετασχηματισμός κυματιδίων (CWT - *Continuous Wavelet Transform*) χρησιμοποιείται για την κατάτμηση μιας συνάρτησης συνεχούς χρόνου σε κυματίδια (αντίστοιχα με το μετασχηματισμό Fourier ο οποίος αναλύει μια συνάρτηση σε ημίτονα). Το πλεονέκτημα που έχει ο CWT έναντι του FT είναι η υψηλή διακριτική ικανότητα τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας. Ο CWT ενός συνεχούς, διπλά ολοκληρώσιμου σήματος $x(t)$, σε κλίμακα $a > 0$ και με μετατόπιση b εκφράζεται από το ολοκλήρωμα:

$$X_w(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) \psi^* \left(t - \frac{b}{a} \right) dt$$

όπου $\psi(t)$ είναι η μητρική συνάρτηση κυματιδίου, της οποίας ο σκοπός είναι η παροχή μιας πηγαίας συνάρτησης για την παραγωγή των θυγατρικών κυματιδίων. Τα τελευταία αποτελούν απλώς μετατοπισμένες και κλιμακωμένες εκδοχές της μητρικής συνάρτησης.

Οι εφαρμογές του CWT εκτείνονται σε πολλούς τομείς. Η πιο δημοφιλής εφαρμογή αφορά την συμπίεση εικόνας λόγω της αυξημένης ποιότητας συμπίεσμνης εικόνας σε σχέση με άλλες τεχνικές συμπίεσης. Χρήσεις του CWT υπάρχουν επίσης στην αναγνώριση προτύπων και στην επεξεργασία ήχου [41].

Οι συντελεστές του συνεχούς μετασχηματισμού κυματιδίων ενός σήματος ροχαλητού παρουσιάζονται στην Εικόνα 2.12.



Εικόνα 2.12: Συντελεστές συνεχούς μετασχηματισμού κυματιδίων σήματος ροχαλητού

2.3.11: Φασματογράφημα

Το φασματογράφημα (spectrogram) είναι μια χρονικά μεταβαλλόμενη φασματική αναπαράσταση που δείχνει τη μεταβολή της φασματικής πυκνότητας ενός σήματος με το χρόνο. Η πιο συνηθισμένη μορφή που έχει ένα φασματογράφημα είναι γραφήματος δύο διαστάσεων όπου ο οριζόντιος άξονας αντιπροσωπεύει το χρόνο και ο κάθετος άξονας τη συχνότητα [42].

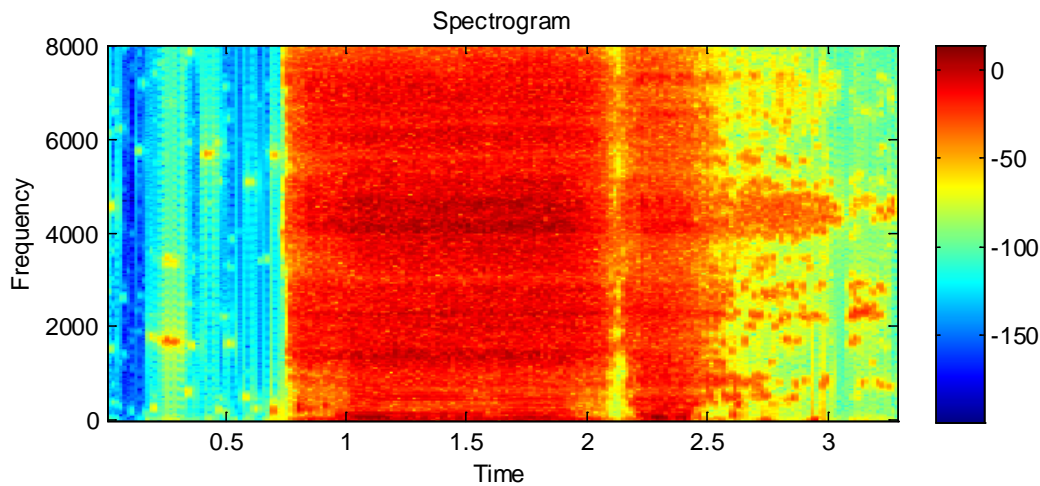
Ο υπολογισμός του φασματογραφήματος γίνεται συνήθως με τη βοήθεια του μετασχηματισμού Fourier βραχέως χρόνου (STFT - Short time Fourier transform). Ο STFT ενός διακριτού σήματος $x[n]$, υπολογίζεται από τη σχέση:

$$STFT\{x[n]\}(m, \omega) \equiv X(m, \omega) = \sum_{n=-\infty}^{+\infty} x[n]w[n - m]e^{-j\omega n}$$

όπου $w[n]$ μια συνάρτηση παραθύρου (συνήθως χρησιμοποιείται παράθυρο Hann ή Gaussian). Το φασματογράφημα με τη σειρά του ορίζεται ως το τετράγωνο του πλάτους του μετασχηματισμού Fourier βραχέως χρόνου:

$$spectrogram\{x(t)\}(\tau, \omega) \equiv |X(\tau, \omega)|^2$$

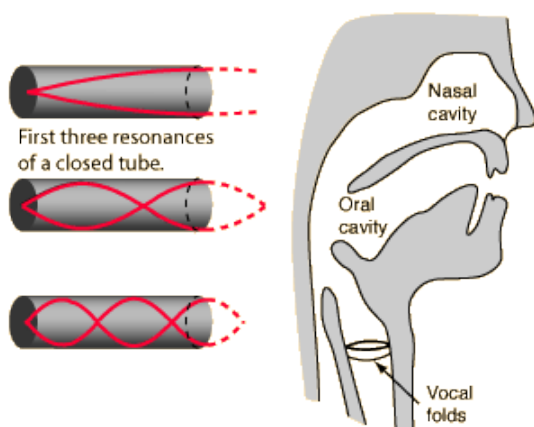
Στην Εικόνα 2.13 παρουσιάζεται το φασματογράφημα ενός σήματος ροχαλητού.



Εικόνα 2.13: Φασματογράφημα σήματος ροχαλητού

2.3.12: Συχνότητες φωνοσυντονισμού

Ο όρος συχνότητες φωνοσυντονισμού (formant frequencies) ορίστηκε από τον G. Fant το 1960 [43] ως οι φασματικές κορυφές του φάσματος ενός ακουστικού σήματος. Στα ηχητικά σήματα που προέρχονται από τον άνθρωπο οι FF καθορίζονται από τα χαρακτηριστικά του φωνητικού σωλήνα. Το εύρος των formants καθορίζεται από το συνολικό μήκος του φωνητικού σωλήνα από τη γλωττίδα μέχρι τα χείλη ενώ τα σημαντικά formants θεωρούνται οι τρεις πρώτες κορυφές της περιβάλλουσας του φάσματος (F1, F2, F3) (Εικόνα 2.14). Τυπικές τιμές των FF για ομιλίας είναι 500 Hz, 1500 Hz και 2500 Hz για άντρες ομιλητές και 600 Hz, 1800 Hz και 3000 Hz για γυναίκες.

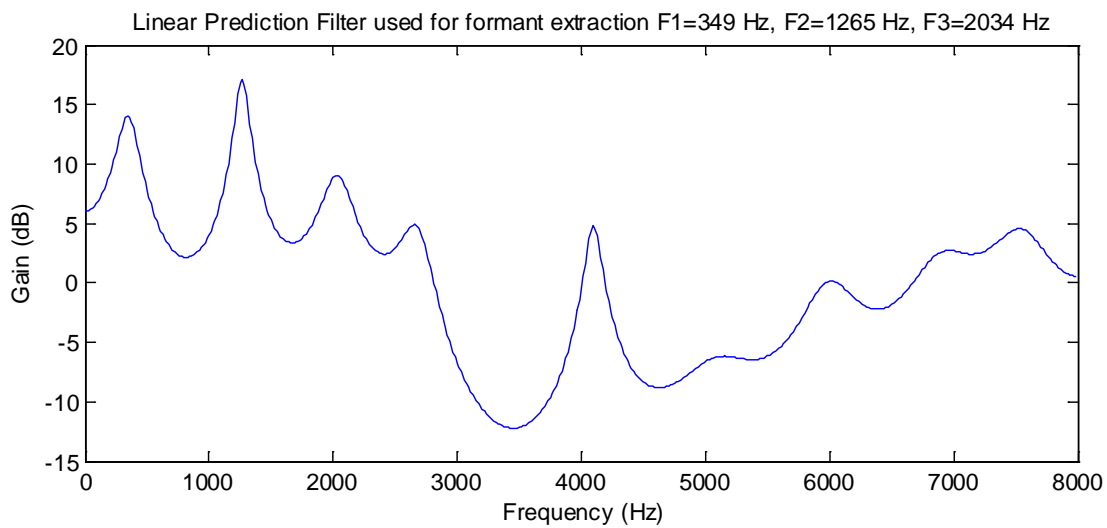


Εικόνα 2.14: Αντιστοιχία FF με τις τρεις πρώτες αρμονικές σήματος σε κλειστό σωλήνα

Για τον υπολογισμό των FF μπορούν να χρησιμοποιηθούν πολλές τεχνικές, επικρατεί όμως η τεχνική υπολογισμού μέσω των συντελεστών γραμμικής πρόβλεψης (Linear Prediction Coefficients - LPC). Στη συγκεκριμένη μέθοδο η εύρεση των formants αφορά τον καθορισμό των κορυφών συντονισμού από το φίλτρο συντελεστών που λαμβάνεται μέσω

της LPC ανάλυσης τμημάτων του ακουστικού σήματος. Αφού υπολογιστεί το πολυώνυμο πρόβλεψης $A(z)$, οι FF καθορίζονται είτε με εύρεση μεγίστων της καμπύλης απόκρισης του φίλτρου, είτε με επίλυση ριζών της εξίσωσης $A(z) = 0$. Σημειώνεται ότι κάθε ζευγάρι μιγαδικών ριζών χρησιμοποιείται για τον υπολογισμό της αντίστοιχης FF και του εύρους της [44].

Στην Εικόνα 2.15 φαίνεται το φίλτρο που χρησιμοποιείται για τον υπολογισμό των FF ενός σήματος ροχαλητού.



Εικόνα 2.15: Φίλτρο υπολογισμού FF σήματος ροχαλητού

Κεφάλαιο 3: Ανάπτυξη λογισμικού Snor-Lux

Στο κεφάλαιο αυτό περιγράφονται τα εργαλεία που χρησιμοποιήθηκαν για την ανάπτυξη του γραφικού διαδραστικού περιβάλλοντος Snor-Lux. Αρχικά εξηγείται η επιλογή του προγραμματισμού του Snor-Lux με το λογισμικό πακέτο MATLAB®. Στη συνέχεια γίνεται μια εισαγωγή στο εργαλείο GUIDE που χρησιμοποιήθηκε για την σχεδίαση των επιφανειών του Snor-Lux και στο εργαλείο DEPLOYTOOL που χρησιμοποιήθηκε για την μετατροπή του Snor-Lux σε αυτόνομη εφαρμογή. Αναφέρονται τέλος τα χαρακτηριστικά του Snor-Lux.

3.1: Επιλογή πλατφόρμας για την ανάπτυξη γραφικού διαδραστικού περιβάλλοντος

Για την ανάπτυξη του γραφικού διαδραστικού περιβάλλοντος (GUI - Graphical User Interface) Snor-Lux επιλέχθηκε η χρήση της πλατφόρμας της Mathworks, του MATLAB®. Οι λόγοι για την επιλογή του MATLAB® είναι η πληθώρα των ήδη υπαρχουσών μεθόδων ανάλυσης σημάτων, η ευρεία χρήση του από μηχανικούς που ασχολούνται με ανάλυση σημάτων και η ταχύτητα εκτέλεσης αριθμητικών πράξεων σε πίνακες μεγάλων διαστάσεων όπως οι πίνακες ακουστικών σημάτων ροχαλητών. Κύριο πλεονέκτημα του MATLAB® είναι το γεγονός ότι οι βασικές μαθηματικές λειτουργίες και τεχνικές επεξεργασίας σημάτων είναι ενσωματωμένες είτε στο βασικό πρόγραμμα είτε στα διάφορα συμπληρωματικά πακέτα (toolboxes). Η υποστήριξη από την εταιρία και την ακαδημαϊκή κοινότητα εξασφαλίζει την ποιότητα των λειτουργιών αυτών, ενώ η τεκμηρίωση τους είναι άψογη. Το MATLAB® υποστηρίζει ακόμα την εύκολη δημιουργία γραφικού διαδραστικού περιβάλλοντος καθιστώντας δυνατή την ανάπτυξη προγραμμάτων που να απευθύνονται και σε χρήστες με ελάχιστη γνώση MATLAB® αλλά και σε χρήστες που είναι σε θέση να υλοποιήσουν τις δικές τους τεχνικές και χρησιμοποιούν το πρόγραμμα για τις έτοιμες βασικές λειτουργίες.

3.2: Ανάπτυξη GUI στο MATLAB®

Με τον όρο γραφικό περιβάλλον αναφερόμαστε σε μια γραφική αναπαράσταση αποτελούμενη από συστατικά τα οποία δίνουν την δυνατότητα στο χρήστη να επιτελέσει διαδραστικούς χειρισμούς. Στο προγραμματιστικό περιβάλλον του MATLAB® η κατασκευή ενός διαδραστικού γραφικού περιβάλλοντος - GUI περιλαμβάνει το σχεδιασμό των διαδραστικών στοιχείων και τη σύνδεσή τους με συγκεκριμένες λειτουργίες που πραγματοποιούνται όταν ο χρήστης εκτελέσει τον εκάστοτε χειρισμό. Η διαδικασία της δημιουργίας του γραφικού περιβάλλοντος μπορεί να πραγματοποιηθεί είτε μέσω του βοηθητικού εργαλείου GUIDE (Graphical User Interface Design Environment) που προσφέρει το MATLAB®, είτε προγραμματιστικά με απευθείας συγγραφή του απαιτούμενου κώδικα. Για την υλοποίηση του Snor-Lux χρησιμοποιήθηκε το βοηθητικό εργαλείο GUIDE. Ο προγραμματισμός του GUI γίνεται με μια σειρά εντολών που προσφέρει το MATLAB® και οι οποίες ελέγχουν τη δημιουργία των διαδραστικών

στοιχείων, την εμφάνιση και τη θέση τους στην οθόνη, τη σύνδεσή τους με συγκεκριμένες λειτουργίες, τις πληροφορίες που παρουσιάζουν και άλλες ιδιότητες.

Κάθε GUI αποτελείται συνήθως από διάφορα παράθυρα, τα οποία είναι δυνατόν να καλούν το ένα το άλλο, να δέχονται δεδομένα από το χρήστη, να μεταβιβάζουν τα δεδομένα και να επιτελούν διάφορες λειτουργίες. Το GUIDE για να το πετύχει αυτό δημιουργεί για κάθε παράθυρο δύο αρχεία, το αρχείο fig και το αρχείο m. Το αρχείο fig είναι ουσιαστικά το παράθυρο όπου το MATLAB® αποθηκεύει τα στοιχεία ελέγχου και την ακριβή θέση τους και αφορά την εμφάνιση του παραθύρου. Το αρχείο m περιέχει τον πηγαίο κώδικα που ενσωματώνεται στα στοιχεία ελέγχου που περιέχει το παράθυρο (π.χ. κουμπιά και λίστες), ώστε αυτά να επιτελούν τις επιθυμητές λειτουργίες.

3.3: Μετατροπή GUI σε αυτόνομη εφαρμογή

Η ανάπτυξη ενός διαδραστικού προγράμματος σε MATLAB® δεν πρέπει να αποτελεί μειονέκτημα λόγω της απαίτησης ύπαρξης του MATLAB® εγκατεστημένου στον υπολογιστή των χρηστών. Για το λόγο αυτό έγινε μετατροπή του λογισμικού Snor-Lux σε αυτόνομη, ανεξάρτητη εφαρμογή. Για τη μετατροπή ενός προγράμματος MATLAB® σε αυτόνομη εφαρμογή χρησιμοποιείται ο μεταγλωττιστής Matlab (Matlab Compiler) ο οποίος μειώνει σημαντικά το χρόνο ανάπτυξης της εφαρμογής αφού εξαλείφει τη διαδικασία μεταγλώττισης του κώδικα MATLAB® σε C ή C++. Η δημιουργία του αυτόνομου πακέτου Snor-Lux έγινε με το βοηθητικό εργαλείο deploytool.

Το τελικό αρχείο που προκύπτει με τη χρήση του βοηθητικού εργαλείου deploytool είναι ένα συμπιεσμένο αρχείο (.zip) που περιέχει το εκτελέσιμο αρχείο του προγράμματος (.exe) και ένα αρχείο εγκατάστασης του Matlab Component Runtime (MCR). Ο χρήστης που δεν έχει εγκατεστημένο στον υπολογιστή του το MATLAB® μπορεί να εγκαταστήσει εύκολα το MCR και ακολούθως να τρέξει την εφαρμογή χωρίς να απαιτείται πλήρης εγκατάσταση του MATLAB®. Σημειώνεται ότι η εφαρμογή μπορεί να εκτελεστεί και σε περιβάλλον Windows και σε περιβάλλον Unix.

3.4: Χαρακτηριστικά του λογισμικού Snor-Lux

Το λογισμικό Snor-Lux αναπτύχθηκε όπως αναφέρθηκε με τη χρήση του βοηθητικού εργαλείου GUIDE. Περιλαμβάνει πέντε επιφάνειες - παράθυρα:

Κεφάλαιο 3

1. Επιφάνεια αρχικής οθόνης: Είναι η επιφάνεια που εμφανίζεται μόλις ξεκινήσει το πρόγραμμα Snor-Lux
2. Επιφάνεια φόρτωσης αρχείου και αποθρομβοποίησης: Είναι η επιφάνεια στην οποία επιλέγεται το αρχείο που θα επεξεργαστεί και η μέθοδος αποθρομβοποίησης
3. Επιφάνεια κατάτμησης σήματος: Είναι η επιφάνεια όπου γίνεται η κατάτμηση του σήματος
4. Επιφάνεια επεξεργασίας σήματος: Είναι η επιφάνεια όπου γίνεται η κυρίως επεξεργασία του αποθρομβοποιημένου σήματος και εξαγωγή χρονοσυχνοτικών χαρακτηριστικών
5. Επιφάνεια επικοινωνίας: Είναι η τελευταία επιφάνεια του Snor-Lux όπου παρέχονται στοιχεία επικοινωνίας

Στο Παράρτημα Α παρουσιάζεται το εγχειρίδιο λειτουργίας του λογισμικού Snor-Lux και παραθέτονται εικόνες των πέντε επιφανειών.

Το λογισμικό Snor-Lux λειτουργεί σε οθόνες ανάλυσης Full-HD (ανάλυση 1920 × 1080 pixels) για λόγους καλής ευκρίνειας και αφού οι πλείστες οθόνες σήμερα υποστηρίζουν την ανάλυση αυτή.

Στην έκδοση διανομής (distribution version) του λογισμικού Snor-Lux που επισυνάπτεται στο ψηφιακό δίσκο (CD) περιέχονται ο πηγαίος κώδικας του προγράμματος σε MATLAB® στο φάκελο "snor-lux matlab source code", το αυτόνομο πρόγραμμα "snorlux.exe" μαζί με το πρόγραμμα εγκατάστασης του MCR στο συμπιεσμένο αρχείο "snorlux-standalone.zip" και μερικά ακουστικά σήματα ροχαλητών στο φάκελο "sample data". Όλες οι εκδοχές του Snor-Lux συνοδεύονται από αρχεία που επεξηγούν τη διαδικασία εκτέλεσης τους, στα αγγλικά και στα ελληνικά (README-EN.txt και README-EL.txt αντίστοιχα).

Κεφάλαιο 4: Υλικό - Μέθοδοι

Στο κεφάλαιο αυτό περιγράφονται τα χαρακτηριστικά του συστήματος καταγραφής, η προεπεξεργασία στην οποία υποβλήθηκαν τα δεδομένα και τα χαρακτηριστικά του υλικού που χρησιμοποιήθηκε και η μεθοδολογία που ακολουθήθηκε για την εξαγωγή αποτελεσμάτων. Ως υλικό εννοούνται οι εξεταζόμενοι που προσήλθαν στο Κέντρο Μελέτης Ύπνου του νοσοκομείου "Ευαγγελισμός" και υποβλήθηκαν σε νυκτερινή μελέτη ύπνου που περιλαμβάνει πολυσωματογράφημα και παράλληλη ηχογράφηση των ροχαλητών τους. Το υλικό και η μεθοδολογία χωρίζονται σε δύο επιμέρους τμήματα. Το πρώτο περιγράφει τη μελέτη επίδρασης του συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών σε απνοϊκούς ασθενείς και τη διαδικασία εξαγωγής της μέσης και μέγιστης έντασης ροχαλητού (σε dB) και της συχνότητας. Το δεύτερο τμήμα περιγράφει αναλυτικά την διαδικασία εξαγωγής των συχνοτήτων φωνοσυντονισμού και των τριών ενδεικτικών συχνοτήτων του φάσματος (μέγιστη συχνότητα: $f_{90\%}$, μέση συχνότητα $f_{50\%}$ και συχνότητα κορυφής f_{peak}), και τις παραμέτρους ανάλυσης των ηχογραφήσεων ροχαλητών που προέρχονται από εξεταζόμενους με διαφορετικά ΑΗΙ. Στο τμήμα αυτό αναλύεται και η μέθοδος στατιστικής επεξεργασίας των αποτελεσμάτων με καμπύλες λειτουργικού χαρακτηριστικού δέκτη (ROC).

4.1: Σύστημα καταγραφής

Τα ακουστικά σήματα ροχαλητών που χρησιμοποιήθηκαν στην διπλωματική εργασία, έχουν όλα ηχογραφηθεί στο Κέντρο Μελέτης Ύπνου (ΚΜΥ) του νοσοκομείου "Ευαγγελισμός" υπό την επίβλεψη των Δρ. Ε. Βαγιάκη και Δρ. Ι. Κουτσοελάκη. Οι εξεταζόμενοι είναι εθελοντές που υποβάλλονται σε PSG για διάγνωση του OSAHS ή άλλων διαταραχών στον ύπνο και ταυτόχρονα ηχογραφούνται, πάντοτε με τη συναίνεση τους. Τα αποτελέσματα της μελέτης ύπνου περιλαμβάνουν τα μεγέθη που αναφέρονται στο υποκεφάλαιο 1.2 καθώς και άλλες χρήσιμες πληροφορίες όπως συνολικό χρόνο ύπνου, αριθμό αφυπνίσεων κ.λπ.

Για την ηχογράφηση των ροχαλητών ενσωματώθηκε σε ένα δωμάτιο μελέτης ύπνου του ΚΜΥ (Δωμάτιο 4), ένα υψηλής πιστότητας σύστημα ηχογράφησης και αποθήκευσης ήχου. Αποτελείται από δύο πυκνωτικά μικρόφωνα Sontronics STC-1 με απόκριση συχνότητας στη ζώνη 25 Hz - 20 kHz και αντίσταση μικρότερη ή ίση των 200 Ω, που τοποθετούνται περίπου 15 cm πάνω από το κεφάλι του εξεταζόμενου. Τα μικρόφωνα συνδέονται με προενισχυτή M-Audio Fast Track Ultra ο οποίος είναι ενωμένος απευθείας με επιτραπέζιο υπολογιστή (Intel Core2 Duo E8500 @3.16 GHz, 2GB RAM με λειτουργικό Windows XP - SP3) στο δωμάτιο ελέγχου. Η σύνδεση γίνεται μέσω διπλού, θωρακισμένου ομοαξονικού καλωδίου Sontronics Ultima για ελαχιστοποίηση των ηλεκτρομαγνητικών παρεμβολών. Για τη μελέτη ύπνου υπάρχει εγκατεστημένο σύστημα καταγραφής PSG, Embla S-7000. Ο εξοπλισμός τόσο για την καταγραφή του PSG όσο και για την καταγραφή των ακουστικών σημάτων ροχαλητών βρίσκεται στο δωμάτιο ελέγχου ενώ το δωμάτιο εξέτασης ικανοποιεί όλες τις απαραίτητες προδιαγραφές σύμφωνα με τα τελευταία πρότυπα [13]. Η επεξεργασία των δεδομένων έγινε με υπολογιστή Intel Core i7 Q740 @1.73 GHz, 6GB RAM και λειτουργικό Windows 7 Professional - 64 bit.

4.2: Προεπεξεργασία δεδομένων

Τα δεδομένα των ηχογραφήσεων ροχαλητών υποβλήθηκαν σε προεπεξεργασία πριν την εξαγωγή των απαιτούμενων χαρακτηριστικών κάθε μελέτης. Αναφέρεται ότι τα δεδομένα ηχογραφήθηκαν με τη βοήθεια του λογισμικού Audacity® 1.3.14-beta (Unicode) στο ΚΜΥ. Οι ηχογραφήσεις περιλαμβάνουν ακουστικό σήμα δύο καναλιών (stereo) με συχνότητα δειγματοληψίας 16000 Hz.

