ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
& ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΜΕΘΟΔΟΙ ΑΥΤΟΜΑΤΗΣ ΤΑΞΙΝΟΜΗΣΗΣ ΦΩΝΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ

ΣΠΗΛΙΩΤΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Εξεταστική Επιτροπή:
Διγαλάκης Βασίλειος (επιβλέπων)
Ποταμιάνος Αλέξανδρος
Λαγουδάκης Μιχαήλ

Χανιά, Νοέμβριος 2011

στους γονείς μου
και στον αδερφό μου

στη μνήμη του Ιατρού Stefan Schalber
(14.03.1984 – 22.02.2011)

στα παιδιά

Ειδικές Ευχαριστίες

Ευχαριστώ πολύ *τους γονείς μου* και πάλι για την αμέριστη στήριξή τους καθ'όλη την διάρκεια των σπουδών μου.

Ευχαριστώ πολύ τον *Καθηγητή κύριο Βασίλη Διγαλάκη* για αυτό το αξέχαστο ταξίδι στην επιστήμη της Στατιστικής Μοντελοποίησης και Αναγνώρισης Προτύπων και φυσικά για την στήριξη του με κάθε τρόπο από την πρώτη μέρα ανάληψης της Διπλωματικής μου Εργασίας.

Ευχαριστώ πολύ τον *Δρ. Βασίλη Διακολουκά* για τη αδιάκοπη βοήθειά του και τον συνεχή υποστηρικτικό και συμβουλευτικό του ρόλο όλους αυτούς τους μήνες.

Ευχαριστώ πολύ τον *Επίκουρο Καθηγητή Ιατρό Ιωάννη Γερμανάκη* για την συμμετοχή μου στο Erasmus Pedocardiology Intensive Program, τον Οκτώβριο του 2010 στο Ηράκλειο και για όλες τις πολύτιμες γνώσεις και συμβουλές μου μεταλαμπάδευσε σε αυτή την περίοδο.

Ευχαριστώ θέρμα *όλους που με τον δικό τους τρόπο στάθηκαν δίπλα μου* όλο αυτό το χρονικό διάστημα της εκπόνησης της Διπλωματικής μου Εργασίας.

Περιεχόμενα

1 ΕΙΣΑΓΩΓΗ 6

1.1 ΓΕΝΙΚΑ 6

1.2 ΓΕΝΙΚΑ ΠΕΡΙ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΤΥΠΩΝ ΚΑΙ ΑΝΑΠΤΥΞΗΣ ΣΥΣΤΗΜΑΤΩΝ 7

1.3 ΤΑ ΦΥΣΗΜΑΤΑ ΣΤΑ ΠΑΙΔΙΑ ΚΑΙ Η ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΤΟΥΣ 10

1.4 ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ ΚΑΙ ΕΡΕΥΝΕΣ 11

1.5 ΣΥΝΕΙΣΦΟΡΑ ΚΑΙ ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ 13

1.6 ΟΡΓΑΝΩΣΗ ΤΗΣ ΑΝΑΦΟΡΑΣ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ 14

2 ΑΝΙΧΝΕΥΣΗ ΚΑΡΔΙΟΠΑΘΕΙΩΝ ΜΕΣΩ ΗΧΗΤΙΚΩΝ ΣΗΜΑΤΩΝ ΚΑΡΔΙΑΚΗΣ ΛΕΙΤΟΥΡΓΙΑΣ 15

2.1 ΤΟ ΚΑΡΔΙΑΚΟ ΗΧΗΤΙΚΟ ΣΗΜΑ ΚΑΙ ΤΑ ΦΥΣΙΟΛΟΓΙΚΑ ΤΟΥ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ 17

2.2 ΜΗ ΦΥΣΙΟΛΟΓΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ ΚΑΡΔΙΑΚΟΥ ΣΗΜΑΤΟΣ 19

2.3 Η ΚΑΡΔΙΑΚΗ ΑΚΡΟΑΣΗ 24

2.4 ΠΡΟΒΛΗΜΑΤΑ ΚΑΙ ΔΥΣΚΟΛΙΕΣ ΤΗΣ ΚΑΡΔΙΑΚΗΣ ΑΚΡΟΑΣΗΣ 29

3 ΣΥΛΛΟΓΗ ΚΑΙ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ 30

3.1 ΓΕΝΙΚΑ 32

3.2 ΣΥΛΛΟΓΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ 33

3.3 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΚΑΤΑΤΜΙΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ 34