Το ακουστικό σήμα αρχικά διασπάστηκε σε δύο επιμέρους κανάλια (mono) και χρησιμοποιήθηκε μόνον η πληροφορία από το δεύτερο κανάλι για την εξαγωγή των δεδομένων, το οποίο περιείχε την ουσιαστική πληροφορία. Το πρώτο βήμα επεξεργασίας αποτέλεσε η αποθορυβοποίηση των ηχογραφήσεων που πραγματοποιήθηκε με τους αλγόριθμους αφαίρεσης φάσματος που περιγράφηκαν στο Κεφάλαιο 2.1. Η αποθορυβοποίηση έγινε μέσω του λογισμικού Audacity®. Αφού επιλέχθηκαν τμήματα της συνολικής ηχογράφησης που περιείχαν μόνο θόρυβο, η φασματική πληροφορία αυτών αφαιρέθηκε από το σύνολο του σήματος. Στη συνέχεια αφαιρέθηκαν χειροκίνητα από κάθε ηχογράφηση περίοδοι παρατεταμένης ησυχίας ή τμήματα που περιείχαν ήχους άλλους από ροχαλητό (π.χ. ομιλίες, βαριές αναπνοές, θόρυβο από πόρτες, βήματα κ.λπ.). Το αποθορυβοποιημένο σήμα ροχαλητών που προέκυψε για κάθε ηχογράφηση, αποθηκεύτηκε ως αρχείο μορφότυπου wav μέσω του Audacity® και φορτώθηκε ακολούθως στο λογισμικό Snor-Lux. Εκεί εκτελέστηκε ο αλγόριθμος κατάτμησης σήματος, ο οποίος έδωσε ως αποτέλεσμα τα επιμέρους ροχαλητά κάθε ηχογράφησης σε μορφή πίνακα. Οι πίνακες αυτοί χρησιμοποιήθηκαν για την εξαγωγή των ζητούμενων χαρακτηριστικών.

4.3: Μελέτη επίδρασης ρινικού αποσυμφορητικού σε συνδυασμό με κορτικοστεροειδή σε ασθενείς με OSAHS

Στο υποκεφάλαιο αυτό παρουσιάζεται αρχικά το πρωτόκολλο μελέτης επίδρασης ρινικού αποσυμφορητικού σε συνδυασμό με κορτικοστεροειδή. Στη συνέχεια παραθέεται το υλικό που χρησιμοποιήθηκε για τη μελέτη αυτή και ακολούθως η μεθοδολογία που ακολουθήθηκε για την εξαγωγή των μεγεθών ενδιαφέροντος (Μέση - μέγιστη ένταση και συχνότητα ροχαλητών).

4.3.1: Πρωτόκολλο μελέτης επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή

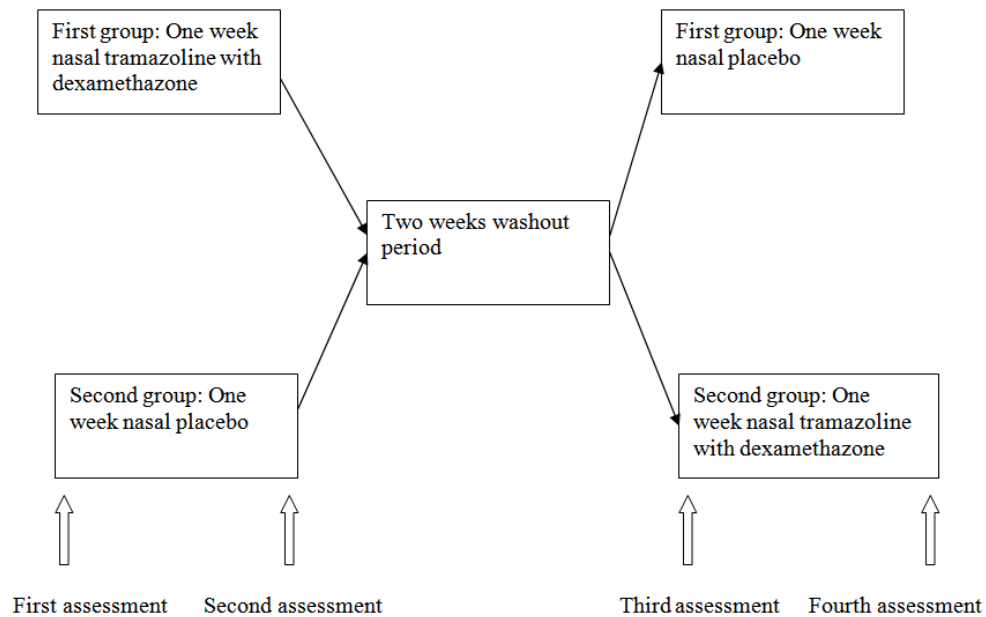
Η ρινική απόφραξη επηρεάζει την ποιότητα ύπνου και μπορεί να προδιαθέσει ροχαλητό και ακόμα και σύνδρομο αποφρακτικής άπνοιας - υπόπνοιας στον ύπνο. Παρά της σχέσης όμως μεταξύ ρινικής απόφραξης και του OSAHS, η θεραπευτική επίδραση της βελτίωσης βατότητας των ρινικών αεραγωγών σε ασθενείς με OSAHS παραμένει σημείο εικασίας. Η μελέτη που εκπονείται στο ΚΜΥ του νοσοκομείου "Ευαγγελισμός" αποσκοπεί στην

Κεφάλαιο 4

διερεύνηση του κατά πόσον η φαρμακολογική πρόληψη της ρινικής απόφραξης μπορεί να βελτιώσει την αναπνευστική δίοδο και να οδηγήσει σε μείωση του AHI σε ασθενείς με OSAHS και φυσιολογική ρινική αντίσταση. Για τη μελέτη χρησιμοποιείται συνδυασμός ρινικού αποσυμφορητικού (τραμαζολίνη - tramazoline) με κορτικοστεροειδή (δεξαμεθαζόνη - dexamethazone). Η τραμαζολίνη, η οποία είναι παράγωγο της ιμιδαζολίνης, είναι ρινικό αποσυμφορητικό που επιτυγχάνει τη μέγιστη ισχύ του 30 με 210 λεπτά μετά τη χορήγηση, ενώ η δεξαμεθαζόνη χρησιμοποιείται για την μετρίαση της ρινικής φλεγμονής που σχετίζεται με το OSAHS.

Η μελέτη γίνεται χρησιμοποιώντας μια διπλά τυφλή, τυχαιοποιημένη, ελεγχόμενη με εικονικό φάρμακο (placebo), αλληλεπικαλυπτόμενη σχεδίαση που φαίνεται στην Εικόνα 4.1. Οι ασθενείς χωρίζονται τυχαία σε δύο ομάδες και υποβάλλονται ο καθένας σε τέσσερις εκτιμήσεις. Κάθε αξιολόγηση αποτελείται από μια ολονύκτια πολυσωματογραφική μελέτη με ταυτόχρονο μέτρηση της διαδρομής αναπνοής και της ρινικής αντίστασης. Παράλληλα γίνεται ηχογράφηση του ασθενούς για ακουστική επεξεργασία των ροχαλητών. Με το πέρας της αρχικής εκτίμησης, χορηγείται στους ασθενείς της πρώτης ομάδας για μία εβδομάδα ρινική τραμαζολίνη με δεξαμεθαζόνη (2 φορές τη μέρα) ενώ στους ασθενείς της δεύτερης ομάδας χορηγείται για μία εβδομάδα ένα πανομοιότυπο ρινικό εικονικό φάρμακο (διάλυμα χλωριούχου νατρίου 0.9%, 2 φορές τη μέρα). Στο τέλος της εβδομάδας πραγματοποιείται η δεύτερη εκτίμηση. Ακολουθεί περίοδος έκπλυσης δύο εβδομάδων για όλους τους ασθενείς και πραγματοποιείται η τρίτη αξιολόγηση με το πέρας της περιόδου έκπλυσης. Στη συνέχεια εναλλάσσονται τα χορηγούμενα φάρμακα στις δύο ομάδες (ομάδα 1 - εικονικό ρινικό φάρμακο, ομάδα 2 - τραμαζολίνη με δεξαμεθαζόνη) για μία εβδομάδα (χορήγηση 2 φορές τη μέρα) και πραγματοποιείται η τέταρτη και τελευταία αξιολόγηση στο τέλος της εβδομάδας [45].

Για την αξιολόγηση των αποτελεσμάτων της μελέτης επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή επιλέχθηκαν τα μεγέθη της συχνότητας ροχαλητών (ροχαλητά ανά ώρα) και της μέσης και μέγιστης έντασης (σε dB).



Εικόνα 4.1: Σχεδιασμός μελέτης επίδρασης ρινικού αποσυμφορητικού σε συνδυασμό με κορτικοστεροειδή.

4.3.2: Υλικό μελέτης επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή

Για τους σκοπούς της πρώτης μελέτης που αφορά την επίδραση του συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών χρησιμοποιήθηκαν δεδομένα από τρεις εξεταζόμενους μόνο, αφού η μελέτη αυτή βρίσκεται ακόμα υπό διεξαγωγή τη στιγμή κατάθεσης της εργασίας. Οι εξεταζόμενοι επιλέχθηκαν χρησιμοποιώντας έξι κριτήρια αποκλεισμού:

- Πρόσφατη εγχείρηση στις άνω αναπνευστικές οδούς
- Περισσότερες από τρεις κεντρικές άπνοιες ανά ώρα ή ποσοστό κεντρικών απνοιών μεγαλύτερο του 5% των συνολικών απνοιών.
- Συνολικό χρόνο ύπνου κατά τη διάρκεια του πολυσωματογραφήματος λιγότερο από τρεις ώρες
- Τρέχουσα θεραπεία με ρινικά αποσυμφορητικά και τοπικά στεροειδή
- Νόσους των άνω ή κάτω αναπνευστικών οδών, περιλαμβανομένου και ιστορικού ρινικής αλλεργίας
- Κάπνισμα

Κεφάλαιο 4

Τα αναλυτικά χαρακτηριστικά των εξεταζομένων για την μελέτη επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή φαίνονται στον Πίνακα 4.1.

Πίνακας 4.1: Χαρακτηριστικά εξεταζομένων (Μελέτη επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή)

A/A	Φύλο	Ημερομηνία Καταγραφής	AHI (ε/ώ)	BMI (kg/m^2)	Ηλικία (έτη)
1	Θ	04/02/2012	17.4	25.0	62
2	A	22/03/2012	14.1	23.4	40
3*	A	05/04/2012	13.9	23.4	40

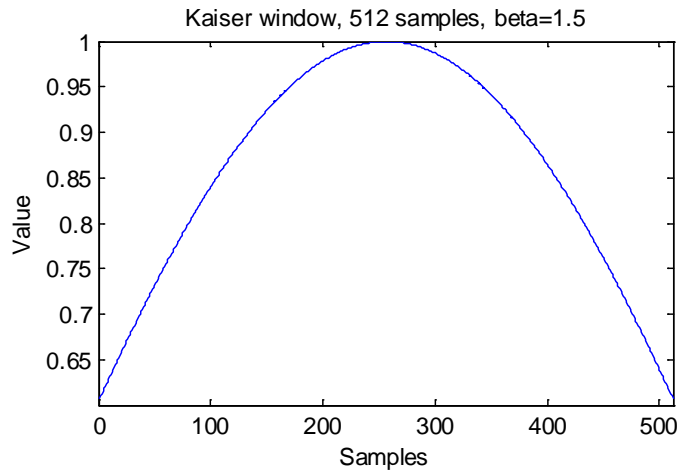
A: Άρρεν, Θ: Θήλυ, BMI: Body Mass Index

*Η ηχογράφιση (3) αποτελεί επανεξέταση του ίδιου ατόμου (2) σε άλλη ημερομηνία

4.3.3: Μέθοδος υπολογισμού συχνότητας και μέσης - μέγιστης έντασης ροχαλητών

Η κλινική μελέτη του ΚΜΥ επίδρασης του συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών απαιτεί για την αξιολόγηση της την εξαγωγή της μέσης και μέγιστης έντασης όλων των ροχαλητών κάθε νυκτερινής ηχογράφησης καθώς και της συχνότητας ροχαλητού. Για τις ανάγκες τις έρευνας χρησιμοποιήθηκε το λογισμικό Snor-Lux σε συνδυασμό με αλγόριθμο που αναπτύχθηκε στο λογισμικό MATLAB®. Για κάθε επιμέρους ροχαλητό εκτελούνται τα εξής βήματα:

1. Μετατροπή του ηχητικού σήματος έντασης σε dB.
2. Αντικατάσταση των τιμών του dB σήματος που είναι κάτω από ένα κατώφλι με ένα κατώτατο όριο για αποφυγή σφαλμάτων που οφείλονται σε πολύ μικρές τιμές σήματος.
3. Παραθύρωση ανά 512 δείγματα του σήματος με επικάλυψη 511 δείγματα με χρήση παραθύρου Kaiser με τιμή $\beta=1.5$ για καθορισμό βαρυτήτων σε κάθε τιμή του τμήματος. Ένα τέτοιο παράθυρο φαίνεται στην Εικόνα 4.2.
4. Εύρεση της μέσης έντασης του παραθυρομένου αυτού τμήματος και ορισμός της ως της τιμής της έντασης.
5. Επανάληψη των βημάτων 1 - 4 σε κάθε επιμέρους ροχαλητό κάθε ηχογράφησης



Εικόνα 4.2: Παράδειγμα παραθύρου Kaiser μήκους 512 δειγμάτων, με $\beta=1.5$

Σημειώνεται ότι μετά τον υπολογισμό του ομαλοποιημένου σήματος έντασης για κάθε επιμέρους ροχαλητό γινόταν εκκαθάριση των άχρηστων μεταβλητών στο MATLAB® για λόγους υπολογιστικής ταχύτητας και για αποφυγή έλλειψης μνήμης.

Για τον καθορισμό της συχνότητας ροχαλητού χρησιμοποιήθηκε η πληροφορία από το PSG του συνολικού χρόνου ύπνου και ο συνολικός αριθμός ροχαλητών κάθε ηχογράφησης. Συγκεκριμένα:

$$\text{Συχνότητα Ροχαλητού} = \frac{\text{Αριθμός ροχαλητών}}{\text{Χρόνος Ύπνου (secs)}} \cdot 3600 \left(\frac{\text{ροχαλητά}}{\text{ώρα}} \right)$$

4.4: Μελέτη διαγνωστικής ικανότητας χρονοσυχνοτικών χαρακτηριστικών ως προς το OSAHS

Στο υποκεφάλαιο αυτό παρουσιάζονται αρχικά τα χρονοσυχνοτικά μεγέθη των οποίων θα μελετηθεί η διαγνωστική ικανότητα ως προς το OSAHS. Στη συνέχεια παραθέεται το υλικό που χρησιμοποιήθηκε για τη μελέτη αυτή και ακολούθως η μεθοδολογία που ακολουθήθηκε για την εξαγωγή των μεγεθών ενδιαφέροντος (FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}). Τέλος αναλύεται η μέθοδος στατιστικής επεξεργασίας των αποτελεσμάτων που χρησιμοποιήθηκε, η ανάλυση με καμπύλες λειτουργικού χαρακτηριστικού δέκτη.

4.4.1: Επιλογή χρονοσυχνοτικών χαρακτηριστικών για εξέταση της διαγνωστικής τους ικανότητας

Για τις ανάγκες της μελέτης διαγνωστικής ικανότητας χρονοσυχνοτικών χαρακτηριστικών είναι απαραίτητη η επιλογή χαρακτηριστικών που διαφέρουν μεταξύ υγιών ατόμων και ατόμων που πάσχουν από OSAHS. Σύμφωνα με την [15] τα ακουστικά χαρακτηριστικά των σημάτων ροχαλητών και ειδικότερα το φασματικό τους περιεχόμενο διαφέρει σε υγιή και απνοϊκά άτομα. Επιλέχθηκαν για την ανάλυση της διαγνωστικής τους ικανότητας τρεις ενδεικτικές συχνότητες του φάσματος ισχύος, η μέση συχνότητα ($f_{50\%}$), η μέγιστη συχνότητα ($f_{90\%}$) και η συχνότητα κορυφής (f_{peak}). Οι συχνότητες $f_{50\%}$, $f_{90\%}$ και f_{peak} ορίζονται ως εξής:

- Μέση συχνότητα $f_{50\%}$: είναι η συχνότητα που περιέχει το 50% της συνολικής ισχύος του φάσματος ισχύος του σήματος.
- Μέγιστη συχνότητα $f_{90\%}$: είναι η συχνότητα που περιέχει το 90% της συνολικής ισχύος του φάσματος ισχύος του σήματος.
- Συχνότητα Κορυφής f_{peak} : είναι η συχνότητα στην οποία το φάσμα ισχύος του σήματος εμφανίζει τη μέγιστη του τιμή.

Σύμφωνα με την [19] το OSAHS συσχετίζεται με ανατομικές και λειτουργικές ανωμαλίες του άνω αεραγωγού. Συγκεκριμένα σε ασθενείς με OSAHS ο άνω αεραγωγός είναι στενότερος και καταρρέει ευκολότερα. Ο άνω αεραγωγός μπορεί να θεωρηθεί ένα ακουστικό μεταβλητό φίλτρο, στο οποίο παράγονται τα ακουστικά σήματα ροχαλητών, που επιτρέπει τη μέγιστη ροή ενέργειας στις συχνότητες φωνοσυντονισμού. Επιλέχθηκαν επομένως και οι τρεις συχνότητες φωνοσυντονισμού (FF1, FF2, FF3) για την μελέτη της διαγνωστικής τους ικανότητας ως προς το OSAHS.

Οι έξι συχνότητες ενδιαφέροντος εξάγονται για κάθε ροχαλητό από κάθε ολονύκτια ηχογράφηση και ακολούθως εισάγονται ως δεδομένα σε αλγόριθμο στατιστικής ανάλυσης τους που βασίζεται σε ανάλυση καμπυλών χαρακτηριστικού λειτουργικού δείκτη (ROC).

4.4.2: Υλικό μελέτης διαγνωστικής ικανότητας των FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}

Για τη μελέτη της ικανότητας διάγνωσης του OSAHS μέσω των FF και των $f_{90\%}$, $f_{50\%}$ και f_{peak} , χρησιμοποιήθηκαν δεδομένα από 11 εξεταζόμενους. Αναλύθηκαν 200 ροχαλητά,

τα οποία επιλέχθηκαν τυχαία από το σύνολο της νυκτερινής ηχογράφησης κάθε εξεταζόμενου με τη βοήθεια των λογισμικών Snor-Lux και MATLAB®. Για την μελέτη αυτή χρησιμοποιήθηκαν όσες ηχογραφήσεις υπήρχαν διαθέσιμες από το ΚΜΥ κατά το χρόνο εκπόνησης της διπλωματικής. Συγκεκριμένα αναλύθηκαν τα ροχαλητά 10 απνοϊκών εξεταζομένων και 1 υγιούς με βάση το κριτήριο διαχωρισμού υγιών - απνοϊκών εξεταζομένων το $AHI \geq 10$ ε/ώ. Σημειώνεται ότι στην καταγραφή (10) ο εξεταζόμενος δεν κοιμήθηκε αρκετή ώρα γι' αυτό αναλύθηκαν μόνο 100 ροχαλητά λόγω μη ύπαρξης περισσότερων. Τα αναλυτικά χαρακτηριστικά των εξεταζομένων για την μελέτη διαγνωστικής ικανότητας των FF , $f_{50\%}$, $f_{90\%}$ και f_{peak} φαίνονται στον Πίνακα 4.2.

Πίνακας 4.2: Χαρακτηριστικά εξεταζομένων (Μελέτη διαγνωστικής ικανότητας των FF , $f_{50\%}$, $f_{90\%}$ και f_{peak})

A/A	Φύλο	Ημερομηνία Καταγραφής	AHI (ε/ώ)	BMI (kg/m^2)	Ηλικία (έτη)
1	Θ	04/02/2012	17.4	25.0	62
2	Θ	11/03/2012	51.9	30.8	56
3	A	30/01/2012	45.5	40.0	45
4	A	03/02/2012	58.4	33.4	47
5	Θ	26/03/2012	1.6	22.0	49
6	Θ	12/03/2012	10.1	25.3	55
7	A	09/02/2012	41.7	27.5	34
8	A	22/03/2012	14.1	23.4	40
9*	A	05/04/2012	13.9	23.4	40
10	A	06/03/2012	104.3	33.2	52
11	A	18/04/2012	28.2	26.0	35

A: Άρρεν, Θ: Θήλυ, BMI: Body Mass Index

*Η ηχογράφηση (9) αποτελεί επανεξέταση του ίδιου ατόμου (8) σε άλλη ημερομηνία

Το γεγονός ότι στο σύνολο των καταγραφών υπάρχουν δεδομένα μόνο για έναν υγιή εξεταζόμενο (5) καθιστά την ανάλυση των δεδομένων αναξιόπιστη. Για το λόγο αυτό πέρα από την ανάλυση και σύγκριση των μεγεθών με βάση το διαχωρισμό του $AHI \geq 10$ ε/ώ γίνεται μια δεύτερη ανάλυση, αυξάνοντας το κατώφλι διαχωρισμού εξεταζομένων σε $AHI \geq 15$ ε/ώ, κατατάσσοντας μαζί έτσι τους υγιείς και τους απνοϊκούς με ήπιο και μέτριο OSAHS σε μια κατηγορία. Η κατάταξη αυτή αποσκοπεί στην εξέταση της ικανότητας των εξαγόμενων μεγεθών για διαχωρισμό των υγιών και μέτριων παθόντων OSAHS από αυτούς που πάσχουν από σοβαρό OSAHS.

4.4.3: Μέθοδος υπολογισμού συχνοτήτων φωνοσυντονισμού ροχαλητών

Για την εξαγωγή των FF από κάθε ροχαλητό έγινε επεξεργασία 200 ροχαλητών από κάθε εξεταζόμενο εκτός της καταγραφής 10, (100 ροχαλητά), τα οποία επιλέχθηκαν τυχαία από το σύνολο των ροχαλητών.

Όπως αναφέρθηκε στο Κεφάλαιο 2.3.12 υπάρχουν διάφοροι τρόποι για την εξαγωγή των FF. Στην διπλωματική εργασία χρησιμοποιήθηκε η μέθοδος γραμμικής πρόβλεψης (LPC). Η τάξη προβλέπτη που χρησιμοποιήθηκε, δεδομένου του ότι έχουμε συχνότητα δειγματοληψίας 16 KHz, είναι $\alpha = 18$ που προκύπτει από τον εμπειρικό κανόνα:

$$\text{Τάξη προβλέπτη } (\alpha) = 2 + \frac{f_s(\text{Hz})}{1000} = 2 + 16 = 18$$

Ακολουθως για κάθε επιμέρους ροχαλητό εκτελούνταν με τη βοήθεια των λογισμικών Spnr-Lux και MATLAB® τα εξής βήματα:

1. Εξαγωγή των συντελεστών γραμμικής πρόβλεψης με ελαχιστοποίηση του σφάλματος πρόβλεψης χρησιμοποιώντας την τεχνική ελαχίστων τετραγώνων.
2. Εξαγωγή ενός φίλτρου με αριθμητή μονάδα και παρονομαστή όρους με συντελεστές τους συντελεστές γραμμικής πρόβλεψης ως εξής:

$$H(e^{j\omega}) = \frac{1}{\alpha_1 + \alpha_2 e^{-j\omega} + \dots + \alpha_{18} e^{-j17\omega}}$$

3. Εύρεση των ριζών του ανωτέρω φίλτρου, των σημείων δηλαδή που μηδενίζεται η πρώτη παράγωγος του φίλτρου.
4. Παράλειψη των ριζών που έχουν φανταστικό μέρος μικρότερο του 0.01.
5. Εύρεση των συχνοτήτων που αντιστοιχούν σε κάθε ζεύγος ριζών μέσω της εξής σχέσης:

$$f_i = \arctan2\left(\text{imag}(r_i), \text{real}(r_i) * \frac{f_s}{2\pi}\right)$$

όπου r_i το ζεύγος ριζών, $\text{imag}()$ το φανταστικό μέρος, $\text{real}()$ το πραγματικό μέρος, f_s η συχνότητα δειγματοληψίας και $\arctan2$ η συνάρτηση τόξου εφαπτομένης με εύρος γωνίας από $-\pi$ έως π .

6. Οι FF: F1, F2 και F3 είναι οι τρεις πρώτες συχνότητες που προκύπτουν. Η ταξινόμηση των συχνοτήτων με αύξουσα σειρά.

7. Επανάληψη των βημάτων 1 - 6 σε κάθε επιμέρους ροχαλητό κάθε ηχογράφησης.

4.4.4: Μέθοδος υπολογισμού $f_{50\%}$, $f_{90\%}$ και f_{peak}

Για την εξαγωγή των $f_{50\%}$, $f_{90\%}$ και f_{peak} ακολουθήθηκαν τα εξής βήματα από αλγόριθμο που αναπτύχθηκε στο MATLAB®:

1. Εύρεση του αριθμού σημείων n για τον FT μέσω της σχέσης:

$$n = \text{ceil}(\log_2(\text{length}(x)))$$

όπου x το σήμα ροχαλητού προς ανάλυση, ceil η συνάρτηση στρογγυλοποίησης στον κοντινότερο ακέραιο και length η συνάρτηση που επιστρέφει τον αριθμό στοιχείων που περιέχει το όρισμα της.

2. Εύρεση του φάσματος ισχύος του σήματος με τη βοήθεια του γρήγορου μετασχηματισμού Fourier (FFT) σε n σημεία.

3. Εύρεση της συνολικής ισχύος του σήματος, αθροίζοντας τις τιμές του φάσματος ισχύος.

4. Καθορισμός κατωφλίων για εύρεση της μέσης και μέγιστης συχνότητας ως το 50% και 90% της συνολικής ισχύος αντίστοιχα.

5. Υπολογισμός του πίνακα αθροιστικής ισχύος του σήματος.

6. Εύρεση της μέσης και μέγιστης συχνότητας ως τις συχνότητες που απέχουν το λιγότερο στον πίνακα αθροιστικής ισχύος με χρήση των κατωφλίων που καθορίστηκαν στο βήμα 4.

7. Εύρεση της συχνότητας κορυφής με εύρεση της μέγιστης ισχύος στο προκύπτον φάσμα.

8. Επανάληψη των βημάτων 1 - 7 σε κάθε επιμέρους ροχαλητό κάθε ηχογράφησης.

4.4.5: Καμπύλες λειτουργικού χαρακτηριστικού δέκτη

Για τον έλεγχο των αποτελεσμάτων της επεξεργασίας κάθε ροχαλητού απαιτείται η στατιστική ανάλυση των αποτελεσμάτων. Για το σκοπό αυτό επιλέχθηκαν οι καμπύλες

Κεφάλαιο 4

λειτουργικού χαρακτηριστικού δέκτη (Receiver Operating Characteristic - ROC) (Εικόνα 4.3).

Οι καμπύλες ROC αποτελούν χρήσιμη τεχνική για την οργάνωση, επιλογή και απεικόνιση ταξινομητών με βάση τη γραφική τους παράσταση. Χρησιμοποιούνται ευρέως στη διαγνωστική ιατρική ενώ πρόσφατα έχουν υιοθετηθεί στους τομείς της μηχανικής μάθησης και στην εξόρυξη δεδομένων. Ο ορισμός της καμπύλης ROC είναι το μοναδιαίο τετράγωνο $[0,1] \times [0,1]$ όπου το σημείο $(0,0)$ είναι το σημείο όπου το σημείο απόφασης είναι μεγαλύτερο από όλες τις μετρήσεις σήματος και θορύβου και το σημείο $(1,1)$ το σημείο όπου το σημείο απόφασης είναι μικρότερο από όλες τις μετρήσεις. Το εμβαδόν που ορίζεται κάτω από την καμπύλη αποτελεί ποιοτικό μέτρο διαχωρισμού θορύβου - σήματος και χρησιμοποιείται στη στατιστική συμπερασματολογία των καμπυλών ROC.

Στην ιατρική έρευνα αναπτύσσονται συχνά διαγνωστικοί έλεγχοι για το διαχωρισμό υγιών και ασθενών, όπως συμβαίνει και στην διπλωματική εργασία. Η χρησιμότητα μιας τέτοιας έρευνας προσδιορίζεται από τη διακριτική της ικανότητα, την ακρίβεια δηλαδή, με την οποία διαχωρίζει τους εξεταζόμενους σε ασθενείς ή μη. Οι κυριότερες συνιστώσες αξιολόγησης μιας έρευνας είναι η ευαισθησία της δοκιμασίας (sensitivity) ή ποσοστό αληθώς θετικών αποτελεσμάτων (true positive rate - TPR) και η ειδικότητα (specificity) ή ποσοστό αληθώς αρνητικών αποτελεσμάτων (true negative rate - TNR). Η ευαισθησία (SE) και η ειδικότητα (SPC) δίνονται από τις παρακάτω σχέσεις:

$$SE = TPR = \frac{TP}{P} = \frac{TP}{TP + FN}, \quad SPC = TNR = \frac{TN}{N} = \frac{TN}{FP + TN}$$

όπου TP (True Positive) τα αληθώς θετικά αποτελέσματα, TN (True Negative) τα αληθώς αρνητικά αποτελέσματα, FP (False Positive) τα ψευδώς θετικά αποτελέσματα και FN (False Negative) τα ψευδώς αρνητικά αποτελέσματα.

Τα μεγέθη της ευαισθησίας και ειδικότητας έχουν σχέση αντιστρόφως ανάλογη. Βελτιώνοντας το ποσοστό ευαισθησίας μειώνεται η ειδικότητα και αντίστροφα. Η απαίτηση για μια αξιόπιστη έρευνα είναι να είναι ταυτόχρονα υψηλές οι τιμές και των δύο μεγεθών.

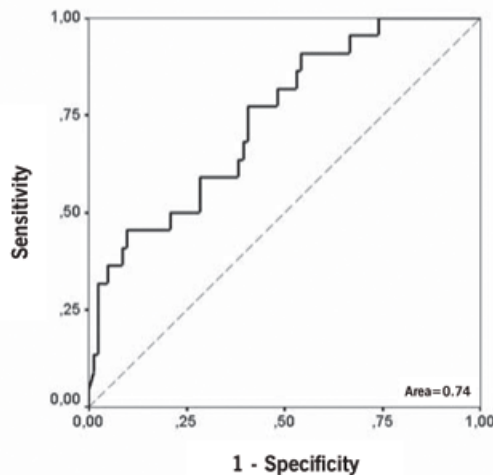
Σημειώνεται ότι τα δεδομένα που χρησιμοποιούνται για την ανάλυση ROC στην διπλωματική εργασία θεωρούνται ως εξέταση αναφοράς (Gold Standard) αφού είναι

γνωστός ο διαχωρισμός μεταξύ απνοϊκών και υγιών εξεταζομένων, λόγω του παράλληλου πολυσωματογραφήματος που παρέχει την ακριβή τιμή ΑΗΙ.

Λόγω του μικρού αριθμού δείγματος που συμμετέχει συνήθως σε κλινικές έρευνες, κάτι που συμβαίνει και στα δεδομένα προς ανάλυση της διπλωματικής εργασίας, είναι απαιτούμενος ο έλεγχος ακρίβειας των αποτελεσμάτων, ευαισθησίας και ειδικότητας. Για το λόγο αυτό οι τιμές που προκύπτουν παραθέτονται μαζί με τις τιμές του διαστήματος εμπιστοσύνης 95% (Confidence Interval - CI). Η τιμές του 95% CI υπολογίζονται ως εξής:

$$95\% \text{ CI value} = p \pm 1.96 \sqrt{\frac{p(1-p)}{N}}$$

όπου p η ευρεθείσα τιμές ευαισθησίας ή ειδικότητας και N ο αριθμός εξεταζομένων.



Εικόνα 4.3: Παράδειγμα καμπύλης ROC
[Πηγή: <http://www.scielo.br>]

Το εμβαδόν κάτω από την καμπύλη ROC (AUC - Area under the curve) εκφράζει την πιθανότητα ένα ζεύγος μετρήσεων, μιας από τους υγιείς και μιας από τους ασθενείς να ταξινομηθεί με τη σωστή σειρά, δηλαδή η μέτρηση του ασθενούς να είναι υψηλότερη από αυτή του υγιούς.

Τέλος μέσω της ανάλυσης των καμπυλών ROC μπορεί να γίνει επιλογή του βέλτιστου σημείου απόφασης (cut-off point) για διαχωρισμό ασθενών - υγιών. Το επικρατές κριτήριο για την επιλογή του σημείου απόφασης στηρίζεται στη μεγιστοποίηση των ορθών αποφάσεων (θετικών και αρνητικών) και είναι αυτό που απέχει τη μεγαλύτερη απόσταση από τη διαγώνιο σε μια καμπύλη ROC. Το σημείο αυτό πετυχαίνει τις μεγαλύτερες δυνατές ταυτόχρονες τιμές ευαισθησίας και ειδικότητας [46].

Κεφάλαιο 5: Αποτελέσματα

Σε αυτό το κεφάλαιο καταγράφονται τα αποτελέσματα των δύο μελετών που αναφέρθηκαν και περιγράφηκαν στα προηγούμενα κεφάλαια. Αρχικά παρουσιάζονται οι μέσες και μέγιστες εντάσεις και συχνότητες που προέκυψαν από τις ηχογραφήσεις της μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού με κορτικοστεροειδή [45]. Εν συνεχεία παρουσιάζονται τα αποτελέσματα της δεύτερης μελέτης που αφορά τις FF και τις $f_{50\%}$, $f_{90\%}$ και f_{peak} . Παραθέτονται οι μέσες τιμές και τυπικές αποκλίσεις των έξι μεγεθών ανάλυσης και ενδεικτικές γραφικές παραστάσεις των μεγεθών για δύο περιπτώσεις. Στη συνέχεια παρουσιάζονται τα αποτελέσματα της ανάλυσης με καμπύλες ROC των έξι συχνοτήτων ενδιαφέροντος για δύο τιμές κατωφλίου AHI. Στην πρώτη περίπτωση θεωρείται κατώφλι για το διαχωρισμό υγιών - απνοϊκών ατόμων το AHI ≥ 10 ε/ώ ενώ στη δεύτερη το AHI ≥ 15 ε/ώ.

5.1: Αποτελέσματα μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών

Στον Πίνακα 5.1 παρουσιάζονται οι τιμές συχνοτήτων, μέσων και μέγιστων εντάσεων που προέκυψαν από την επεξεργασία των σημάτων ροχαλητού της [45].

Πίνακας 5.1: Αποτελέσματα μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού και κορτικοστεροειδών

A/A	AHI (ε/ώ)	Αριθμός ροχαλητών	Ώρα ύπνου (min)	Συχνότητα ροχαλητών (ροχαλητά/ώρα)	Μέγιστη Ένταση (dB)	Μέση Ένταση (dB)
1	17.4	1342	307	262.3	83.4	52.8
2	14.1	3941	414.5	570.5	93.9	60.5
3	13.9	2757	397.5	416.2	96.6	61.1

Η εξαγωγή συμπερασμάτων από τα μέχρι τώρα αποτελέσματα της μελέτης δυστυχώς δεν είναι εφικτή, καθώς δεν έχουν ολοκληρωθεί ακόμα οι ηχογραφήσεις και οι αντίστοιχες πολυσωματογραφικές μελέτες στους συμμετέχοντες στην έρευνα. Τα διαθέσιμα δεδομένα κατά το χρόνο εκπόνησης της διπλωματικής, δηλαδή οι τρεις ηχογραφήσεις που παρουσιάζονται, δεν είναι αρκετές ούτε και αντιπροσωπευτικές αφού δεν περιλαμβάνουν τα τέσσερα στάδια ηχογραφήσεων που περιγράφονται στο Κεφάλαιο 4.3.1 ώστε να μπορούν να αναλυθούν. Όπως εξηγήθηκε, η έρευνα είναι σε διεξαγωγή στο ΚΜΥ του νοσοκομείου "Ευαγγελισμός" και θα συνεχιστεί η συνεργασία με το ΚΜΥ και μετά το πέρας της διπλωματικής ενώ τα αποτελέσματα θα παρουσιαστούν σε δημοσίευση όταν τελειώσει η έρευνα.

Για τις ανάγκες της μελέτης [45] αναπτύχθηκε λογισμικό φιλικό προς το χρήστη σε περιβάλλον MATLAB®, το οποίο εγκαταστάθηκε στους υπολογιστές του ΚΜΥ μαζί με το λογισμικό Snor-Lux, ώστε να είναι δυνατή η εξαγωγή των απαιτούμενων μεγεθών (συχνότητα και μέση - μέγιστη ένταση) από το ιατρικό προσωπικό που διεξάγει τη μελέτη. Τα μεγέθη θα χρησιμοποιηθούν ως κριτήριο αξιολόγησης της επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή σε ασθενείς με OSAHS.

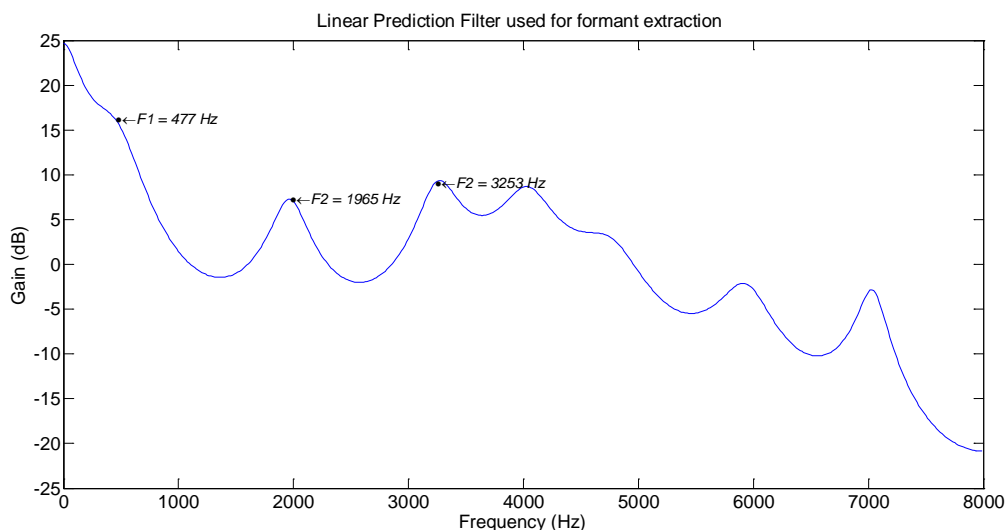
5.2: Αποτελέσματα εξαγωγής FF , $f_{50\%}$, $f_{90\%}$ και f_{peak}

Στον Πίνακα 5.2 παρουσιάζονται οι μέσες τιμές και τυπικές αποκλίσεις των FF , $f_{50\%}$, $f_{90\%}$ και f_{peak} για κάθε εξεταζόμενο της μελέτης διαγνωστικής ικανότητας των συχνοτήτων αυτών.

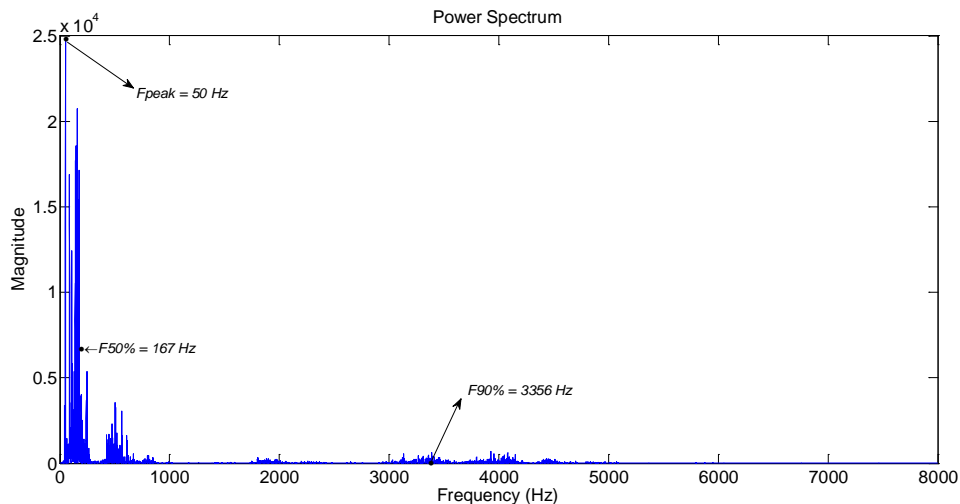
Πίνακας 5.2: Μέση τιμή \pm Τυπική απόκλιση (STD) των FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}

A/A	AHI (ε/ώ)	F1 \pm STD (Hz)	F2 \pm STD (Hz)	F3 \pm STD (Hz)	$f_{50\%} \pm$ STD (Hz)	$f_{90\%} \pm$ STD (Hz)	$f_{peak} \pm$ STD (Hz)
1	17.4	1333 \pm 217	1680 \pm 131	2273 \pm 364	2521 \pm 902	5778 \pm 1308	1860 \pm 687
2	51.9	241 \pm 44	1214 \pm 103	2014 \pm 259	252 \pm 82	1747 \pm 1621	155 \pm 110
3	45.5	1407 \pm 26	2317 \pm 108	2891 \pm 250	4189 \pm 46	5994 \pm 342	3056 \pm 1832
4	58.4	928 \pm 315	1876 \pm 384	2954 \pm 429	617 \pm 307	3971 \pm 1419	67 \pm 18
5	1.6	122 \pm 14	1022 \pm 160	2107 \pm 149	92 \pm 10	239 \pm 52	71 \pm 21
6	10.1	351 \pm 78	971 \pm 48	2100 \pm 18	667 \pm 444	4093 \pm 1199	259 \pm 220
7	41.7	258 \pm 36	1088 \pm 102	1923 \pm 142	278 \pm 61	910 \pm 432	156 \pm 130
8	14.1	359 \pm 104	1318 \pm 415	2041 \pm 204	348 \pm 163	2843 \pm 1858	220 \pm 153
9	13.9	260 \pm 57	1441 \pm 176	2090 \pm 48	227 \pm 80	374 \pm 62	192 \pm 115
10	104.3	651 \pm 206	1748 \pm 582	2928 \pm 633	421 \pm 259	4447 \pm 1118	196 \pm 236
11	28.2	270 \pm 61	648 \pm 66	1903 \pm 25	283 \pm 57	840 \pm 314	70 \pm 53

Παρατηρώντας τις μέσες τιμές και τυπικές αποκλίσεις των έξι συχνοτήτων, αλλά και από τη φύση των σημάτων ροχαλητού των ηχογραφήσεων 1 και 3 (αντί για ροχαλητά οι εξεταζόμενοι ανάπνεαν βαριά) κρίνεται αναγκαίο να μην συμπεριληφθούν στην ανάλυση καμπύλες ROC ως εσφαλμένες μετρήσεις. Συγκεκριμένα, όπως παρατηρείται στον Πίνακα 5.2, η πρώτη συχνότητα φωνοσυντονισμού, η μέση συχνότητα, η μέγιστη συχνότητα αλλά και η συχνότητα κορυφής δείχνουν ότι τα σήματα για τους εξεταζόμενους 1 και 3 διαφέρουν κατά πολύ από τα σήματα των υπόλοιπων ηχογραφήσεων. Στις Εικόνες 5.1 και 5.2 παρουσιάζονται οι F1, F2, F3 και $f_{50\%}$, $f_{90\%}$ και f_{peak} αντίστοιχα, για οπτικοποίηση και καλύτερη κατανόηση των μεγεθών που αναλύονται.



Εικόνα 5.1: Οι τρεις πρώτες συχνότητες φωνοσυντονισμού για σήμα ροχαλητού



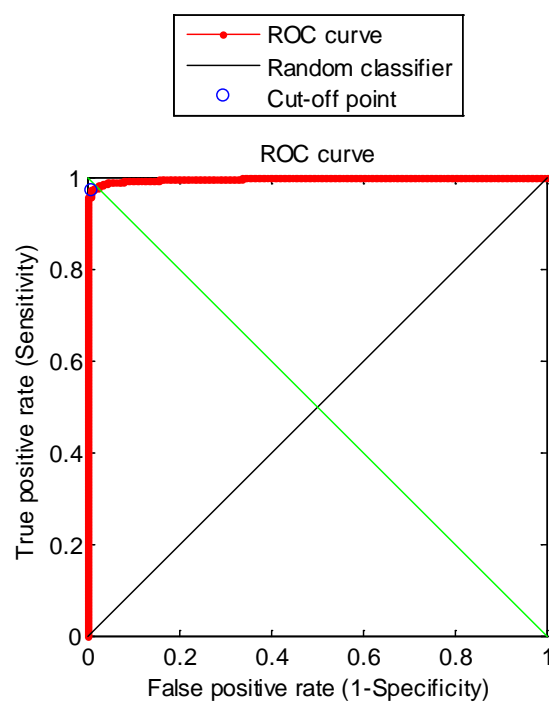
Εικόνα 5.2: Οι συχνότητες $f_{50\%}$, $f_{90\%}$ και f_{peak} στο φάσμα σήματος ροχαλητού

5.3: Ανάλυση με καμπύλες ROC (Κριτήριο AHI \geq 10 ε/ώ)

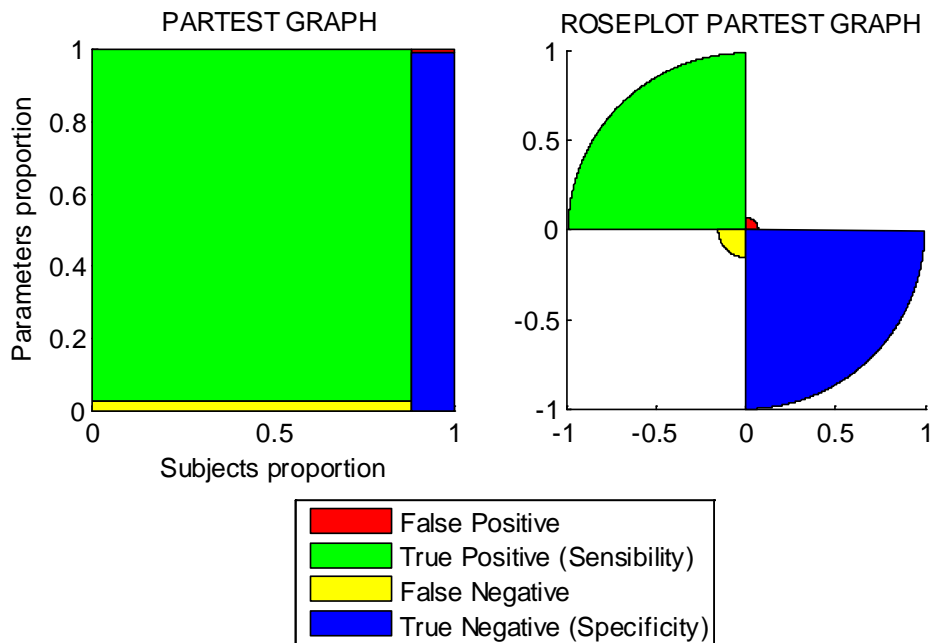
Στο υποκεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα της ανάλυσης με καμπύλες ROC για τα έξι μεγέθη ανάλυσης με κριτήριο διαχωρισμού υγιών - απνοϊκών AHI \geq 10 ε/ώ.