3.3.1 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ 34

3.3.2 ΚΑΤΑΤΜΙΣΗ ΣΕ ΚΥΚΛΟΥΣ ΚΑΙ ΤΜΗΜΑΤΑ 36

3.4 ΕΞΑΓΩΓΗ ΤΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΏΝ 42

4 ΑΝΑΠΤΥΞΗ ΤΑΞΙΝΟΜΗΤΩΝ 49

4.1 ΑΝΑΠΤΥΞΗ ΤΟΥ GMM ΤΑΞΙΝΟΜΗΤΗ 49

4.2 ΠΑΡΟΥΣΙΑΣΗ ΚΑΙ ΧΡΗΣΗ ΤΩΝ ΕΡΓΑΛΕΙΩΝ WEKA 53

5 ΠΕΙΡΑΜΑΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΣΥΣΤΗΜΑΤΩΝ 57

5.1 ΔΙΑΔΙΚΑΣΙΕΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ 1ΟΥ ΠΕΙΡΑΜΑΤΙΚΟΥ ΜΕΡΟΥΣ 57

5.2 ΔΙΑΔΙΚΑΣΙΕΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ 1ΟΥ ΠΕΙΡΑΜΑΤΙΚΟΥ ΜΕΡΟΥΣ 67

5.3 ΣΥΓΚΡΙΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ 70

6 ΣΥΜΠΕΡΑΣΜΑΤΑ 72

7 ΒΙΒΛΙΟΓΡΑΦΙΑ 75

8 ΣΥΧΝΕΣ ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ 78

1 Εισαγωγή

1.1 Γενικά

Τα τελευταία χρόνια η παρουσία της σύγχρονης τεχνολογίας στην ιατρική αποτελεί ένα γεγονός. Ολοένα και περισσότερα συστήματα σχεδιάζονται, με σκοπό να συνεπικουρήσουν στο δύσκολο έργο του Ιατρού. Σε πολλές και δύσκολες περιπτώσεις σε διάφορες παθήσεις, αξιόπιστα συστήματα υποστηρίζουν με τις ενδείξεις τους τον Ιατρό κατά την διάρκεια της διάγνωσης, του εντοπισμού του ευρήματος και της αξιολόγησης του. Πρόσφατα μάλιστα, ξεκίνησαν να λειτουργούν πιλοτικά και στην χώρα μας συστήματα ρομποτικών χειρουργών. Η ακρίβεια και η σταθερότητα τέτοιων συστημάτων, ωθούν ολοένα και περισσότερους ερευνητές στο σχεδιασμό και γενικότερα στη μελέτη τους. Φυσικά δεν τίθενται θέματα ηθικής ή δεοντολογίας, αφού τέτοιες προσπάθειες σκοπό έχουν να διαδραματίσουν υποστηρικτικό και εκπαιδευτικό ρόλο κι όχι να αντικαταστήσουν τον ίδιο τον Ιατρό. Διάφορες τέτοιες τεχνο-ιατρικές μελέτες και έρευνες, έχουν αναπτυχθεί σε αντικείμενα όπως η αποκωδικοποίηση ακολουθιών του DNA, η σύνθεση των πρωτεϊνών, η Καρδιολογία, η μοριακή Βιολογία, η καταπολέμηση μορφών καρκίνου, οι μέθοδοι μή επεμβατικής διάγνωσης και πολλά άλλα.

Ένας από τους ιατρικούς κλάδους που κεντρίζει το ενδιαφέρον πολλών ερευνητών μηχανικών, είναι αυτός της Καρδιολογίας. Καρδιολογία είναι ο κλάδος της Ιατρικής επιστήμης, που ασχολείται με τις παθήσεις της καρδιάς, δηλαδή με τις διάφορες συγγενείς καρδιακές παθήσεις, τις αρτηριακές παθήσεις, την καρδιακή ανεπάρκεια, τις βαλβιδικές παθήσεις καθώς επίσης και την Ηλεκτροφυσιολογία. Όπως σε όλους τους κλάδους της Ιατρικής επιστήμης, έτσι και στην Καρδιολογία η διάγνωση των διαφόρων παθήσων, γίνεται μέσα από την παρατήρηση και σχετική αξιολόγηση των ευρημάτων – συμπτωμάτων. Πολλά συμπτώματα σχετίζονται με καρδιακές βλάβες, όπως πόνος στο στήθος, πυρετός και άλλα καθώς και πολλά ευρήματα όπως το φύσημα, οι κλαγγές, οι επιπρόσθετοι ήχοι και άλλα που παρατηρούνται κατά τη διάρκεια σχετικών εξετάσεων.