5.3.1: Στατιστική ανάλυση F1 (Κριτήριο AHI \geq 10 ε/ώ)

Στην Εικόνα 5.3 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F1 και στην Εικόνα 5.4 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.3 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.3: Καμπύλη ROC για F1 (AHI \geq 10 ε/ώ)



Εικόνα 5.4: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (F1)

Πίνακας 5.3: Αποτελέσματα ανάλυσης ROC για F1 (AHI ≥ 10 ε/ώ)

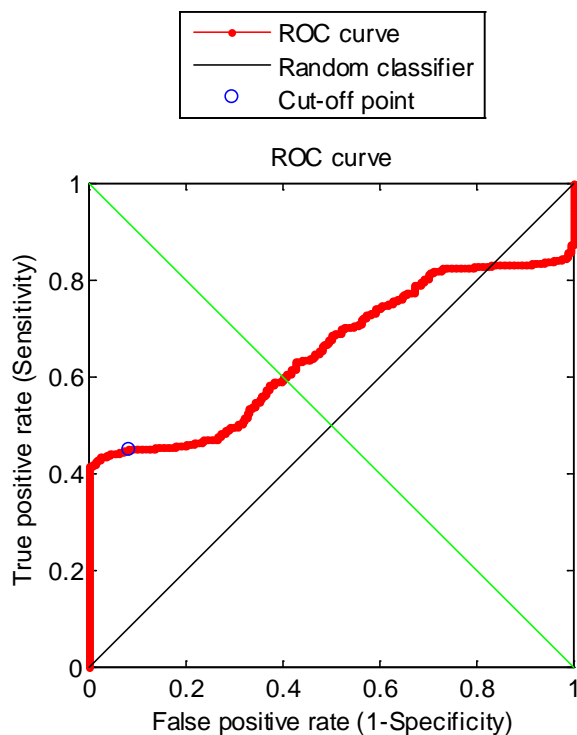
	n	A.P.	AHI \pm STD (ε/ώ)	F1 \pm STD (Hz)	Κατώφλι F1 (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
A	8	1500	40.3 \pm 31.6	399 \pm 269	161	99.83	97.7	99.5	Εξαιρετική
Y	1	200	1.6 \pm 0	122 \pm 14					
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	99.69	96.9	98.5	
					Άνω όριο	99.98	98.4	100	

A: Απνοιικοί, Y: Υγιείς, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

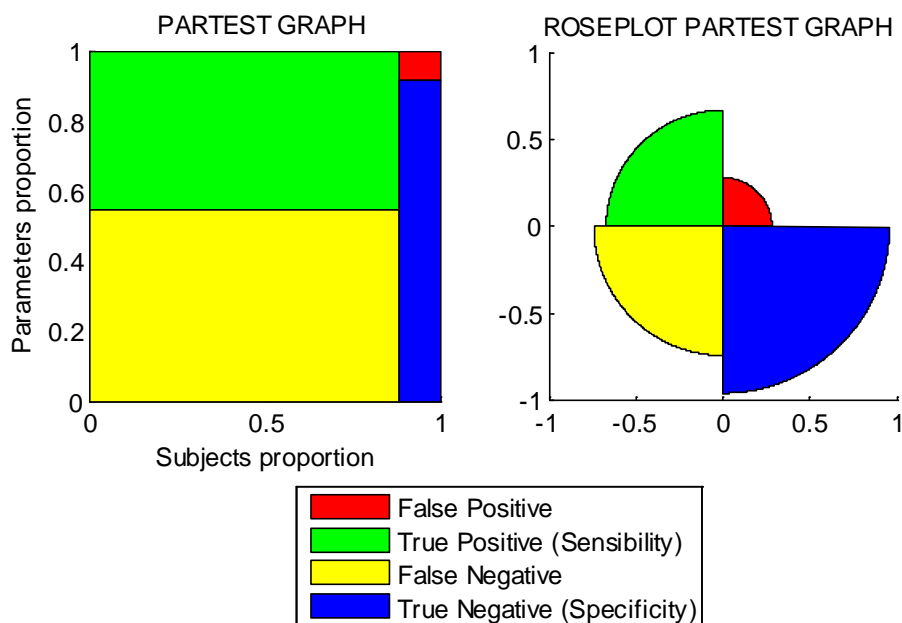
Όπως παρατηρείται από το αποτέλεσμα της ανάλυσης ROC, η πρώτη συχνότητα φωνοσυντονισμού F1 μπορεί να αποτελέσει διαγνωστικό εργαλείο για το OSAHS. Η ειδικότητα και ευαισθησία που προκύπτουν (99.5% και 97.7% αντίστοιχα) έχουν πολύ υψηλές τιμές καθιστώντας την ανάλυση εξαιρετική. Η συχνότητα κατωφλίου για διαχωρισμό υγιών - απνοικών ατόμων (161 Hz) είναι παρόλα αυτά αρκετά χαμηλή συχνότητα και δεν συμπίπτει με τη συχνότητα κατωφλίου (470 Hz) που βρέθηκε σε παρόμοια ανάλυση FF ακουστικών σημάτων ροχαλητού στην [19].

5.3.2: Στατιστική ανάλυση F2 (Κριτήριο AHI \geq 10 ε/ώ)

Στην Εικόνα 5.5 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F2 και στην Εικόνα 5.6 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.4 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.5: Καμπύλη ROC για F2 (AHI \geq 10 ε/ώ)



Εικόνα 5.6: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (F2)

Πίνακας 5.4: Αποτελέσματα ανάλυσης ROC για F2 (AHI ≥ 10 ε/ώ)

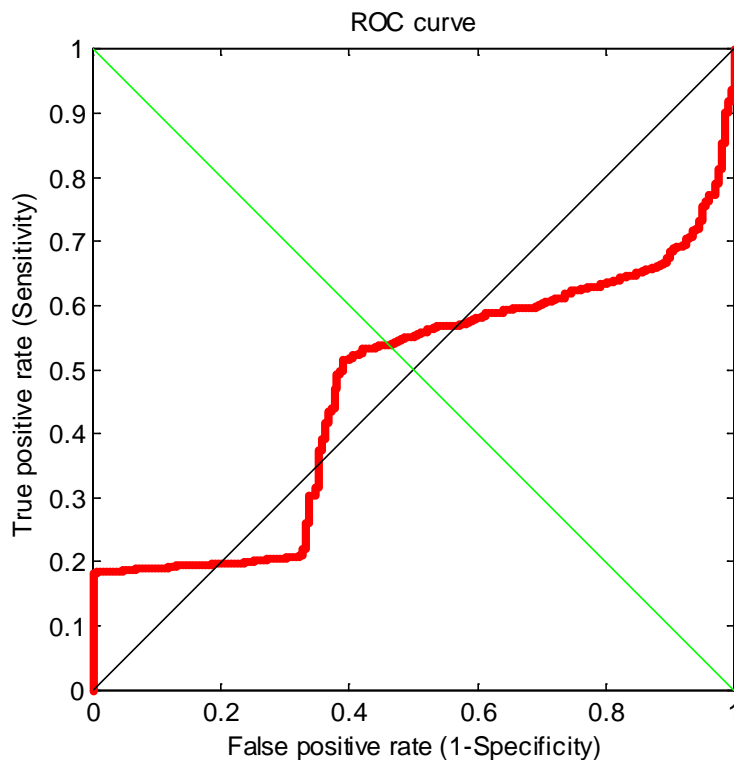
	n	A.P.	AHI \pm STD (ε/ώ)	F2 \pm STD (Hz)	Κατώφλι F2 (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
A	8	1500	40.3 \pm 31.6	1257 \pm 457	1220	65.42	45.0	92.0	Ανεπαρκής
Υ	1	200	1.6 \pm 0	1022 \pm 160					
Διάστημα Εμπιστοσύνης 95%									
					Άνω όριο	69.09	47.5	95.8	

A: Απνοιϊκοί, Υ: Υγιείς, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

Το αποτέλεσμα της ανάλυσης ROC για την δεύτερη συχνότητα φωνοσυντονισμού F2, κρίνεται αναξιόπιστο λόγω των χαμηλών τιμών ευαισθησίας και AUC (45% και 65.4% αντίστοιχα). Αυτό οδηγεί στην απόρριψη της χρήσης της F2 ως διαγνωστικό εργαλείο για το OSAHS.

5.3.3: Στατιστική ανάλυση F3 (Κριτήριο AHI ≥ 10 ε/ώ)

Στην Εικόνα 5.7 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F3. Στον Πίνακα 5.5 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.7: Καμπύλη ROC για F3 (AHI ≥ 10 ε/ώ)

Κεφάλαιο 5

Πίνακας 5.5: Αποτελέσματα ανάλυσης ROC για F3 (AHI ≥ 10 ε/ώ)

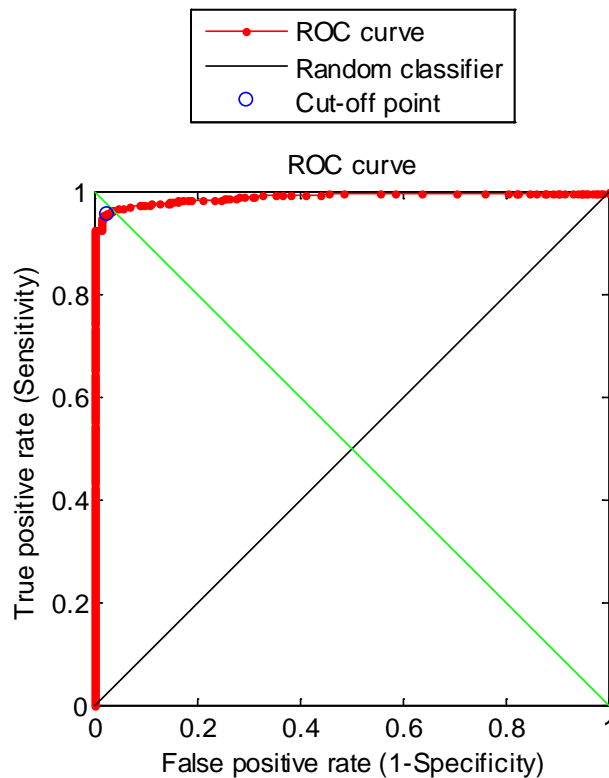
	n	A.P.	AHI \pm STD (ε/ώ)	F3 \pm STD (Hz)	Κατώφλι F3 (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
A	8	1500	40.3 \pm 31.6	2196 \pm 461	-	46.77	-	-	Αποτυχημένη
Υ	1	200	1.6 \pm 0	2107 \pm 149	-	-	-	-	
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	42.45	-	-	
					Άνω όριο	51.09	-	-	

A: Απνοϊκοί, Υ: Υγιείς, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

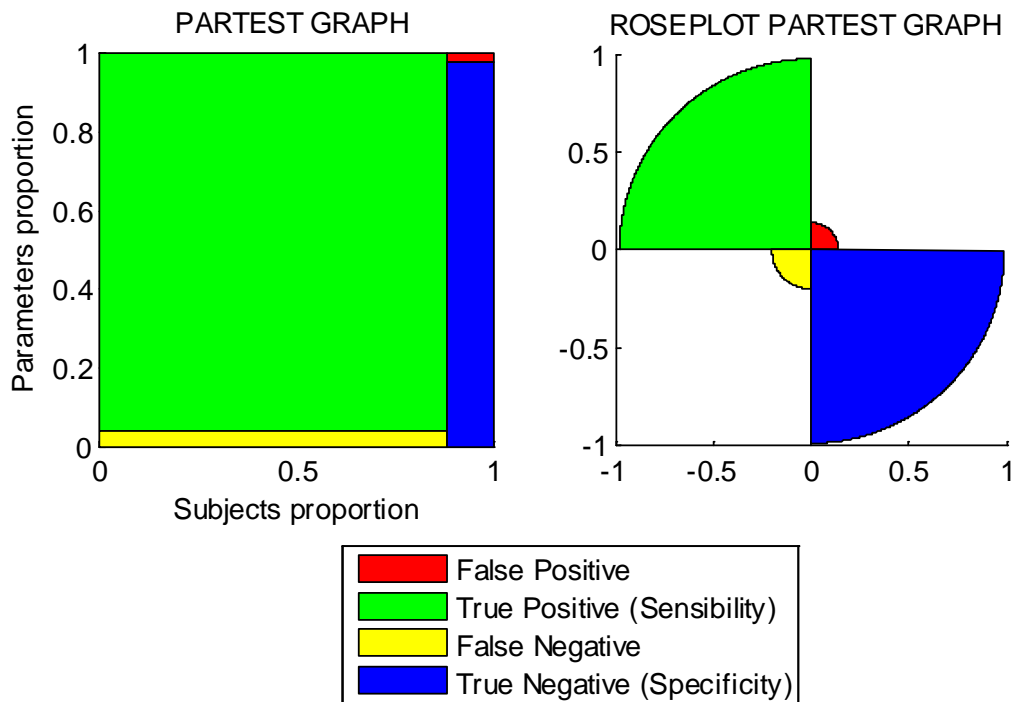
Στην περίπτωση της τρίτης συχνότητας φωνοσυντονισμού F3, η ανάλυση ROC δεν έδωσε αποτελέσματα και κρίνεται αποτυχημένη. Αυτό οφείλεται στην ομοιότητα που παρουσιάζουν οι συχνότητες F3 όπως φαίνεται στον Πίνακα 5.5 (Μέση τιμή \pm Τυπική απόκλιση: Απνοϊκοί - 2196 \pm 461, Υγιείς - 2107 \pm 149). Η F3 κρίνεται επομένως αναξιόπιστη για την διάγνωση του OSAHS.

5.3.4: Στατιστική ανάλυση $f_{50\%}$ (Κριτήριο AHI ≥ 10 ε/ώ)

Στην Εικόνα 5.8 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων $f_{50\%}$ και στην Εικόνα 5.9 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.6 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.8: Καμπύλη ROC για $f_{50\%}$ (AHI ≥ 10 ε/ώ)



Εικόνα 5.9: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN ($f_{50\%}$)

Πίνακας 5.6: Αποτελέσματα ανάλυσης ROC για $f_{50\%}$ (AHI ≥ 10 ε/ώ)

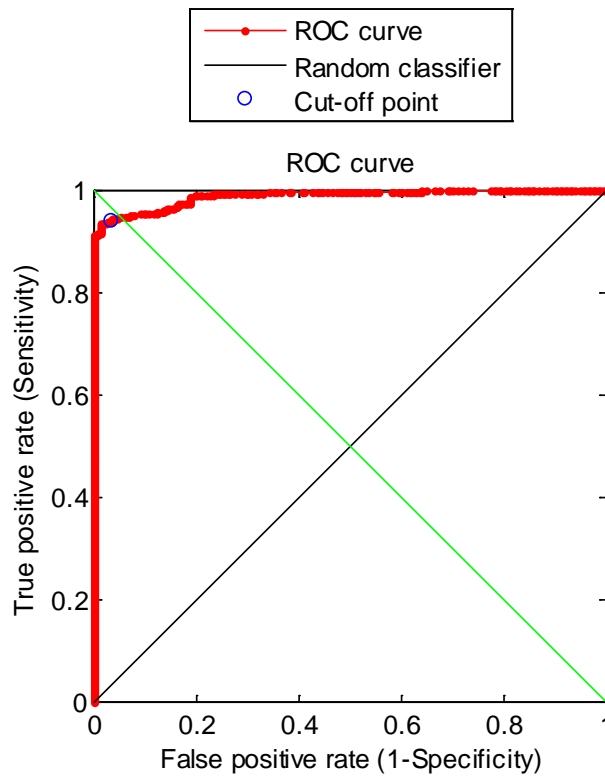
	n	A.P.	AHI \pm STD (ε/ώ)	$f_{50\%} \pm$ STD (Hz)	Κατώφλι $f_{50\%}$ (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
A	8	1500	40.3 \pm 31.6	384 \pm 276	113	99.19	96.0	98.0	Εξαιρετική
Y	1	200	1.6 \pm 0	92 \pm 10					
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	98.84	95.0	96.1	
					Άνω όριο	99.53	97.0	99.9	

A: Απνοϊκοί, Y: Υγιείς, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

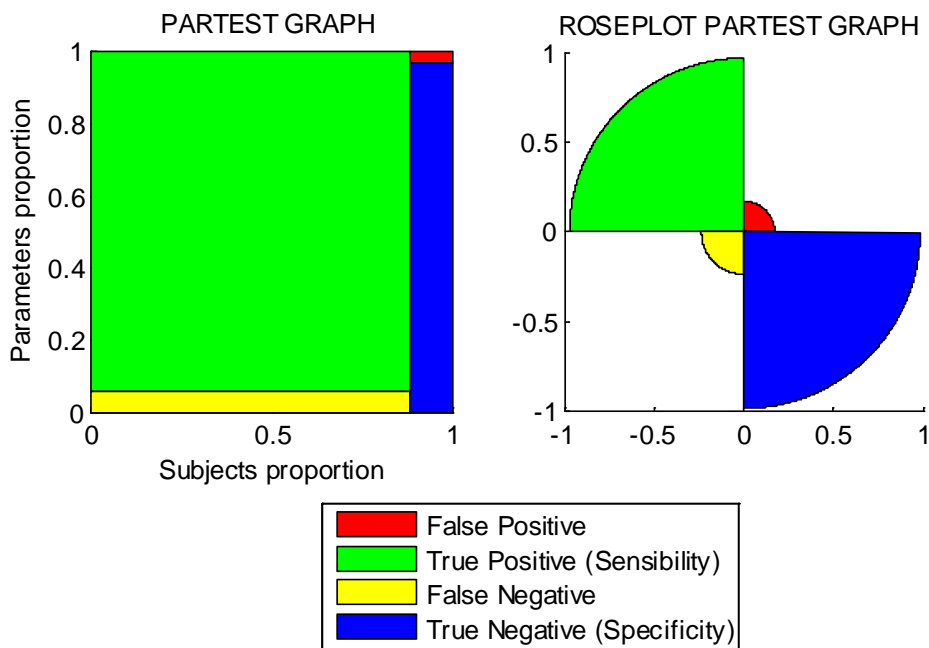
Η μέση συχνότητα $f_{50\%}$ εμφανίζει πολύ καλά αποτελέσματα στην ανάλυση ROC. Με ευαισθησία, ειδικότητα και AUC 96%, 98% και 99.2% αντίστοιχα, η ανάλυση κρίνεται εξαιρετική και επομένως μπορεί να χρησιμοποιηθεί για τη διάγνωση του OSAHS. Σαν κατώφλι διαχωρισμού υγιών - απνοϊκών προκύπτει η συχνότητα 113 Hz.

5.3.5: Στατιστική ανάλυση $f_{90\%}$ (Κριτήριο $AHI \geq 10$ ε/ώ)

Στην Εικόνα 5.10 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων $f_{90\%}$ και στην Εικόνα 5.11 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.7 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.10: Καμπύλη ROC για $f_{90\%}$ ($AHI \geq 10$ ε/ώ)



Εικόνα 5.11: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN ($f_{90\%}$)

Πίνακας 5.7: Αποτελέσματα ανάλυσης ROC για $f_{90\%}$ (AHI ≥ 10 $\epsilon/\acute{\omega}$)

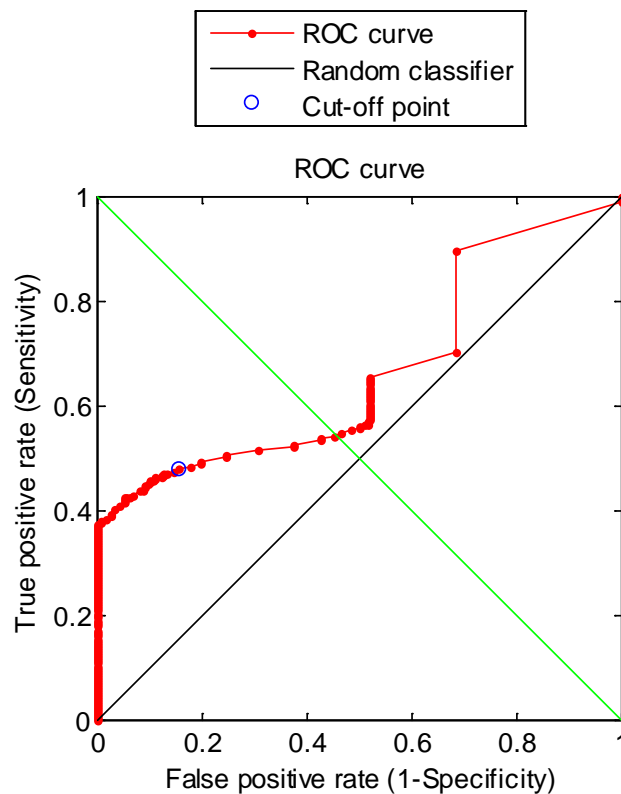
	n	A.P.	AHI \pm STD ($\epsilon/\acute{\omega}$)	$f_{90\%} \pm$ STD (Hz)	Κατώφλι $f_{90\%}$ (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
A	8	1500	40.3 \pm 31.6	2267 \pm 1900	344	99.02	94.5	97.0	Εξαιρετική
Y	1	200	1.6 \pm 0	239 \pm 52					
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	98.65	93.3	94.6	
					Άνω όριο	99.40	95.6	99.4	

A: Απνοϊκοί, Y: Υγιείς, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

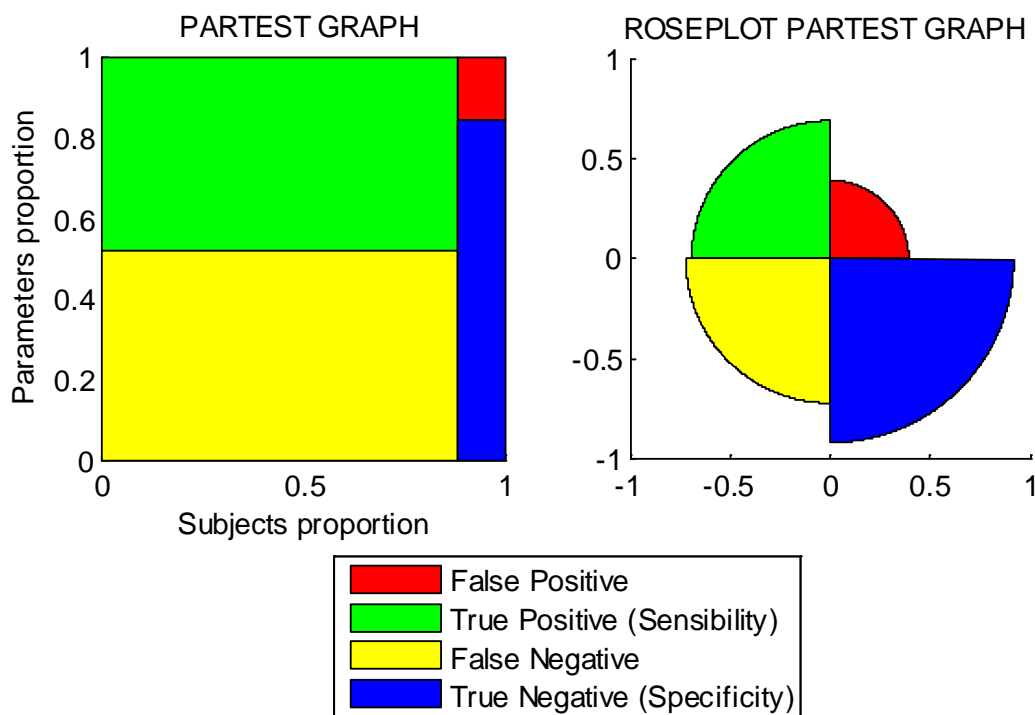
Η μέγιστη συχνότητα $f_{90\%}$ παρουσιάζει, όπως και η μέση συχνότητα, πολύ καλά αποτελέσματα στην ανάλυση ROC. Με ευαισθησία, ειδικότητα και AUC 94.5%, 97% και 99% αντίστοιχα, η ανάλυση κρίνεται εξαιρετική και επομένως μπορεί να χρησιμοποιηθεί για τη διάγνωση του OSAHS. Σαν κατώφλι διαχωρισμού υγιών - απνοϊκών προκύπτει η συχνότητα 344 Hz.

5.3.6: Στατιστική ανάλυση f_{peak} (Κριτήριο AHI ≥ 10 $\epsilon/\acute{\omega}$)

Στην Εικόνα 5.12 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων f_{peak} και στην Εικόνα 5.13 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.8 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.12: Καμπύλη ROC για f_{peak} (AHI ≥ 10 $\epsilon/\acute{\omega}$)



Εικόνα 5.13: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (f_{peak})

Πίνακας 5.8: Αποτελέσματα ανάλυσης ROC για f_{peak} (AHI ≥ 10 ε/ώ)

	n	A.P.	AHI \pm STD (ε/ώ)	$f_{peak} \pm$ STD (Hz)	Κατώφλι f_{peak} (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
A	8	1500	40.3 \pm 31.6	162 \pm 153	89	66.70	48.0	84.5	Ανεπαρκής
Y	1	200	1.6 \pm 0	71 \pm 21					
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	63.11	45.5	79.5	
					Άνω όριο	70.30	50.5	89.5	

A: Απνοιικοί, Y: Υγείς, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

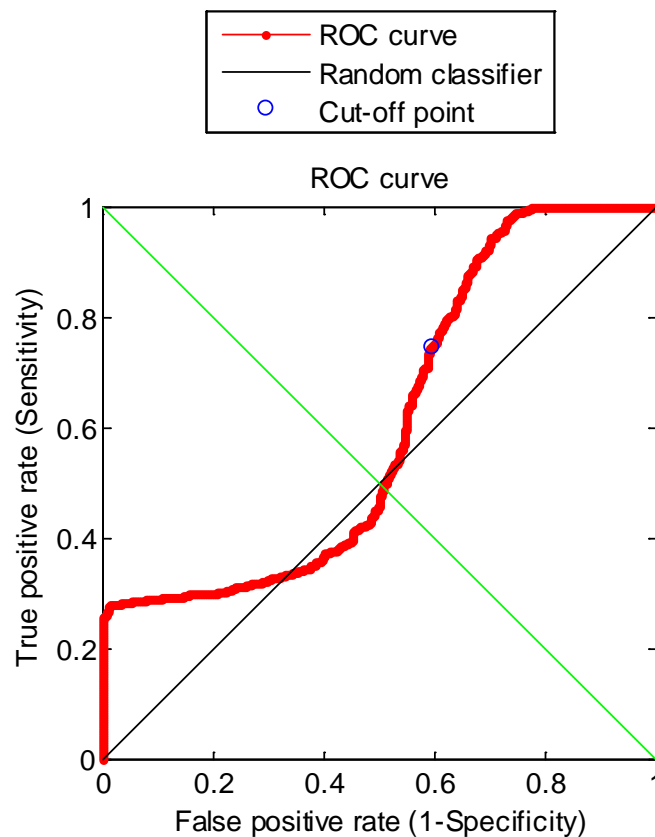
Το αποτέλεσμα της ανάλυσης ROC για συχνότητα κορυφής f_{peak} , κρίνεται αναξιόπιστο λόγω των χαμηλών τιμών ευαισθησίας και AUC (48% και 66.7% αντίστοιχα). Αυτό οδηγεί στην απόρριψη της χρήσης της f_{peak} ως διαγνωστικό εργαλείο για το OSAHS.

5.4: Ανάλυση με καμπύλες ROC (Κριτήριο $AHI \geq 15$ ε/ώ)

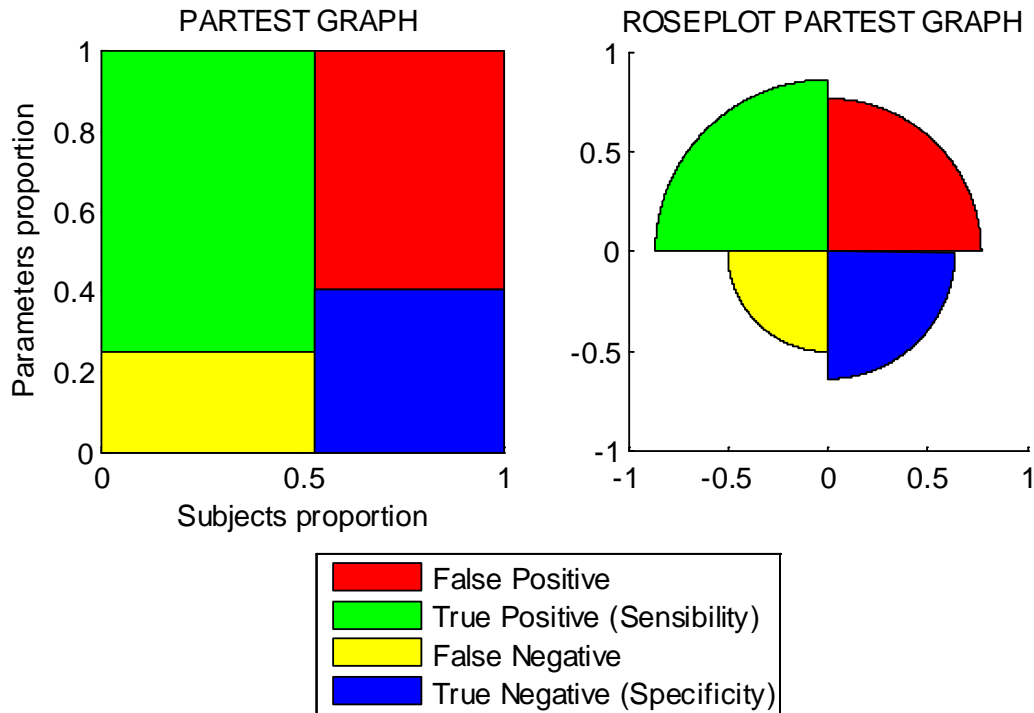
Όπως αναφέρθηκε η παρουσία δεδομένων από μόνο ένα υγιές εξεταζόμενο καθιστά τη μελέτη αμφισβητήσιμη. Για το λόγο αυτό θεωρήθηκε ως κατώφλι διαχωρισμού "υγιών" - "ασθενών" η τιμή AHI 15 ε/ώ. Με τον τρόπο αυτό κατατάσσονται στην ίδια κατηγορία οι υγιείς εξεταζόμενοι και αυτοί με ήπιο έως μέτριο OSAHS (Ομάδα 1 - O1) και θεωρούνται "ασθενείς" οι εξεταζόμενοι με σοβαρό OSAHS (Ομάδα 2 - O2). Εξαιρούνται και πάλι από την ανάλυση με καμπύλες ROC οι εξεταζόμενοι 1 και 3 λόγω των μεγάλων αποκλίσεων που παρουσιάζουν οι συχνότητες ενδιαφέροντος (Πίνακας 5.2). Η ανάλυση αυτή αποσκοπεί στη εξέταση της ικανότητας διάγνωσης και διαχωρισμού ατόμων με σοβαρό OSAHS από τους υγιείς ή πάσχοντες από ήπιο - μέτριο OSAHS.

5.4.1: Στατιστική ανάλυση F1 (Κριτήριο $AHI \geq 15$ ε/ώ)

Στην Εικόνα 5.14 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F1 και στην Εικόνα 5.15 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.9 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.14: Καμπύλη ROC για F1 ($AHI \geq 15$ ε/ώ)



Εικόνα 5.15: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (F1)

Πίνακας 5.9: Αποτελέσματα ανάλυσης ROC για F1 (AHI ≥ 15 ε/ώ)

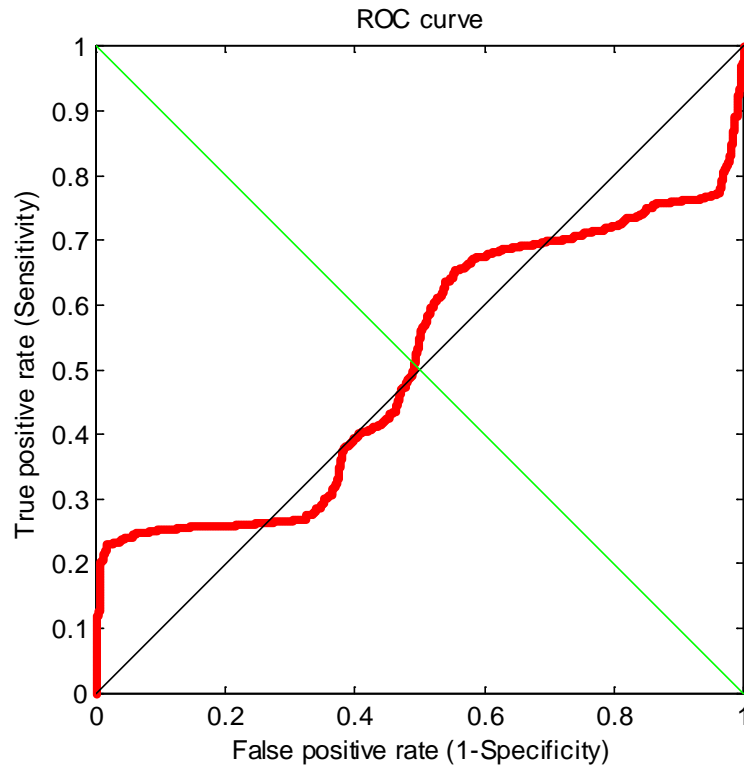
	n	A.P.	AHI \pm STD (ε/ώ)	F1 \pm STD (Hz)	Κατώφλι F1 (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
O1	5	900	56.90 \pm 28.85	449 \pm 330	239	60.83	74.8	40.8	Ανεπαρκής
O2	4	800	9.93 \pm 5.85	273 \pm 119					
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	58.18	71.9	37.3	
					Άνω όριο	63.49	77.6	44.2	

O1: Ομάδα 1, O2: Ομάδα 2, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

Η ανάλυση της F1 με κριτήριο διαχωρισμού AHI ≥ 15 ε/ώ δίνει αποτελέσματα που κρίνουν τη μελέτη ανεπαρκή. Συγκεκριμένα οι τιμές ευαισθησίας, ειδικότητας και AUC είναι αρκετά χαμηλές (74.8%, 40.8% και 60.8% αντίστοιχα) καθιστώντας το αποτέλεσμα μη αξιόπιστο. Σημειώνεται ότι η ευρεθείσα τιμή κατωφλίου διαχωρισμού για την F1, 239 Hz, πάλι απέχει από την αντίστοιχη τιμή που βρέθηκε στην [19] (470 Hz). Ο διαχωρισμός Ομάδας 1 και Ομάδας 2 δεν μπορεί να γίνει επομένως με τη βοήθεια της F1.

5.4.2: Στατιστική ανάλυση F2 (Κριτήριο AHI \geq 15 ε/ώ)

Στην Εικόνα 5.16 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F2. Στον Πίνακα 5.10 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.16: Καμπύλη ROC για F2 (AHI \geq 15 ε/ώ)

Πίνακας 5.10: Αποτελέσματα ανάλυσης ROC για F2 (AHI \geq 10 ε/ώ)

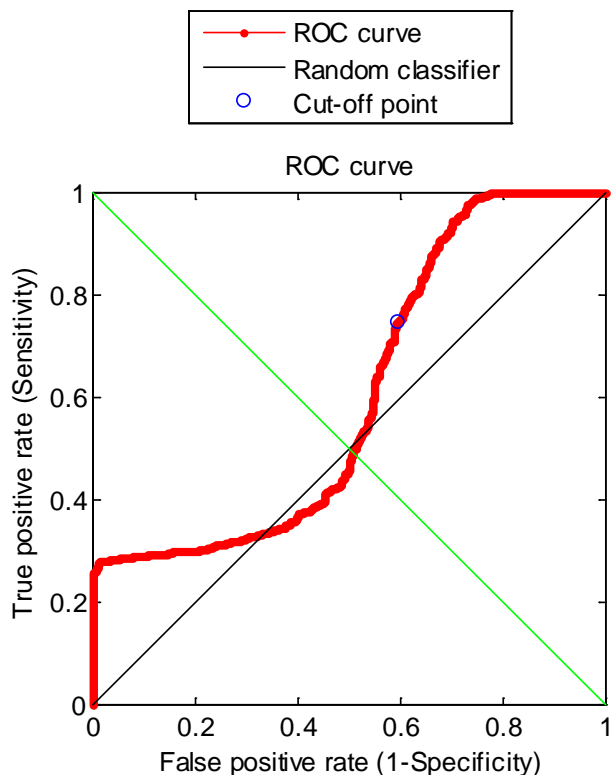
	n	A.P.	AHI \pm STD (ε/ώ)	F2 \pm STD (Hz)	Κατώφλι F2 (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
O1	5	900	56.90 \pm 28.85	1266 \pm 525	-	50.98	-	-	Αποτυχημένη
O2	4	800	9.93 \pm 5.85	1188 \pm 311	-	-	-		
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	48.23	-	-	
					Άνω όριο	53.73	-	-	

O1: Ομάδα 1, O2: Ομάδα 2, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

Η ανάλυση ROC για την F2 με κριτήριο κατηγοριοποίησης AHI \geq 15 ε/ώ δεν έδωσε αποτελέσματα και κρίνεται αποτυχημένη (AUC: 51%). Αυτό οφείλεται στην ομοιότητα που παρουσιάζουν οι συχνότητες F2 όπως φαίνεται στον Πίνακα 5.10 (Μέση τιμή \pm Τυπική απόκλιση: Απνοϊκοί - 1266 \pm 525, Υγιείς - 1188 \pm 311). Η F2 κρίνεται επομένως μη ικανή για το διαχωρισμό Ομάδας 1 και Ομάδας 2.

5.4.3: Στατιστική ανάλυση F3 (Κριτήριο AHI \geq 15 ε/ώ)

Στην Εικόνα 5.17 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F3. Στον Πίνακα 5.11 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.17: Καμπύλη ROC για F3 (AHI \geq 15 ε/ώ)

Πίνακας 5.11: Αποτελέσματα ανάλυσης ROC για F3 (AHI \geq 15 ε/ώ)

	n	A.P.	AHI \pm STD (ε/ώ)	F3 \pm STD (Hz)	Κατώφλι F3 (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
O1	5	900	56.90 \pm 28.85	2280 \pm 572	-	43.45	-	-	Αποτυχημένη
O2	4	800	9.93 \pm 5.85	2084 \pm 131	-	-	-	-	
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	40.72	-	-	
					Άνω όριο	46.18	-	-	

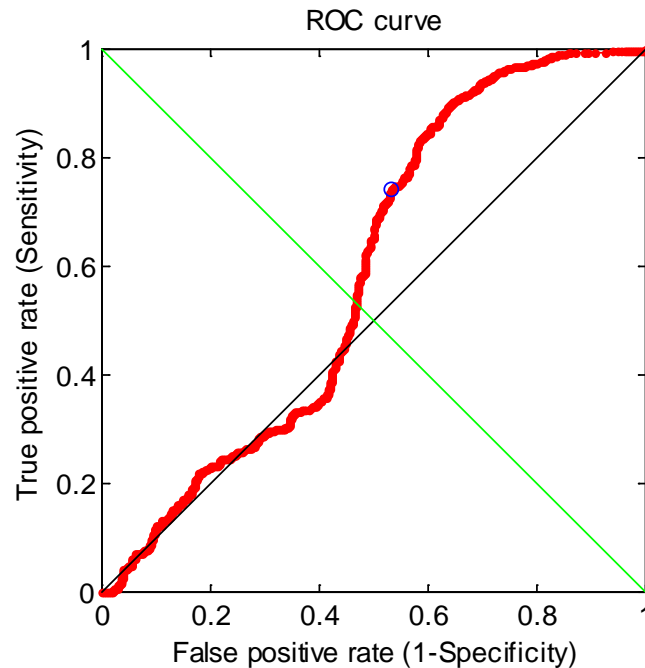
O1: Ομάδα 1, O2: Ομάδα 2, n: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

Η ανάλυση ROC για την F3 με κριτήριο κατηγοριοποίησης AHI \geq 15 ε/ώ δεν έδωσε αποτελέσματα και κρίνεται αποτυχημένη. Αυτό οφείλεται, όπως και στην περίπτωση κατηγοριοποίησης AHI \geq 10 ε/ώ, στην ομοιότητα που παρουσιάζουν οι συχνότητες F3 όπως φαίνεται στον Πίνακα 5.11 (Μέση τιμή \pm Τυπική απόκλιση: Απνοϊκοί - 2280 \pm 572,

Υγείς - 2084 ± 131). Η F3 κρίνεται επομένως αναξιόπιστη για το διαχωρισμό Ομάδας 1 και Ομάδας 2.

5.4.4: Στατιστική ανάλυση $f_{50\%}$ (Κριτήριο $AHI \geq 15$ ε/ώ)

Στην Εικόνα 5.18 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων $f_{50\%}$. Στον Πίνακα 5.12 παραθέτονται τα αποτελέσματα της ανάλυσης.



Εικόνα 5.18: Καμπύλη ROC για $f_{50\%}$ ($AHI \geq 15$ ε/ώ)

Πίνακας 5.12: Αποτελέσματα ανάλυσης ROC για $f_{50\%}$ ($AHI \geq 15$ ε/ώ)

	<i>n</i>	A.P.	AHI ± STD (ε/ώ)	$f_{50\%}$ ± STD (Hz)	Κατώφλι $f_{50\%}$ (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
O1	5	900	56.90 ± 28.85	365 ± 228	-	58.91	-	-	Αποτυχημένη
O2	4	800	9.93 ± 5.85	334 ± 321	-	-	-		
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	56.22	-	-	
					Άνω όριο	61.60	-	-	

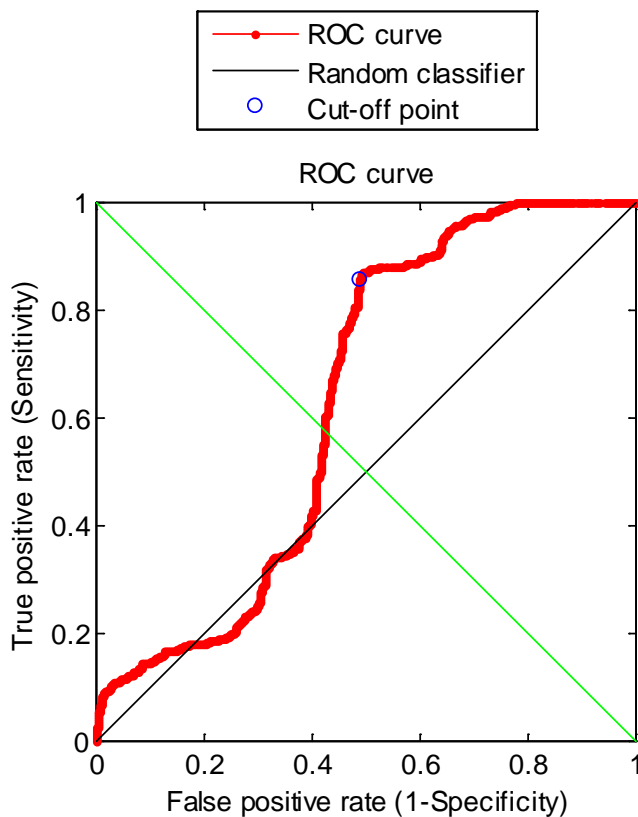
O1: Ομάδα 1, O2: Ομάδα 2, *n*: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα

Η ανάλυση ROC για την F3 με κριτήριο κατηγοριοποίησης $AHI \geq 15$ ε/ώ δεν έδωσε αποτελέσματα και κρίνεται αποτυχημένη. Αυτό οφείλεται στην ομοιότητα που παρουσιάζουν οι συχνότητες $f_{50\%}$ όπως φαίνεται στον Πίνακα 5.12 (Μέση τιμή ± Τυπική απόκλιση: Απνοϊκοί - 365 ± 228 , Υγείς - 334 ± 321) και στην μεγάλη τυπική απόκλιση που

έχουν οι συχνότητες $f_{50\%}$ για την O2. Η $f_{50\%}$ κρίνεται επομένως αναξιόπιστη για το διαχωρισμό Ομάδας 1 και Ομάδας 2.

5.4.5: Στατιστική ανάλυση $f_{90\%}$ (Κριτήριο $AHI \geq 15$ ε/ώ)

Στην Εικόνα 5.19 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων $f_{90\%}$ και στην Εικόνα 5.20 η κατανομή των TP, TN, FP και FN. Στον Πίνακα 5.13 παραθέτονται τα αποτελέσματα της ανάλυσης.

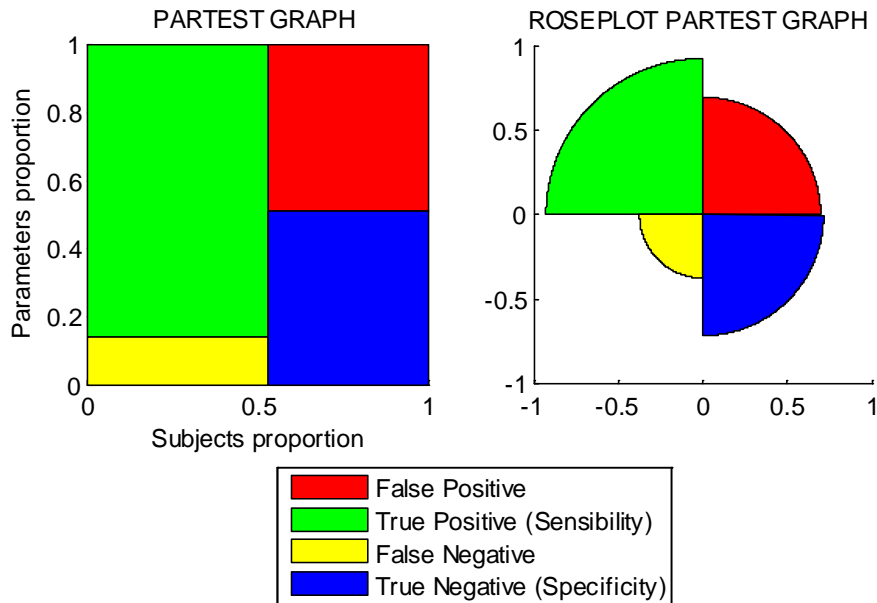


Εικόνα 5.19: Καμπύλη ROC για $f_{90\%}$ ($AHI \geq 15$ ε/ώ)

Πίνακας 5.13: Αποτελέσματα ανάλυσης ROC για $f_{90\%}$ ($AHI \geq 15$ ε/ώ)

	n	A.P.	$AHI \pm STD$ (ε/ώ)	$f_{90\%} \pm STD$ (Hz)	Κατώφλι $f_{90\%}$ (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
O1	5	900	56.90 ± 28.85	2154 ± 1820	526	63.37	86.0	51.2	Ανεπαρκής
O2	4	800	9.93 ± 5.85	1887 ± 1980					
Διάστημα Εμπιστοσύνης 95%									
					Άνω όριο	65.98	88.3	54.7	

O1: Ομάδα 1, O2: Ομάδα 2, n : Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα



Εικόνα 5.20: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN ($f_{90\%}$)

Η ανάλυση της $f_{90\%}$ με κριτήριο διαχωρισμού $AHI \geq 15$ ε/ώ δίνει αποτελέσματα που κρίνουν τη μελέτη ανεπαρκή. Συγκεκριμένα οι τιμές ειδικότητας και AUC είναι αρκετά χαμηλές (51.2% και 63.3% αντίστοιχα) καθιστώντας το αποτέλεσμα μη αξιόπιστο. Ο λόγος για το αποτέλεσμα αυτό πιθανόν να είναι οι πολύ μεγάλες τιμές τυπικής απόκλισης των συχνοτήτων $f_{90\%}$ όπως φαίνεται στον Πίνακα 5.13. Καταλήγουμε στο συμπέρασμα ότι η $f_{90\%}$ δεν είναι αξιοποιήσιμη για το διαχωρισμό Ομάδας 1 και Ομάδας 2.

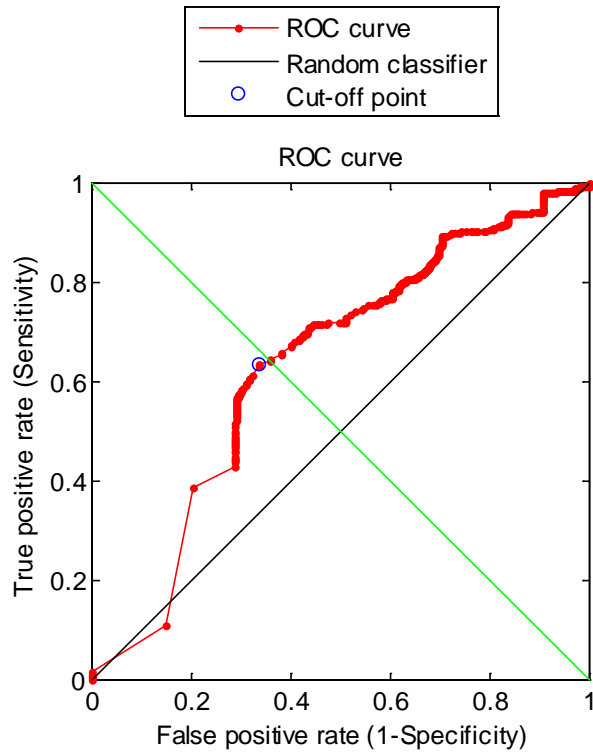
5.4.6: Στατιστική ανάλυση f_{peak} (Κριτήριο $AHI \geq 15$ ε/ώ)

Στην Εικόνα 5.21 παρουσιάζεται η καμπύλη ROC που προκύπτει από την ανάλυση των συχνοτήτων F2 και στην Εικόνα 5.22 η κατανομή των TP, TN, FP και FN που προκύπτουν. Στον Πίνακα 5.14 παραθέτονται τα αποτελέσματα της ανάλυσης.

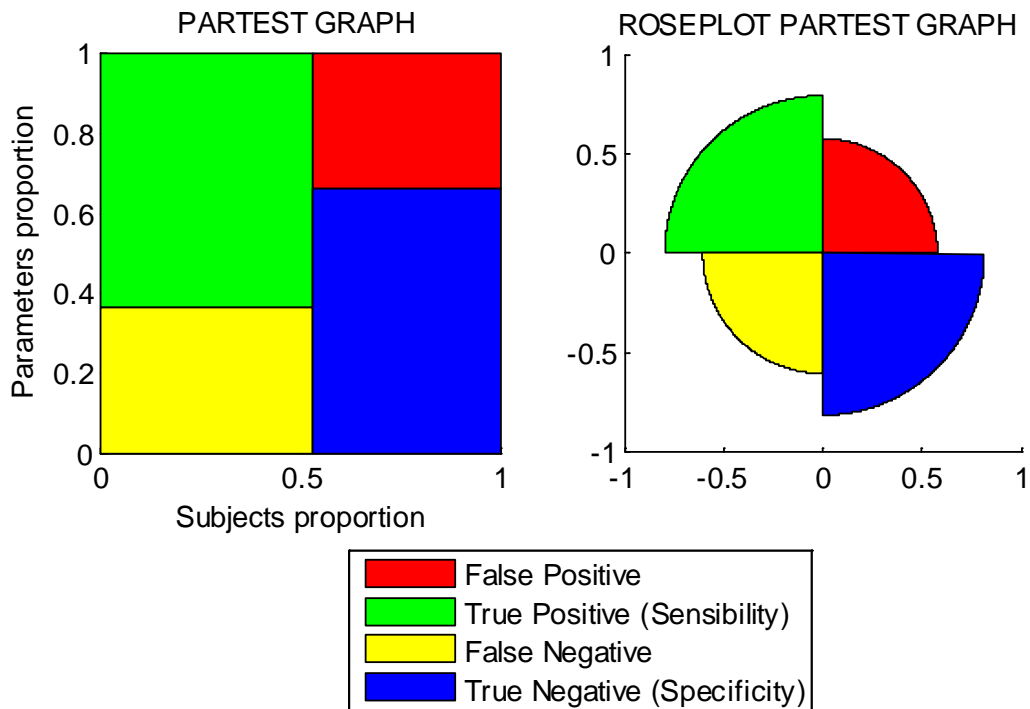
Πίνακας 5.14: Αποτελέσματα ανάλυσης ROC για f_{peak} ($AHI \geq 15$ ε/ώ)

	<i>n</i>	A.P.	AHI ± STD (ε/ώ)	$f_{peak} \pm$ STD (Hz)	Κατώφλι f_{peak} (Hz)	AUC (%)	Sens (%)	Spec (%)	Ποιότητα Ανάλυσης
O1	5	900	56.90 ± 28.85	121 ± 125	87	63.7	63.4	66.5	Ανεπαρκής
O2	4	800	9.93 ± 5.85	185 ± 162					
Διάστημα Εμπιστοσύνης 95%					Κάτω όριο	61.06	60.3	63.2	
					Άνω όριο	66.28	66.6	69.8	

O1: Ομάδα 1, O2: Ομάδα 2, *n*: Αριθμός εξεταζομένων, A.P.: Αριθμός ροχαλητών που αναλύθηκαν, Sens: Ευαισθησία, Spec: Ειδικότητα



Εικόνα 5.21: Καμπύλη ROC για f_{peak} ($AHI \geq 15$ ε/ώ)



Εικόνα 5.22: Γραφικές παραστάσεις της κατανομής των TP, TN, FP και FN (f_{peak})

Η ανάλυση της f_{peak} με κριτήριο διαχωρισμού $AHI \geq 15$ ε/ώ δίνει αποτελέσματα που κρίνουν τη μελέτη ανεπαρκή. Συγκεκριμένα οι τιμές ευαισθησίας, ειδικότητας και AUC είναι αρκετά χαμηλές (63.4%, 66.5% και 63.7% αντίστοιχα) καθιστώντας το αποτέλεσμα

μη αξιόπιστο. Ο λόγος για το αποτέλεσμα αυτό είναι οι πολύ μεγάλες τιμές τυπικής απόκλισης των συχνοτήτων f_{peak} όπως φαίνεται στον Πίνακα 5.14 και η μικρή διαφορά των μέσων τους τιμών. Καταλήγουμε στο συμπέρασμα ότι η f_{peak} δεν είναι αξιοποιήσιμη για το διαχωρισμό Ομάδας 1 και Ομάδας 2.

5.5: Σύνοψη αποτελεσμάτων

Στον Πίνακα 5.15 συνοψίζονται τα αποτελέσματα των 12 αναλύσεων με καμπύλες ROC που προηγήθηκαν.

Πίνακας 5.15: Συγκεντρωτικός πίνακας αποτελεσμάτων ανάλυσης με καμπύλες ROC

Μέγεθος	Όριο AHI (ε/ώ)	n A/Y	A.P. A/Y	AHI ± STD (ε/ώ) A/Y	AUC (%)	Sens (%)	Spec (%)	Ποιότητα ανάλυσης	Κατώφλι (Hz)
F1	10	8/1	1500/200	40.3 ± 31.6/ 1.6 ± 0	99.8	97.7	99.5	Εξαιρετική	161
F2					65.4	45.0	92.0	Ανεπαρκής	1220
F3					49.8	-	-	Αποτυχημένη	-
$f_{50\%}$					99.2	96.0	98.0	Εξαιρετική	113
$f_{90\%}$					99.0	94.5	97.0	Εξαιρετική	344
f_{peak}					66.7	48.0	84.5	Ανεπαρκής	89
F1	15	5/4	900/800	56.90 ± 28.85/ 9.93 ± 5.85	60.8	74.8	40.8	Ανεπαρκής	239
F2					51.0	-	-	Αποτυχημένη	-
F3					43.4	-	-	Αποτυχημένη	-
$f_{50\%}$					58.9	-	-	Αποτυχημένη	-
$f_{90\%}$					63.3	86.0	51.2	Ανεπαρκής	526
f_{peak}					63.7	63.4	66.5	Ανεπαρκής	87

Στον Πίνακα 5.16 παραθέτεται το υπολογιστικό κόστος των διαφόρων σταδίων επεξεργασίας. Το υπολογιστικό κόστος για τη μελέτη επίδρασης ρινικού αποσυμφορητικού σε συνδυασμό με κορτικοστεροειδή (Μελέτη 1) αφορά ανάλυση ηχογράφησης διάρκειας μιας ώρας ενώ η ανάλυση για τη μελέτη της διαγνωστικής ικανότητας των FF, $f_{50\%}$, $f_{90\%}$ και f_{peak} (Μελέτη 2) αφορά ανάλυση ηχογράφησης που περιέχει 200 ροχαλητά.

Κεφάλαιο 5

Πίνακας 5.16: Υπολογιστικό κόστος αλγορίθμων

	Λειτουργία Αλγορίθμου	Χρόνος (secs)	Χαρακτηριστικά σήματος προς ανάλυση	Πρόγραμμα ανάλυσης
Μελέτη 1	Αποθορυβοποίηση	736	Πραγματική ολονύκτια ηχογράφηση διάρκειας 7 ωρών και 14 λεπτών	Audacity® 1.3.14 - beta
	Αποθήκευση	68	Σήμα που περιέχει μόνο ροχαλητά διάρκειας 2 ωρών και 41 λεπτών	Audacity® 1.3.14 - beta
	Φόρτωση, κατάτμηση και αποθήκευση μεμονωμένων ροχαλητών	164.26	Εξαγωγή σε μορφότυπο wav Αποθορυβοποιημένο σήμα 2829 ροχαλητών διάρκειας 2 ωρών και 41 λεπτών	MATLAB® R2011b 7.13.0.564
	Υπολογισμός μεγεθών μέσης, μέγιστης έντασης και συχνότητας ροχαλητών και γραφικές παραστάσεις	1683.16	Εξαγωγή χαρακτηριστικών από 2829 ροχαλητά	MATLAB® R2011b 7.13.0.564
Μελέτη 2	Αποθορυβοποίηση	736	Πραγματική ολονύκτια ηχογράφηση διάρκειας 7 ωρών και 14 λεπτών	Audacity® 1.3.14 - beta
	Αποθήκευση	68	Σήμα που περιέχει 304 ροχαλητά διάρκειας 49 λεπτών	Audacity® 1.3.14 - beta
	Φόρτωση, κατάτμηση και αποθήκευση μόνο 200 μεμονωμένων ροχαλητών	24.905	Εξαγωγή σε μορφότυπο wav Αποθορυβοποιημένο σήμα 304 ροχαλητών διάρκειας 49 λεπτών	MATLAB® R2011b 7.13.0.564
	Φόρτωση συνολικού πίνακα ροχαλητών και εξαγωγή F1, F2, F3	28.723	2873 σήματα ροχαλητών από 11 εξεταζόμενους	MATLAB® R2011b 7.13.0.564
	Φόρτωση συνολικού πίνακα ροχαλητών και εξαγωγή $f_{50\%}$, $f_{90\%}$, f_{peak}	10.107	2873 σήματα ροχαλητών από 11 εξεταζόμενους	MATLAB® R2011b 7.13.0.564
	Εξαγωγή μέσης τιμής και τυπικής απόκλισης για τις F1, F2, F3, $f_{50\%}$, $f_{90\%}$, f_{peak}	0.125	Συχνότητες ενδιαφέροντος από 2100 σήματα ροχαλητών από 11 εξεταζόμενους	MATLAB® R2011b 7.13.0.564
	Διαχωρισμός εξεταζομένων σε υγιείς - απνοϊκούς	0.013	Συχνότητες ενδιαφέροντος από 2100 σήματα ροχαλητών από 11 εξεταζόμενους	MATLAB® R2011b 7.13.0.564
	Ανάλυση με καμπύλες ROC και γραφικές παραστάσεις	1.311	1700 τιμές F1	MATLAB® R2011b 7.13.0.564

Η εξαγωγή των αποτελεσμάτων της Μελέτης όπως φαίνεται στον Πίνακα 5.16 απαιτεί πάρα πολύ χρόνο και υπολογισμούς. Αυτό οφείλεται κυρίως στον μεγάλο όγκο δεδομένων που αναλύεται, αφού γίνεται επεξεργασία μεγάλων σημάτων ήχου υψηλής

ποιότητας με υψηλή συχνότητα δειγματοληψίας ($f_s = 16000$ Hz) μορφότυπου wav. Σημειώνεται ότι το αρχικό μέγεθος του αρχείου που αναλύεται είναι 3.15 GB.

Στην μελέτη της διαγνωστικής ικανότητας των FF, $f_{50\%}$, $f_{90\%}$ και f_{peak} αναλύονται λιγότερα σήματα ροχαλητών γι' αυτό και απαιτούνται μικρότεροι χρόνοι υπολογισμού. Το κύριο ποσοστό του συνολικού χρόνου ανάλυσης αφορά την προεπεξεργασία (αποθορυβοποίηση και αποθήκευση) των αρχικών ογκωδών σημάτων ροχαλητών.

Κεφάλαιο 6: Συζήτηση - Συμπεράσματα

Τα τελευταία χρόνια έχει αυξηθεί το ενδιαφέρον των ερευνητών για τις διαταραχές κατά τη διάρκεια του ύπνου και ιδιαίτερα για το OSAHS. Κύριο αντικείμενο μελέτης αποτελεί η προσπάθεια διάγνωσης του OSAHS χωρίς την ανάγκη PSG, με πρώτο εργαλείο τα ακουστικά σήματα ροχαλητού. Στη συγκεκριμένη διπλωματική εργασία αναπτύχθηκε ένα εργαλείο εξαγωγής χρονοσυχνοτικών χαρακτηριστικών σημάτων ροχαλητού, το Snor-Lux. Ακολούθως χρησιμοποιήθηκαν οι αλγόριθμοι που το αποτελούν για την εξαγωγή χαρακτηριστικών πραγματικών σημάτων ροχαλητού για τις ανάγκες δύο κλινικών μελετών. Η πρώτη εξυπηρετεί τις ανάγκες μιας έρευνας υπό διεξαγωγή στο νοσοκομείο "Ευαγγελισμός" και για τους σκοπούς αυτής εξήχθησαν η μέση και μέγιστη ένταση και η συχνότητα ροχαλητού. Η δεύτερη αποσκοπεί στην εξέταση της διαγνωστικής ικανότητας για το OSAHS, έξι συχνοτήτων που χαρακτηρίζουν τα σήματα ροχαλητού. Στο κεφάλαιο αυτό γίνεται σχολιασμός των αποτελεσμάτων που παρουσιάστηκαν στο προηγούμενο κεφάλαιο και εξήγηση των περιορισμών και δυσκολιών που συναντήθηκαν. Στη συνέχεια εκτιμούνται οι μελλοντικές προοπτικές της εργασίας και διατυπώνονται προτάσεις για την αξιοποίηση και επέκτασή της. Τέλος διατυπώνεται ένα γενικό συμπέρασμα που προκύπτει από το σύνολο της διπλωματικής εργασίας.

6.1: Γενικά σχόλια μελέτης επίδρασης συνδυασμού ρινικού αποσυμφορητικού με κορτικοστεροειδή

Η εξαγωγή μέσης έντασης, μέγιστης έντασης και συχνότητας για τις ανάγκες της [45] έγινε σε τρεις ολονύκτιες ηχογραφήσεις ατόμων στο ΚΜΥ του νοσοκομείου "Ευαγγελισμός". Τα σήματα αρχικά αποθρομβοποιήθηκαν και στη συνέχεια αφαιρέθηκαν χειροκίνητα τμήματα που περιείχαν ήχους άλλους από ροχαλητά. Ακολούθως έγινε κατάτμηση των σημάτων σε επιμέρους ροχαλητά και εξήχθησαν από αυτά τα μεγέθη της μέσης και μέγιστης έντασης. Με δεδομένο το συνολικό χρόνο ύπνου από τη πολυσωματογραφική μελέτη υπολογίστηκε η συχνότητα ροχαλητού. Λόγω του μεγάλου όγκου δεδομένων προς ανάλυση για κάθε ηχογράφιση έγινε βελτιστοποίηση των αλγορίθμων υπολογισμού των ζητούμενων μεγεθών για μείωση του υπολογιστικού κόστους και ταχύτερη εξαγωγή αποτελεσμάτων.

Ο σχολιασμός των αποτελεσμάτων της επίδρασης συνδυασμού ρινικού αποσυμφορητικού με κορτικοστεροειδή δυστυχώς δεν είναι δυνατός αφού οι τρεις αυτές καταγραφές αποτελούν ένα πολύ μικρό ποσοστό των συνολικών ηχογραφήσεων που θα αναλυθούν για της ανάγκες της [45]. Λόγω της διαδικασίας που απαιτεί το πρωτόκολλο (Εικόνα 4.1) απαιτείται μεγάλος χρόνος για τη συλλογή όλων των δεδομένων, ενώ η μελέτη έχει ξεκινήσει το Φεβρουάριο του 2012. Σημειώνεται επίσης ότι έχουν γίνει δύο ακόμα ηχογραφήσεις ατόμων στο ΚΜΥ, που εμπίπτουν στην [45] όμως λόγω εσφαλμένης συνδεσμολογίας καλωδίων οι ηχογραφήσεις περιείχαν μόνο θόρυβο και δεν ήταν δυνατή η ανάλυσή τους. Όπως έχει αναφερθεί, θα συνεχιστεί η συνεργασία με το ΚΜΥ μέχρι το πέρας της έρευνας και τα αποτελέσματα θα παρουσιαστούν σε δημοσίευση. Απώτερος σκοπός των μεγεθών που εξάγονται είναι η αξιολόγηση της επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή σε ασθενείς με OSAHS.

6.2: Γενικά σχόλια μελέτης διαγνωστικής ικανότητας FF, $f_{50\%}$, $f_{90\%}$ και f_{peak}

Τα δεδομένα εξέτασης της διαγνωστικής ικανότητας για το OSAHS των έξι συχνοτήτων: F1, F2, F3, $f_{50\%}$, $f_{90\%}$ και f_{peak} προέρχονται από 11 καταγραφές ατόμων στο ΚΜΥ του νοσοκομείου "Ευαγγελισμός". Αρχικά έγινε αποθρομβοποίηση των σημάτων και αφαίρεση κάθε ήχου που δεν ήταν ροχαλητό. Στη συνέχεια τα σήματα κατατμήθηκαν σε επιμέρους ροχαλητά και υπολογίστηκαν οι έξι συχνότητες ενδιαφέροντος. Με μια απλή

επισκόπηση των συχνοτήτων και με βάση την ακουστική πληροφορία των σημάτων αφαιρέθηκαν δύο καταγραφές λόγω της μεγάλης απόκλισης που παρουσίαζαν οι συχνότητές τους αλλά και επειδή στις συγκεκριμένες καταγραφές αντί ροχαλητών οι εξεταζόμενοι απλώς ανάσπιναν βαριά.

Ο διαχωρισμός των ατόμων σε υγιείς - απνοϊκούς έγινε αρχικά με βάση το κριτήριο $AHI \geq 10$ ε/ώ. Σε αυτή την περίπτωση όμως στην ομάδα υγιών ατόμων ενέπιπτε ένα μόνο άτομο. Στα δεδομένα έγινε ακολούθως ανάλυση καμπυλών ROC. Τα αποτελέσματα του Κεφαλαίου 5.3 φαίνονται ενθαρρυντικά όσον αφορά τη διάγνωση OSAHS από τις ηχογραφήσεις ροχαλητού αφού τα τρία από τα έξι μεγέθη που αναλύθηκαν έδωσαν πολύ καλά αποτελέσματα. Συγκεκριμένα ως υποψήφιος για τη διάγνωση του OSAHS προέκυψαν οι F1 με ευαισθησία 97.7%, ειδικότητα 99.5% και κατώφλι διαχωρισμού τα 161 Hz, η $f_{50\%}$ με ευαισθησία 96%, ειδικότητα 98% και κατώφλι διαχωρισμού τα 113 Hz και η $f_{90\%}$ με ευαισθησία 94.5%, ειδικότητα 97% και κατώφλι διαχωρισμού τα 344 Hz. Τα υπόλοιπα μεγέθη που αναλύθηκαν (F2, F3, f_{peak}) απέτυχαν να δώσουν ικανοποιητικά αποτελέσματα και κρίνονται επομένως ακατάλληλα για τη διάγνωση του OSAHS.

Λόγω της ύπαρξης μόνον ενός υγιούς εξεταζομένου με βάση το κριτήριο $AHI \geq 10$ ε/ώ πραγματοποιήθηκε μια δεύτερη εξέταση των δεδομένων ανεβάζοντας το κριτήριο κατηγοριοποίησης στο $AHI \geq 15$ ε/ώ. Με τον τρόπο αυτό ενσωματώθηκαν σε μια ομάδα (Ομάδα 1) οι υγιείς και οι πάσχοντες από ήπιο - μέτριο OSAHS ενώ τη δεύτερη ομάδα αποτελούν οι ασθενείς με σοβαρό OSAHS. Πραγματοποιήθηκε ανάλυση ROC στα δεδομένα της δεύτερης κατηγοριοποίησης και τα αποτελέσματα παρουσιάστηκαν στο Κεφάλαιο 5.4. Όπως παρατηρείται κανένα από τα μεγέθη εξέτασης δεν έδωσε ικανοποιητικά αποτελέσματα ώστε να μπορεί να θεωρηθεί ικανό για την διάγνωση του OSAHS. Το γεγονός αυτό δεν αναιρεί τα αποτελέσματα της πρώτης κατηγοριοποίησης ($AHI \geq 10$ ε/ώ) αφού έγινε μεταβολή του υλικού (εξεταζομένων) και συμπεριλήφθηκαν στα υγιή άτομα εξεταζόμενοι με αρκετά υψηλές τιμές AHI (14.1 και 13.9 ε/ώ).

Σε όλες τις αναλύσεις ROC που απέτυχαν τα μεγέθη παρουσίαζαν μεγάλες τιμές τυπικής απόκλισης ενώ σε επτά από τις εννιά αναλύσεις που απέτυχαν οι μέσες τιμές των συχνοτήτων ήταν πολύ κοντά μεταξύ των δύο ομάδων ανάλυσης, καθιστώντας αδύνατο το διαχωρισμό τους (F2 - $AHI \geq 10$ ε/ώ, F1 - $AHI \geq 15$ ε/ώ, F2 - $AHI \geq 15$ ε/ώ, F3 - $AHI \geq 15$ ε/ώ, $f_{50\%}$ - $AHI \geq 15$ ε/ώ, $f_{90\%}$ - $AHI \geq 15$ ε/ώ, f_{peak} - $AHI \geq 15$ ε/ώ). Οι μεγάλες τιμές τυπικής

απόκλισης οφείλονται κατά κύριο λόγο στον μεγάλο αριθμό δεδομένων που εξετάστηκαν (200 ροχαλητά από κάθε εξεταζόμενο) σε σχέση με άλλες παρόμοιες μελέτες (10 στις [26] και [17], 40 στις [19] και [20], 80 στην [18]). Η ανάλυση περισσότερων ροχαλητών στην διπλωματική εργασία αποσκοπεί στην πιο ευρεία και πιο αξιόπιστη εξέταση της διαγνωστικής ικανότητας των μεγεθών ανάλυσης.

6.3: Μελλοντικές προοπτικές

Η μελέτη χρονοσυχνοτικών χαρακτηριστικών ακουστικών σημάτων ροχαλητού με στόχο τη διάγνωση του OSAHS έχει προσελκύσει αρκετούς ερευνητές όμως είναι ακόμα περιορισμένη. Η συνέχιση της μελέτης πέρα από το επιστημονικό της ενδιαφέρον επιβάλλεται ώστε να επιβεβαιωθεί η ικανότητα διάγνωσης του OSAHS μέσω κάποιων χρονοσυχνοτικών χαρακτηριστικών των σημάτων ροχαλητού λόγω των πολλών πλεονεκτημάτων που συνεπάγονται.

Η διάγνωση του συνδρόμου γίνεται προς το παρόν μόνο με PSG, το οποίο δεν είναι φιλικό προς τον εξεταζόμενο, είναι ιδιαίτερα χρονοβόρο απαιτεί ειδικό εξοπλισμό και διεξάγεται σε κέντρα μελέτης ύπνου. Οι εξεταζόμενοι κοιμούνται σε περιβάλλον άγνωστο προς αυτούς, συνδεδεμένοι με πολλά καλώδια καταγραφής βιοσημάτων με αποτέλεσμα οι καταγραφόμενες μετρήσεις να μην αντικατοπτρίζουν πάντοτε τα πραγματικά δεδομένα. Δεδομένου ότι υπάρχει περιορισμένος αριθμός ΚΜΥ σε λίγες περιοχές πολλά άτομα δεν μπορούν να εξεταστούν και κατά συνέπεια δεν λαμβάνουν γνώση ότι πάσχουν από OSAHS. Η επαλήθευση της ικανότητας διάγνωσης του συνδρόμου μόνον από σήματα ροχαλητού εξαλείφει την ανάγκη PSG και επομένως την ανάγκη διανυκτέρευσης του εξεταζόμενου σε ΚΜΥ. Σε τέτοια περίπτωση είναι δυνατή η ενσωμάτωση του συστήματος καταγραφής σημάτων ροχαλητού ακόμα και στο κινητό τηλέφωνο του ατόμου που επιθυμεί να εξεταστεί και η αυτόματη αποστολή των δεδομένων προς ανάλυση σε εξειδικευμένο προσωπικό, χωρίς καν να είναι ανάγκη να φύγει ποτέ από το σπίτι του ο εξεταζόμενος μειώνοντας δραματικά το κόστος διάγνωσης.

Όσο αφορά τις μελλοντικές προοπτικές της διπλωματικής εργασίας πρέπει καταρχάς να προσανατολιστούν στην επανάληψη της εξέτασης της διαγνωστικής ικανότητας των FF , $f_{50\%}$, $f_{90\%}$ και f_{peak} ώστε να είναι πιο αξιόπιστη η μελέτη. Θα πρέπει να συμπεριληφθούν

περισσότερες ηχογραφήσεις, ειδικότερα υγιών ατόμων με $AHI \leq 10$ ε/ώ αφού ο ένας μόνο υγιής εξεταζόμενος δεν εγγυάται την ορθότητα της μελέτης.

Περαιτέρω μελέτη απαιτεί και η αποθορυβοποίηση των σημάτων ροχαλητού στο στάδιο της προεπεξεργασίας, αφού οι αλγόριθμοι φασματικής αφαίρεσης που χρησιμοποιούνται ίσως να αφαιρούν από το φάσμα του κάθε ροχαλητού και πληροφορία που είναι χρήσιμη για την εξαγωγή των χρονοσυχνοτικών χαρακτηριστικών στη συνέχεια. Ως ερευνητικό πεδίο προτείνεται η αποθορυβοποίηση σήματος με τη βοήθεια κυματιδίων (wavelet denoising).

Τα μεγέθη που εξετάζονται ως προς την διαγνωστική τους ικανότητα για το OSAHS μπορούν να διευρυνθούν. Γι' αυτό μπορεί να βοηθήσει το πακέτο Snor-Lux το οποίο παρέχει αλγορίθμους για διάφορες επεξεργασίες σημάτων ροχαλητού.

Βελτίωση επιδέχονται και οι αλγόριθμοι που χρησιμοποιήθηκαν στην διπλωματική εργασία, ως προς το υπολογιστικό τους κόστος. Ειδικότερα, μπορεί να βελτιωθεί το υπολογιστικό κόστος των αλγορίθμων αποθορυβοποίησης και εξαγωγής των ζητούμενων μεγεθών. Η βελτίωση μπορεί να γίνει με την παραλληλοποίηση των αλγορίθμων αυτών για αξιοποίηση των πολλαπλών πυρήνων και νημάτων που διαθέτουν οι πλείστοι προσωπικοί υπολογιστές σήμερα.

Ενδιαφέρον θα έχει η μελέτη ροχαλητών από συγκεκριμένες φάσεις του ύπνου (NON-REM, REM1, REM2, REM3) και με γνωστές περαιτέρω πληροφορίες (π.χ. θέση ύπνου: ύπτια, μπρούμυτα, πλάγια). Για τη συγκεκριμένη μελέτη θα πρέπει να γίνει ενσωμάτωση της πληροφορίας του PSG στην ηχογράφιση ροχαλητών, κάτι το οποίο δεν είναι δυνατό τώρα στο ΚΜΥ του νοσοκομείου "Ευαγγελισμός" λόγω του ότι τα δύο συστήματα καταγραφής είναι ανεξάρτητα.

6.4: Συμπέρασμα

Η παρούσα διπλωματική εργασία ασχολείται με ακουστικά σήματα ροχαλητού και χρονοσυχνοτικών χαρακτηριστικών που εξάγονται από αυτά με σκοπό την αξιολόγηση της μελέτης επίδρασης ρινικών αποσυμφορητικών σε συνδυασμό με κορτικοστεροειδή και την προσπάθεια διάγνωσης του OSAHS μέσω των συχνοτήτων φωνοσυντονισμού και χαρακτηριστικών συχνοτήτων του φάσματος των σημάτων. Η υλοποίησή της περιλαμβάνει και συνδυάζει μεθοδολογία που χρησιμοποιήθηκε σε προηγούμενες

Κεφάλαιο 6

μελέτες, καθώς και μεθοδολογία που δεν έχει εφαρμοστεί στο παρελθόν. Το πόρισμα της μελέτης που αφορά τη διάγνωση του OSAHS είναι ότι η πρώτη συχνότητα φωνοσυντονισμού F1, η μέση συχνότητα $f_{50\%}$ και η μέγιστη συχνότητα $f_{90\%}$ μπορούν να διαγνώσουν το OSAHS χωρίς την ανάγκη PSG μέσω ηχογραφήσεων ροχαλητού. Η διεξαγωγή περαιτέρω έρευνας σε μεγαλύτερο όγκο δεδομένων είναι απαραίτητη για βελτίωση της αξιοπιστίας της μελέτης.

Βιβλιογραφία

- [1] D. Pevernagie, R. M. Aarts and M. D. Meyer, "The acoustics of snoring," *Sleep Med. Rev.*, pp. 131-144, 2010.
- [2] Boston Houghton Mifflin Company, *The American Heritage dictionary of the English language*, 4th ed., 2000.
- [3] K. Strohl and S. Redline, "Recognition of obstructive sleep apnea," *Am J Respir Crit Care Med*, no. 154, pp. 279-289, 1996.
- [4] Ελληνική Εταιρία Διαταραχών Ύπνου (ΕΛΕΔΥ), "Θέσεις ομοφωνίας για τη διάγνωση και θεραπεία των διαταραχών της αναπνοής στον ύπνο," *Πνεύμων*, vol. 22, 2009.
- [5] R. Albert, S. Spiro and J. Jett, *Κλινική Πνευμονολογία*, Εκδόσεις Πασχαλίδη, 2004.
- [6] W. Ruehland, P. Rochford, F. O'Donoghue, R. Pierce, P. Singh and A. Thornton, "The new AASM criteria for scoring hypopneas: Impact on the apnea hypopnea index," *Sleep*, vol. 11, no. 32, pp. 150-159, 2009.
- [7] Δ. Κυρούσης, Α. Οικονόμου and Ε. Πιπίνη, *Άπνοια στον ύπνο και καρδιαγγειακά νοσήματα*, Αθήνα: Εγκυκλοπαίδεια Υγείας, 2010.
- [8] E. Nena, V. Tsara and P. Steiropoulos, "Sleep-disordered breathing and quality of life of railway drivers in Greece," *Chest*, no. 134, pp. 79-86, 2008.
- [9] P. Deegan and W. McNicholas, "Pathophysiology of obstructive sleep apnoea," *Eur Respir J*, no. 8, pp. 1161-1178, 1995.
- [10] American Thoracic Society, "Severity of motor crashes in obstructive sleep apnea patients," in *ATS Int Conf*, 2007.
- [11] J. Dille, "Snoring can be fatal for your marriage and for you," *Aviat. Space Environ. Med.*, vol. 12, no. 58, p. 1234, 1987.
- [12] E. Pipini, A. Ikonou and D. Kyroussis, "Sleep apnea and cardiovascular disease," *Ιατρικά Ανάλεκτα*, no. 3, pp. 285-290, 2010.
- [13] Ε. Βαγιάκης, "Η πολυσωματοκαταγραφική μελέτη ύπνου (πολυυπνογραφία) και το εργαστήριο ύπνου," *Πνεύμων*, no. 2, 2004.
- [14] Χ. Μερμίγκης, Σ. Σχίζα and Ε. Βαγιάκης, "Θεραπευτική αντιμετώπιση αποφρακτικού συνδρόμου απνοιών-υποπνοιών κατά τον ύπνο (ΣΑΥΥ) με CPAP. Απαιτούμενος κλινικός-εργαστηριακός έλεγχος/ εμπλεκόμενοι φορείς/ προϋποθέσεις συνταγογράφησης συσκευών CPAP/ Τύποι συσκευών/ Παρακολούθηση ασθενών," *Ελληνική Πνευμονολογική Εταιρία*, Αθήνα, 2008.
- [15] J. A. Fiz, J. Abad, R. Jane, M. Riera, M. A. Mananas, P. Caminal, D. Rodenstein and J. Morera, "Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnoea," *Eur Respir J*, vol. 9, pp. 2365-2370, 1996.
- [16] U. Abeyratne, A. Wakwella and C. Hukins, "Pitch jump probability measures for the analysis of snoring sounds in apnea," *Physiol. Meas.*, vol. 26, pp. 779-798, 2005.
- [17] A. N. Keong, C. H. Wong and T. S. Koh, "Bispectral analysis of snore signals for obstructive sleep apnea detection," in *IEEE EMBS Conf.*, Lyon, 2007.

Βιβλιογραφία

- [18] D. Matsiki, X. Deligianni, E. Vlachogianni-Daskalopoulou and L. Hadjileontiadis, "Wavelet-based analysis of nocturnal snoring in apneic patients undergoing polysomnography," in *IEEE EMBS Conf.*, Lyon, 2007.
- [19] A. Keong, T. Koh, E. Baey, T. H. Lee, U. R. Abeyratne and K. Puvanendran, "Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?," *Sleep Med.*, vol. 9, pp. 894-898, 2008.
- [20] A. N. Keong, T. S. Koh, U. R. Abeyratne and K. Puvanendra, "Investigation of obstructive sleep apnea using interactions in nonstationary snore signals," *Annals of Biomed. Eng.*, vol. 37, no. 9, pp. 1796-1806, 2009.
- [21] A. Karunajeewa, U. Abeyratne and C. Hukins, "Multi-feature snore sound analysis in obstructive sleep apnea-hypopnea syndrome," *Physiol. Meas.*, vol. 32, pp. 83-97, 2010.
- [22] A. Sakakura, "Acoustic analysis of snoring sounds with chaos theory," *Int. Cong. Ser.*, pp. 227-230, 2003.
- [23] J. Sola-Soler, R. Jane, J. A. Fiz and J. Morera, "Spectral envelope analysis in snoring signals from simple snorers and patients with obstructive sleep apnea," in *IEEE EMBS Conf.*, Mexico, 2003.
- [24] W. D. Duckitt, S. K. Tuomi and T. R. Niesler, "Automatic detection, segmentation and assessment of snoring from ambient acoustic data," *Physiol. Meas.*, vol. 27, pp. 1047-1056, 2006.
- [25] M. Cavusoglu, M. Kamasak, O. Erogul, T. Ciloglu, Y. Serinagaoglu and T. Akcam, "An efficient method for snore/nonsnore classification of sleep sounds," *Physiol. Meas.*, vol. 28, pp. 841-853, 2007.
- [26] A. N. Keong, S. K. Tong, K. Puvanendran and U. R. Abeyratne, "Snore signal enhancement and activity detection via translation-invariant wavelet transform," *IEEE Trans. on Biomed. Eng.*, vol. 55, no. 10, 2008.
- [27] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 27, pp. 113-120, 1979.
- [28] M. Berouti, R. Schwartz and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," *Proc. of ICASSP*, pp. 208-211, 1979.
- [29] S. Kamath and P. Loizou, "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," in *Proc. of ICASSP*, Orlando FL, 2002.
- [30] P. Scalart and J. Vieira Filho, "Speech Enhancement based on a priori signal to noise estimation," *IEEE Intl. Conf Acoust. Speech, Signal Process.*, vol. 2, pp. 629-632, 1996.
- [31] C. Plapous, "Improved signal-to-noise ratio estimation for speech enhancement," *IEEE Trans. Acoust., Speech, Signal Process.*, 2006.
- [32] T. Giannakopoulos, "A method for silence removal and segmentation of speech signals, implemented in Matlab," University of Athens, Athens, 2009.
- [33] A. V. Oppenheim, A. S. Willsky and H. S. Nawab, *Signals and Systems*, Prentice Hall, 1997.
- [34] J. w. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, no. 19, pp. 297-301, 1965.
- [35] S. K. Banchhor and A. Khan, "Musical instrument recognition using zero crossing rate and short-time energy," *IJAIS*, vol. 1, no. 3, 2012.

- [36] R. Serway and J. Faughn, College physics, Fort Worth: Saunders College Publishing, 1991.
- [37] U. Greb and M. Rusbridge, "The interpretation of the bispectrum and bicoherence for non-linear interactions of continuous spectra," *Plasma Phys. Control. Fusion*, vol. 5, no. 30, pp. 537-549, 1988.
- [38] B. P. Bogert, M. J. R. Healy and J. W. Tukey, "The quefreny analysis of time series for echoes: cepstrum, pseudo autocovariance, cross-cepstrum and saphe cracking," in *Proc. of the Symposium on Time Series Analysis*, New York: Wiley, 1963.
- [39] M. A. Noll and M. R. Schroeder, "Cepstrum pitch determination," *J. Acoust. Soc. Am.*, vol. 41, no. 2, pp. 293-309, 1967.
- [40] X. Min, "HMM-based audio keyword generation," in *Adv. Multimed. Inf. Proces.*, 2004.
- [41] P. Robi, The wavelet tutorial, 2008.
- [42] S. Haykin, Advances in spectrum analysis and array processing, Prentice-Hall, 1991.
- [43] G. Fant, Acoustic theory of speech production, Netherlands: Mouton & Co., 1960.
- [44] R. C. Snell and F. Milinazzo, "Formant location from LPC analysis data," *IEEE Trans. Speech and Audio Process.*, vol. 1, no. 2, 1993.
- [45] E. Vagiakis and S. Zakynthinos, "The effect of nasal tramazoline with dexamethasone in obstructive sleep apnoea patients: a randomized placebo-controlled cross-over trial," Athens, 2011.
- [46] E. Βαλάντη, *Ανάλυση των καμπυλών ROC και εφαρμογή τους σε πραγματικά βιοϊατρικά δεδομένα*, Αθήνα: Ε.Μ.Π., 2011.

Παράρτημα Α: Εγχειρίδιο Λειτουργίας του πακέτου Snor-Lux

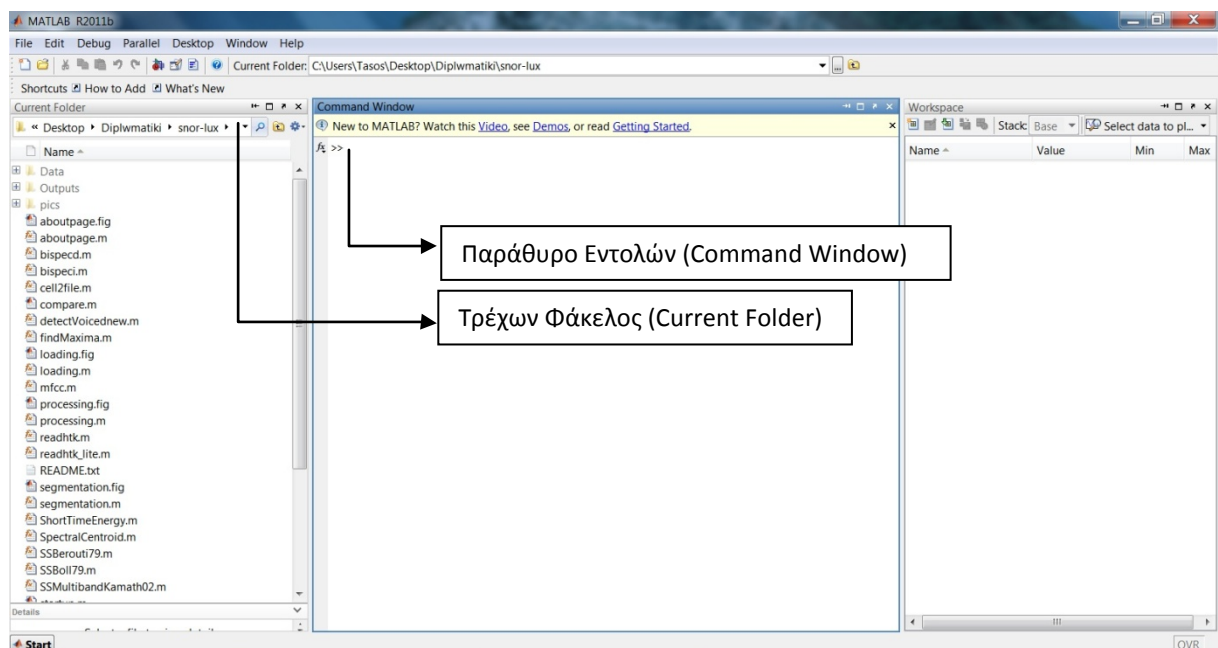
Στο κεφάλαιο αυτό παρουσιάζεται ένα σύντομο εγχειρίδιο χρήσης του γραφικού διαδραστικού περιβάλλοντος Snor-Lux για επεξεργασία ακουστικών σημάτων ροχαλητού.

A.1: Εισαγωγή

Το φιλικό προς το χρήστη πακέτο Snor-Lux αναπτύχθηκε στο πρόγραμμα MATLAB®, έκδοση R2011b. Αποτελεί ένα πακέτο γραφικού διαδραστικού περιβάλλοντος (GUI) και επικεντρώνεται στην επεξεργασία ακουστικών σημάτων ροχαλητού. Η επεξεργασία αυτή περιλαμβάνει αποθορυβοποίηση, κατάτμηση και εξαγωγή διαφόρων μεγεθών των σημάτων ροχαλητού, τα οποία αναφέρονται στο Κεφάλαιο 2. Σημειώνεται ότι η επεξεργασία αφορά αρχεία μορφότυπου *.wav. Ακολουθεί ένα σύντομο εγχειρίδιο λειτουργίας του Snor-Lux.

A.2: Εκκίνηση του Snor-Lux

Το Snor-Lux απαιτεί την εγκατάσταση του λογισμικού MATLAB® στον υπολογιστή στον οποίο θα χρησιμοποιηθεί. Για την εκκίνηση του γραφικού περιβάλλοντος γίνεται εκκίνηση του MATLAB® και πλοήγηση στη διεύθυνση (directory) όπου βρίσκονται τα αρχεία του Snor-Lux (Εικόνα Α.1).

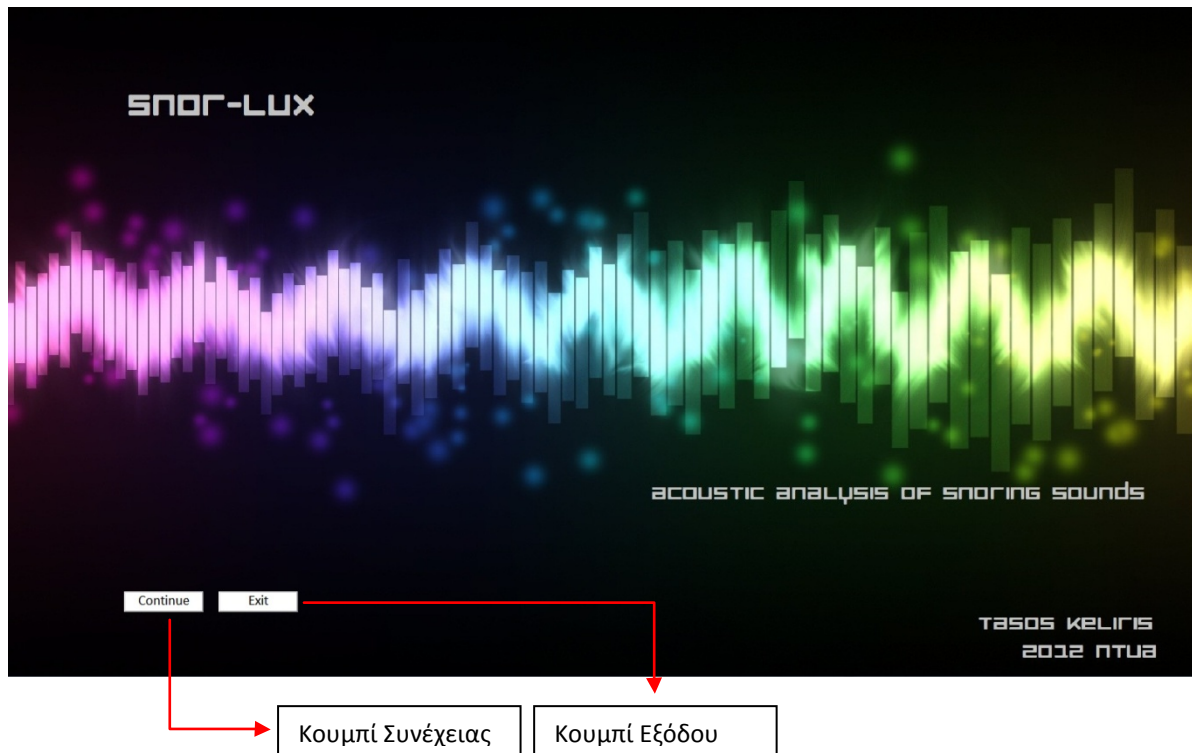


Εικόνα Α.1: Αρχική οθόνη περιβάλλοντος εργασίας στο MATLAB®

Ακολούθως η εκκίνηση του Snor-Lux γίνεται με πληκτρολόγηση στο command window, της εντολής `snor_lux`.

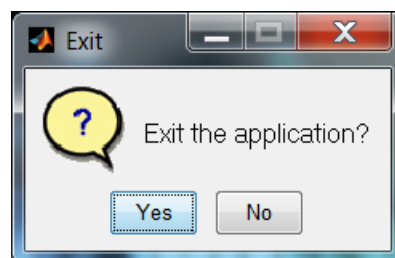
A.3: Επιφάνεια αρχικής οθόνης

Αφού εκκινήσει το Snor-Lux εμφανίζεται η αρχική επιφάνεια του πακέτου (Εικόνα A.2).



Εικόνα A.2: Αρχική επιφάνεια του Snor-Lux

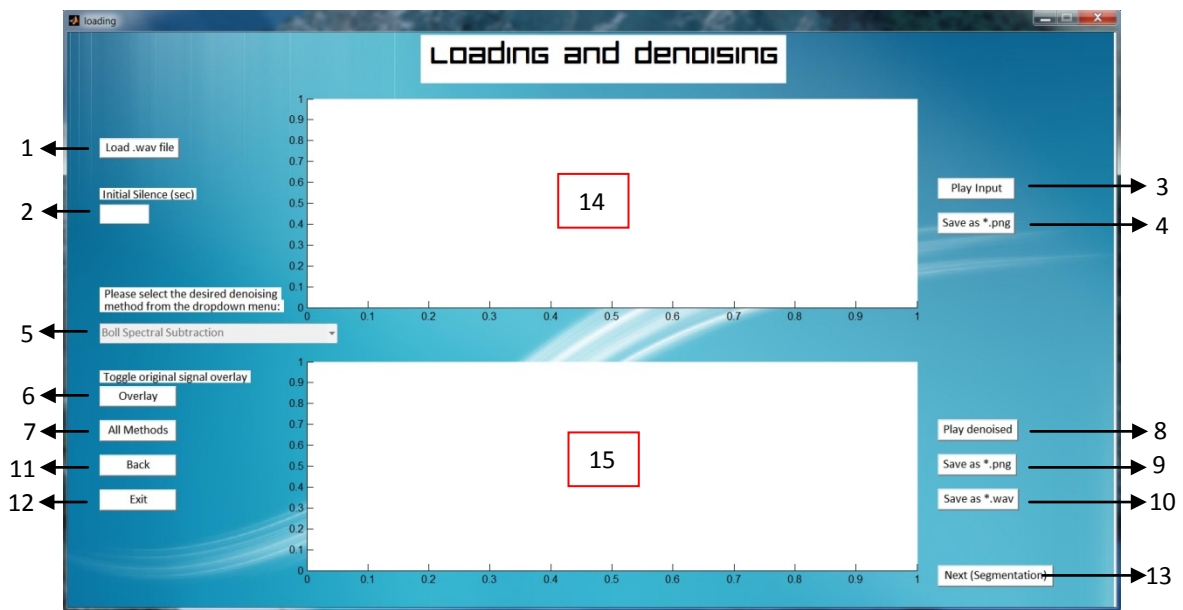
Το κουμπί συνέχειας οδηγεί στην επόμενη επιφάνεια του πακέτου, ενώ το κουμπί εξόδου μετά από μήνυμα επιβεβαίωσης, κλείνει το πρόγραμμα (Εικόνα A.3).



Εικόνα A.3: Μήνυμα επιβεβαίωσης εξόδου από το Snor-Lux

A.4: Επιφάνεια φόρτωσης αρχείου και αποθορυβοποίησης

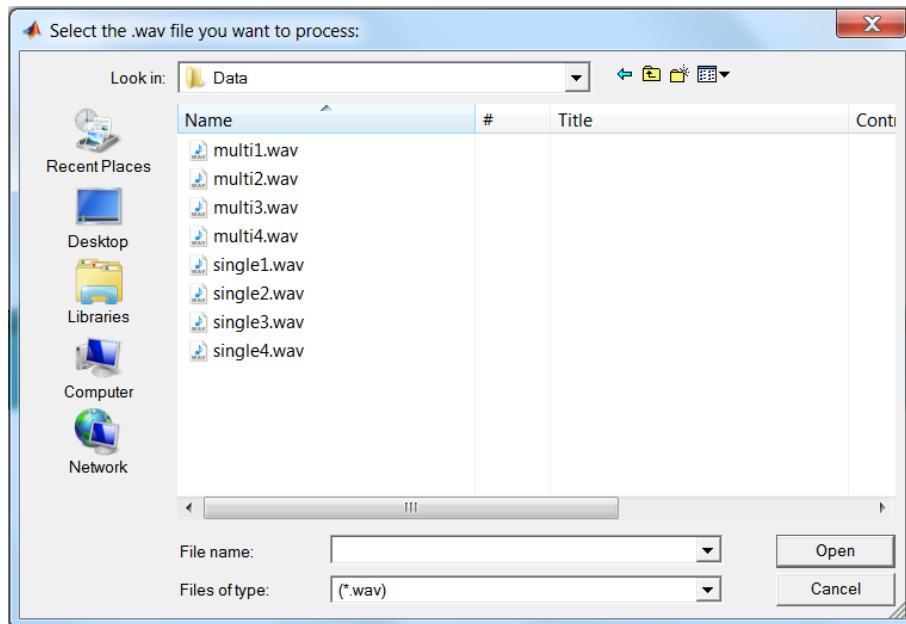
Στη δεύτερη επιφάνεια του λογισμικού γίνεται επιλογή του αρχείου προς επεξεργασία και επιλογή της επιθυμητής μεθόδου αποθορυβοποίησής του, από μια λίστα που περιλαμβάνει τις μεθόδους που αναπτύχθηκαν στο Κεφάλαιο 2. Απαραίτητη είσοδος για την λειτουργία των μεθόδων αποθορυβοποίησης είναι η εισαγωγή του χρόνου αρχικής ησυχίας (initial silence) του σήματος ώστε να γνωρίζει ο αλγόριθμος το προφίλ του θορύβου.



Εικόνα A.4: Επιφάνεια φόρτωσης αρχείου και αποθορυβοποίησης

Στην Εικόνα A.4 φαίνεται η δομή της επιφάνειας φόρτωσης και αποθορυβοποίησης αρχείου. Παρακάτω εξηγούνται τα στοιχεία που την αποτελούν (1 - 15) και η λειτουργία τους, ενώ στην Εικόνα A.7 φαίνεται συμπληρωμένη η επιφάνεια φόρτωσης και αποθορυβοποίησης.

1. Κουμπί φόρτωσης αρχείου: Με το πάτημα του κουμπιού αυτού εμφανίζεται ένα παράθυρο πλοήγησης (Εικόνα A.5), από το οποίο ο χρήστης μπορεί να επιλέξει το φάκελο που περιέχει τα αρχεία που θέλει να επεξεργαστεί και ακολούθως το συγκεκριμένο αρχείο προς επεξεργασία. Μόλις γίνει η επιλογή του αρχείου *.wav εμφανίζεται η γραφική παράσταση του σήματος στο χρόνο, στο χώρο 14.



Εικόνα Α.5: Μενού πλοήγησης για την επιλογή αρχείου

2. Πλαίσιο κειμένου για εισαγωγή χρόνου αρχικής ησυχίας: Στο πλαίσιο κειμένου (text box) γίνεται εισαγωγή του χρόνου αρχικής ησυχίας του σήματος, αφού προηγηθεί σύντομη οπτική επισκόπηση του χρονικού προφίλ του σήματος από το χρήστη. Το κείμενο πρέπει να αποτελείται από αριθμούς και εάν πρόκειται για δεκαδικό αριθμό ο διαχωρισμός του ακέραιου από το δεκαδικό μέρος να γίνεται με το σύμβολο "." (τελεία).

3. Κουμπί αναπαραγωγής αρχικού σήματος: Το κουμπί αυτό αναπαράγει ηχητικά το σήμα εισόδου, χωρίς καμία επεξεργασία.

4. Κουμπί αποθήκευσης αρχικού σήματος ως εικόνα: Το κουμπί αυτό αποθηκεύει το χρονικό αποτύπωμα του σήματος εισόδου σε φάκελο και με όνομα τα οποία επιλέγει ο χρήστης. Ο μορφότυπος αποθήκευσης της εικόνας είναι *.png.

5. Λίστα μεθόδων αποθρομβοποίησης: Η αναπτυσσόμενη λίστα αυτή χρησιμοποιείται για την επιλογή της μεθόδου αποθρομβοποίησης του σήματος. Οι επιλογές που προσφέρονται είναι:

- Boll Spectral Subtraction
- Berouti Spectral Subtraction
- Kamath Multiband Spectral Subtraction
- TSNR Wiener Filters

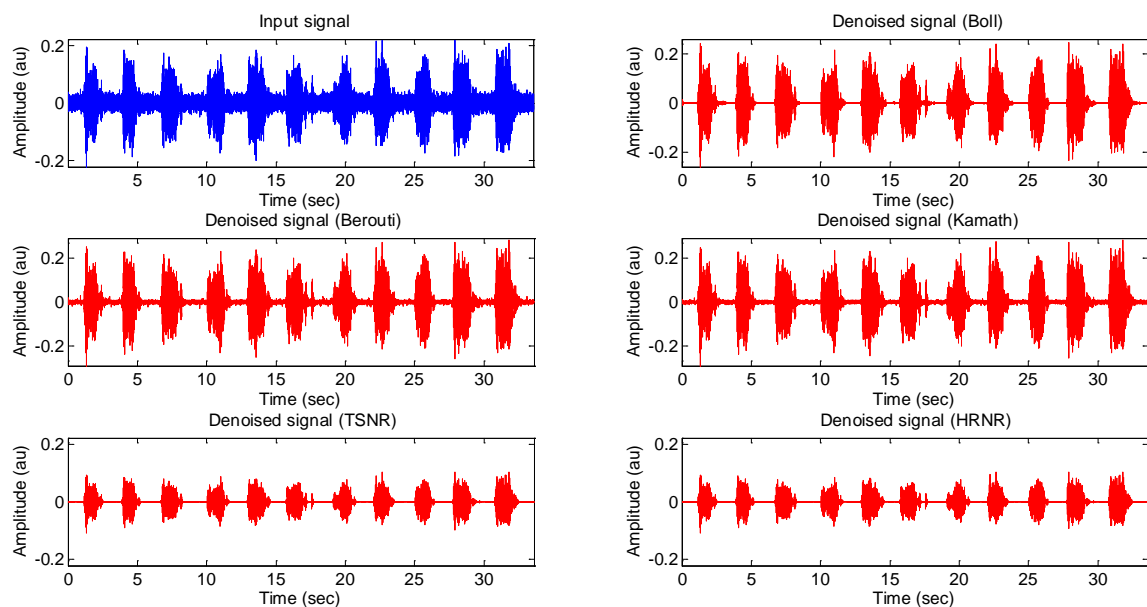
Παράρτημα Α

○ HRNR Wiener Filters

Μόλις γίνει επιλογή κάποιας μεθόδου από τη λίστα, εμφανίζεται η γραφική παράσταση του σήματος στο χρόνο, στο χώρο 15.

6. Κουμπί εναλλαγής υπέρθεσης: Το κουμπί αυτό εναλλάσσει τη ρύθμιση υπέρθεσης του αποθρομβοποιημένου σήματος στο αρχικό, στο χώρο 15.

7. Κουμπί εμφάνισης του αποθρομβοποιημένου σήματος με χρήση όλων των μεθόδων: Με το πάτημα του κουμπιού αυτού γίνεται υπολογισμός του αποθρομβοποιημένου σήματος και με τις πέντε μεθόδους και παρουσιάζονται τα αποτελέσματα σε νέο παράθυρο. Το αποτέλεσμα φαίνεται στην Εικόνα Α.6.



Εικόνα Α.6: Παράθυρο σημάτων που προκύπτει από το κουμπί "All Methods"

8. Κουμπί αναπαραγωγής αποθρομβοποιημένου σήματος: Το κουμπί αυτό αναπαράγει ηχητικά το αποθρομβοποιημένο σήμα, ανάλογα με τη μέθοδο που είναι επιλεγμένη στη λίστα μεθόδων αποθρομβοποίησης (5).

9. Κουμπί αποθήκευσης αποθρομβοποιημένου σήματος ως εικόνα: Το κουμπί αυτό αποθηκεύει το χρονικό αποτύπωμα του αποθρομβοποιημένου σήματος ανάλογα με τη μέθοδο που είναι επιλεγμένη στη λίστα μεθόδων αποθρομβοποίησης (5), σε φάκελο και με όνομα τα οποία επιλέγει ο χρήστης. Ο μορφότυπος αποθήκευσης της εικόνας είναι *.png.

10. Κουμπί αποθήκευσης αποθρουβοποιημένου σήματος ως ηχητικό σήμα: Το κουμπί αυτό αποθηκεύει το αποθρουβοποιημένο σήμα ως ήχο ανάλογα με τη μέθοδο που είναι επιλεγμένη στη λίστα μεθόδων αποθρουβοποίησης (5), σε φάκελο και με όνομα τα οποία επιλέγει ο χρήστης. Ο μορφότυπος αποθήκευσης της εικόνας είναι *.wav.

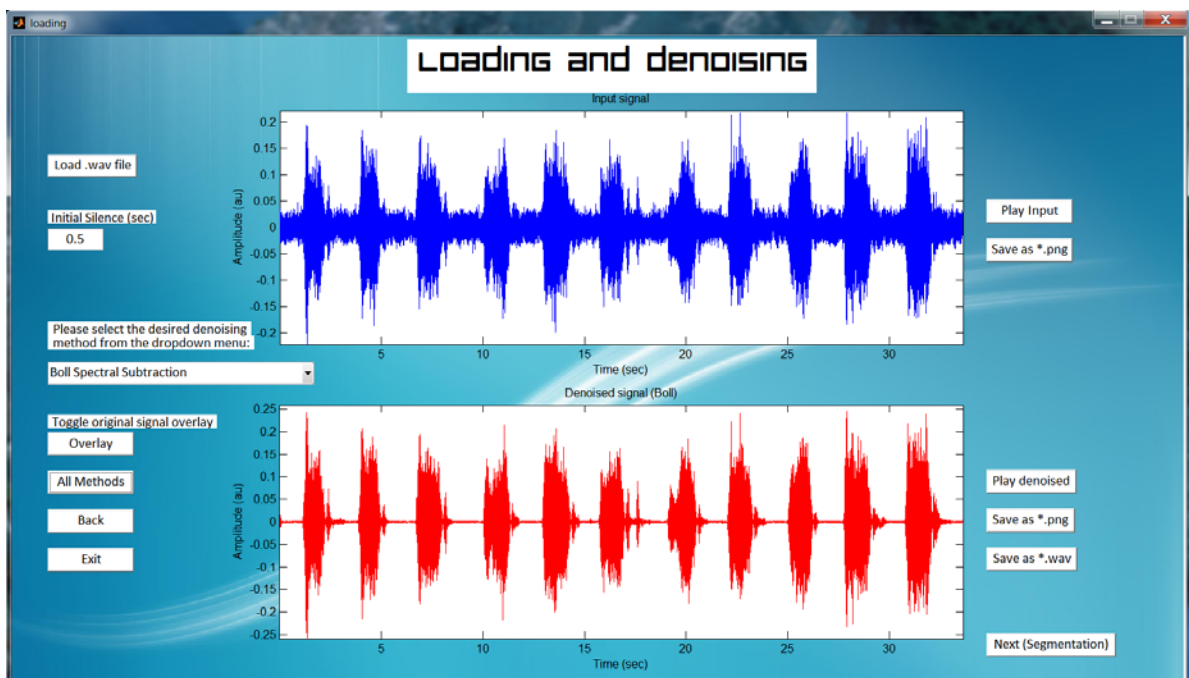
11. Κουμπί επιστροφής (Back): Επιστρέφει στην προηγούμενη επιφάνεια δηλαδή τη αρχική επιφάνεια του Snor-Lux.

12. Κουμπί εξόδου: Κλείνει το πρόγραμμα, μετά από μήνυμα επιβεβαίωσης (Εικόνα A.3).

13. Κουμπί επόμενης επιφάνειας: Προχωρά στην επόμενη επιφάνεια του προγράμματος, τη επιφάνεια κατάτμησης αφού γίνει αυτόματα αποθήκευση των δεδομένων στη μνήμη του υπολογιστή.

14. Χώρος γραφικών παραστάσεων σήματος εισόδου.

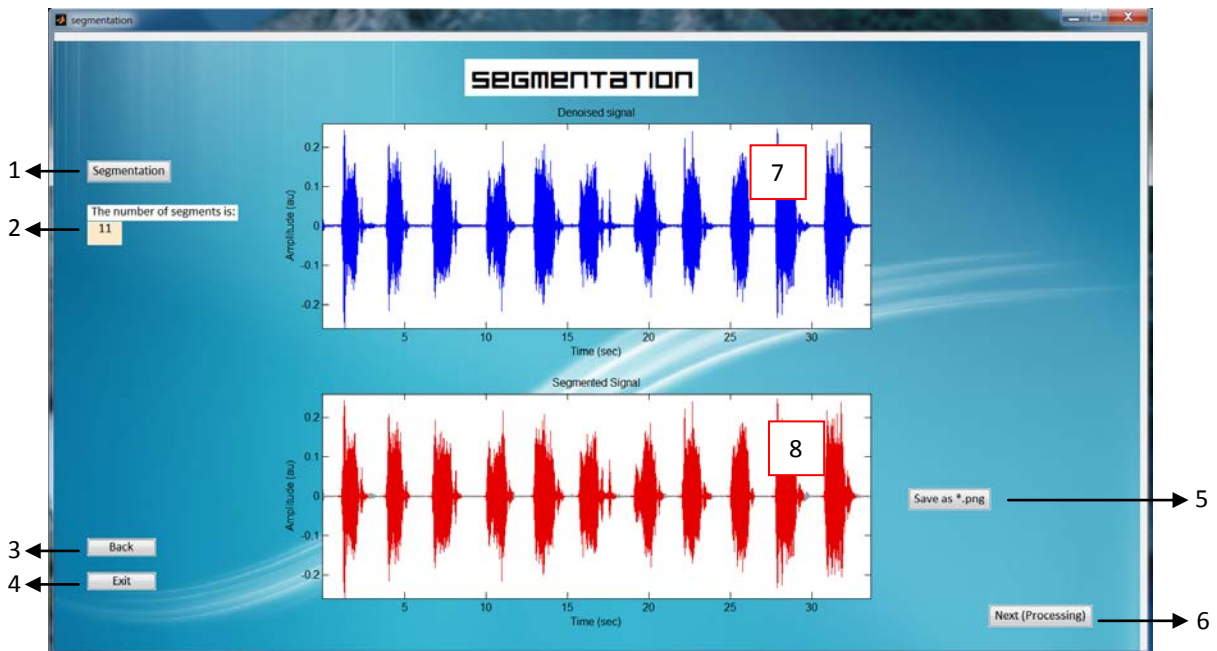
15. Χώρος γραφικών παραστάσεων αποθρουβοποιημένου σήματος.



Εικόνα A.7: Συμπληρωμένη επιφάνεια φόρτωσης και αποθρουβοποίησης αρχείου

A.5: Επιφάνεια κατάτμησης σήματος

Στη επιφάνεια κατάτμησης σήματος γίνεται ο χωρισμός του σήματος που επιλέχθηκε στην προηγούμενη επιφάνεια σε επιμέρους ροχαλητά ενώ φαίνεται επίσης και ο αριθμός των επιμέρους ροχαλητών.



Εικόνα Α.8: Επιφάνεια κατάτμησης σήματος

Στην Εικόνα Α.8 φαίνεται η δομή της επιφάνειας κατάτμησης αρχείου. Παρακάτω εξηγούνται τα στοιχεία που την αποτελούν (1 - 8) και η λειτουργία τους.

1. Κουμπί κατάτμησης (segmentation): Το κουμπί αυτό φορτώνει το αρχείο που επιλέχθηκε και αποθρομβοποιήθηκε στην προηγούμενη επιφάνεια και εκτελεί τον αλγόριθμο κατάτμησης σε επιμέρους ροχαλητά. Με το πάτημα του κουμπιού εμφανίζεται στο χώρο 7 το αρχικό αποθρομβοποιημένο σήμα και στο χώρο 8 η τεμαχισμένη έκδοση του. Τα επιμέρους ροχαλητά εμφανίζονται με κόκκινο χρώμα, ενώ οι ενδιάμεσες περιόδους "ησυχίας" με γκριζό χρώμα.

2. Πλαίσιο κειμένου: Στο πλαίσιο κειμένου εμφανίζεται, αφού γίνει η κατάτμηση του σήματος, ο συνολικός αριθμός των επιμέρους ροχαλητών που προκύπτουν.

3. Κουμπί επιστροφής (Back): Επιστρέφει στην προηγούμενη επιφάνεια δηλαδή τη επιφάνεια φόρτωσης και αποθρομβοποίησης του Snor-Lux.

4. Κουμπί εξόδου: Κλείνει το πρόγραμμα, μετά από μήνυμα επιβεβαίωσης (Εικόνα A.3).

5. Κουμπί αποθήκευσης αποθρορυβοποιημένου σήματος ως εικόνα: Το κουμπί αυτό αποθηκεύει τη γραφική παράσταση που εμφανίζεται στο χώρο 8, δηλαδή το κατατετμημένο σήμα με τα όρια του, σε φάκελο και με όνομα τα οποία επιλέγει ο χρήστης. Ο μορφότυπος αποθήκευσης της εικόνας είναι *.png.

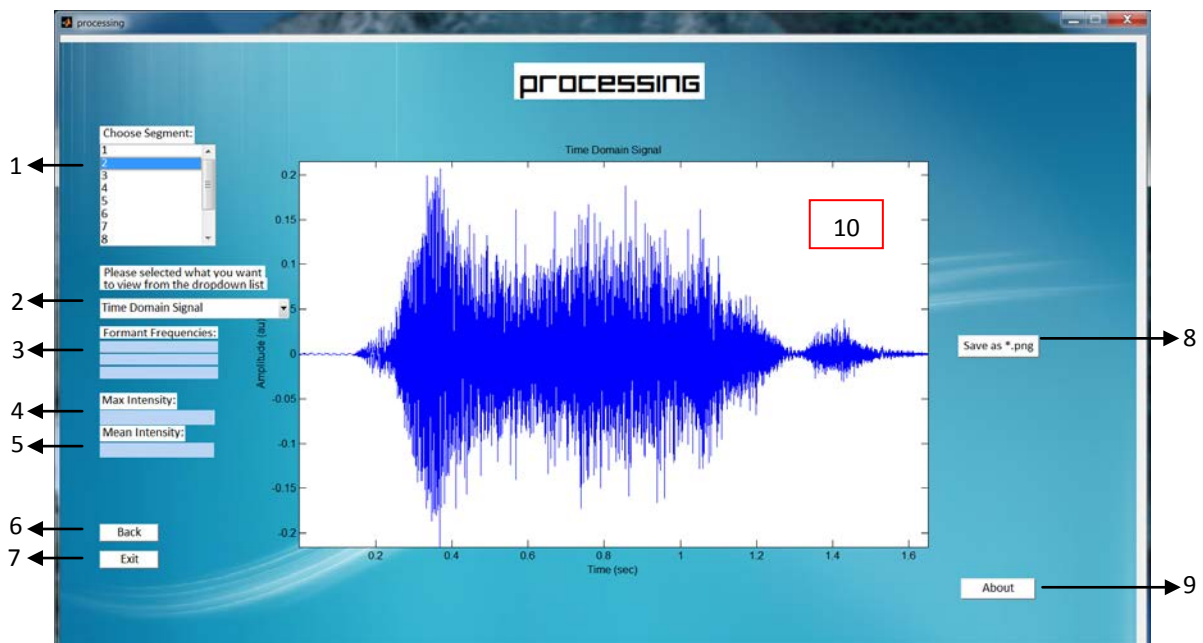
6. Κουμπί επόμενης επιφάνειας: Προχωρά στην επόμενη επιφάνεια του προγράμματος, τη επιφάνεια επεξεργασίας σήματος αφού γίνει αυτόματα αποθήκευση του κατατετμημένου σήματος και των ορίων του, στη μνήμη του υπολογιστή.

7. Χώρος γραφικών παραστάσεων αποθρορυβοποιημένου σήματος.

8. Χώρος γραφικών παραστάσεων κατατετμημένου σήματος.

A.6: Επιφάνεια επεξεργασίας σήματος

Στη επιφάνεια επεξεργασίας σήματος (Εικόνα A.9) γίνεται η κυρίως επεξεργασία των επιμέρους ροχαλητών. Αφού γίνει η επιλογή από το χρήστη του ροχαλητού για το οποίο απαιτείται εξαγωγή χαρακτηριστικών, επιλέγεται η ζητούμενη μέθοδος επεξεργασίας από μια αναπτυσσόμενη λίστα.



Εικόνα A.9: Επιφάνεια επεξεργασίας σήματος

Παράρτημα Α

Η λίστα περιλαμβάνει τις αναπαραστάσεις και μεγέθη που φαίνονται στον Πίνακα Α-1:

Πίνακας Α.1: Μέθοδοι επεξεργασίας που προσφέρει το Spor-Lux

A/A	Ονομασία	Λειτουργία
1	Time Domain Signal	Αναπαράσταση του σήματος στο πεδίο του χρόνου
2	Spectrum (dB)	Αναπαράσταση του λογαριθμικού πλάτους σήματος (σε dB) στο πεδίο της συχνότητας
3	Spectrum	Αναπαράσταση του σήματος στο πεδίο της συχνότητας
4	Spectrogram	Φασματογράφημα του σήματος
5	Cepstrum	Σάφμα του σήματος
6	Autocorrelation	Αυτοσυσχέτιση του σήματος
7	Zero-Crossings	Ρυθμός διελεύσεων από το μηδέν του σήματος
8	Short-time Energy	Ενέργεια βραχέως χρόνου του σήματος
9	Formant Extraction	Εξαγωγή των συχνοτήτων φωνοσυντονισμού του σήματος
10	Log (mel) - FilterBank Energies	Λογαριθμικές mel ενέργειες τραπεζών φίλτρων του σήματος
11	Mel Frequency Cepstrum	Σάφμα mel συχνοτήτων του σήματος
12	Bispectrum (Indirect Method)	Διφάσμα του σήματος υπολογισμένο με την έμμεση μέθοδο
13	Bispectrum (Direct Method)	Διφάσμα του σήματος υπολογισμένο με την άμεση μέθοδο
14	CWT	Συνεχής μετασχηματισμός κυματιδίων του σήματος
15	Intensity	Καμπύλη έντασης σήματος (dB)

Στην Εικόνα Α.9 φαίνεται η δομή της επιφάνειας επεξεργασίας σήματος. Παρακάτω εξηγούνται τα στοιχεία που την αποτελούν (1 - 10) και η λειτουργία τους.

1. Πλαίσιο λίστας: Στο πλαίσιο λίστας παραθέτονται αύξοντες αριθμοί που αντιπροσωπεύουν τα επιμέρους ροχαλητά, όπως χωρίστηκαν στη επιφάνεια κατάτμησης. Όταν γίνει επιλογή του σήματος προς επεξεργασία, εμφανίζεται στο χώρο 10 η χρονική αναπαράσταση του σήματος.

2. Αναπτυσσόμενη λίστα μεθόδων επεξεργασίας: Από την αναπτυσσόμενη λίστα ο χρήστης επιλέγει την επεξεργασία που επιθυμεί. Οι μέθοδοι επεξεργασίας παρουσιάζονται στον Πίνακα Α-1.

3. Πλαίσιο κειμένου συχνοτήτων φωνοσυντονισμού: Στα τρία πλαίσια κειμένου αυτά εμφανίζεται, αφού γίνει επιλογή της εξαγωγής των συγκεκριμένων συχνοτήτων (επιλογή 9 στην αναπτυσσόμενη λίστα - Formant Extraction), οι τρεις πρώτες συχνότητες φωνοσυντονισμού, F1, F2 και F3.

4. Πλαίσιο κειμένου μέγιστης έντασης: Στο πλαίσιο αυτό εμφανίζεται, αφού γίνει επιλογή της εμφάνισης της έντασης του σήματος (επιλογή 15 στην αναπτυσσόμενη λίστα - Intensity), η μέγιστη ένταση του σήματος σε dB.

5. Πλαίσιο κειμένου μέσης έντασης: Στο πλαίσιο αυτό εμφανίζεται, αφού γίνει επιλογή της εμφάνισης της έντασης του σήματος (επιλογή 15 στην αναπτυσσόμενη λίστα - Intensity), η μέση ένταση του σήματος σε dB.

6. Κουμπί επιστροφής (Back): Επιστρέφει στην προηγούμενη επιφάνεια δηλαδή τη επιφάνεια κατάτμησης σήματος του Snor-Lux.

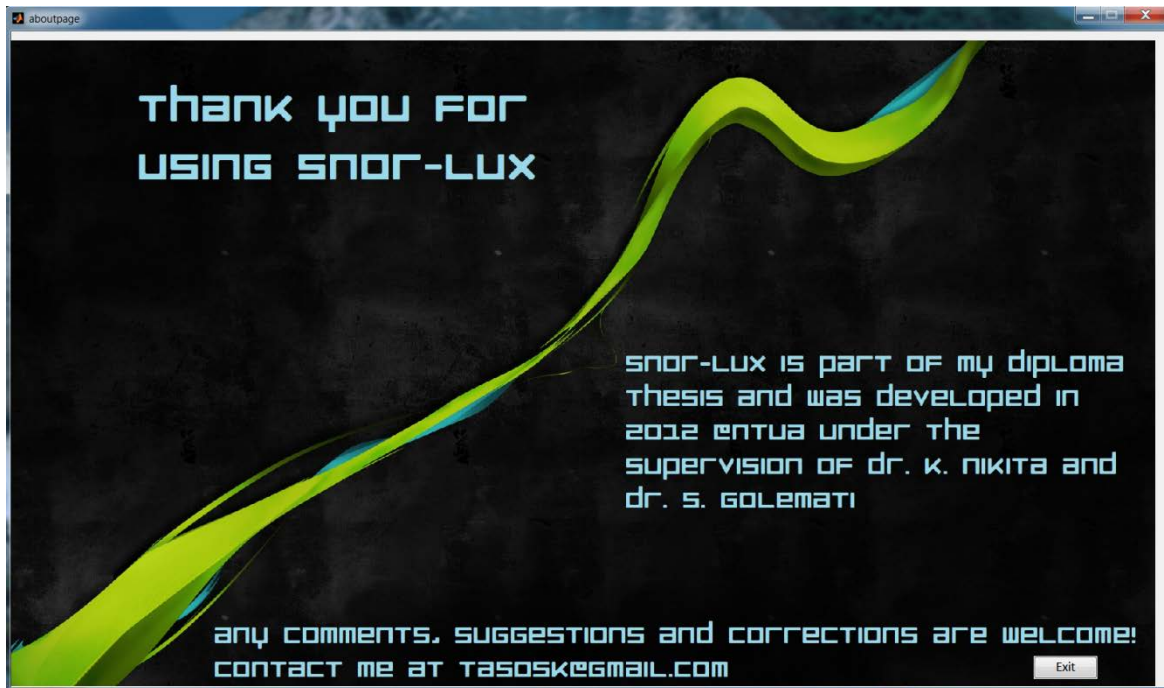
7. Κουμπί εξόδου: Κλείνει το πρόγραμμα, μετά από μήνυμα επιβεβαίωσης (Εικόνα A.3).

8. Κουμπί αποθήκευσης επεξεργασμένου σήματος ως εικόνα: Το κουμπί αυτό αποθηκεύει τη γραφική παράσταση που εμφανίζεται στο χώρο 10, δηλαδή το σήμα αφού έχει τύχει επεξεργασίας, σε φάκελο και με όνομα τα οποία επιλέγει ο χρήστης. Ο μορφότυπος αποθήκευσης της εικόνας είναι *.png.

9. Κουμπί επόμενης επιφάνειας: Προχωρά στην επόμενη επιφάνεια του προγράμματος, τη επιφάνεια ευχαριστιών και επικοινωνίας.

A.7: Επιφάνεια επικοινωνίας

Στη επιφάνεια επικοινωνίας (Εικόνα Α.10) υπάρχουν στοιχεία επικοινωνίας για εισηγήσεις, σχόλια και βελτιώσεις από τους χρήστες του προγράμματος. Υπάρχει επίσης κουμπί εξόδου για επιστροφή στο περιβάλλον εργασίας του προγράμματος MATLAB®.



Εικόνα Α.10: Επιφάνεια επικοινωνίας