Πολλά παιδιά παρουσιάζουν φύσημα κατά την καρδιολογική τους εξέταση. Ωστόσο το μεγαλύτερο ποσοστό αυτών των παιδιών έχει όπως λέγεται φυσιολογικό ή λειτουργικό φυσήμα – πολλοί Ιατροί το αποκαλούν «αθώο», που πρόκειται στο μέλλον με την ανάπτυξή τους να εξαλειφθεί. Όμως υπάρχει ένα ποσοστό παιδιατρικών ασθενών, που το φύσημά τους είναι παθολογικό ή οργανικό όπως καλείται και οφείλεται σε κάποιο υπαρκτό πρόβλημα, όπως για παράδειγμα στη στένωση αρτηρίας, στην ενδοκαρδιακή επικοινωνία και άλλα. Σε περίπτωση που ο παιδίατρος διαγνώσει την ύπαρξη κάποιου τέτοιου παθολογικού φυσήματος, θα πρέπει να παραπέμψει τον παιδιατρικό ασθενή σε Ειδικό Καρδιολόγο για

περαιτέρω εξετάσεις, αξιολόγηση και διάγνωση. Φυσικά, δεν μπορούν όλοι οι ασθενείς για πρακτικούς λόγους να παραπεμπφθούν σε Καρδιολόγους κι έτσι μια αρκετά δύσκολη αποστολή επωμίζεται ο Παιδίατρος. Η τεχνοεπιστημονική κοινότητα, αναγνωρίζοντας το δύσκολο αυτό πρόβλημα διάγνωσης και τις επικρατούσες συνθήκες έχει καταπιαστεί με την μελέτη των καρδιακών σημάτων, την επεξεργασία τους, ακόμη και την προσπάθεια για την αξιολόγησή τους μέσα από πειραματικά συστήματα. Οι τελευταίες μελέτες δείχνουν ότι «καθαρά τεχνολογικοί» τομείς όπως αυτός της Επεξεργασίας Σήματος, της Επεξεργασίας και Αναγνώρισης Φωνής, της Στατιστικής Μοντελοποίησης και άλλοι, μπορούν να επιστρατευτούν έτσι ώστε να παρέχουν έναν ακριβή και πολυσχιδή υποστηρικτικό ρόλο στο δύσκολο Ιατρικό έργο.

1.2 Γενικά περί Αναγνώρισης Προτύπων και Ανάπτυξης Συστημάτων

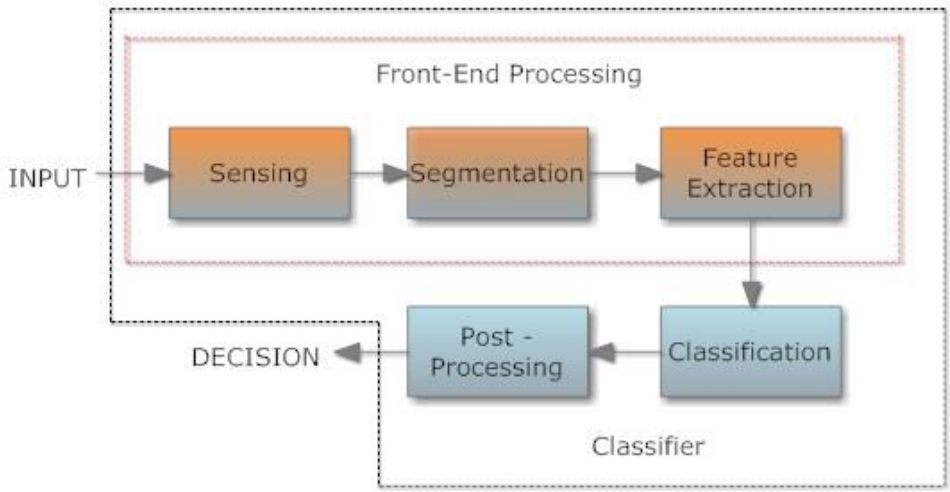
Αναγνώριση Προτύπων καλείται ο κλάδος της επιστήμης, που καταπιάνεται με την ταξινόμηση αντικειμένων σε ένα αριθμό κατηγοριών, σύμφωνα με τον ορισμό του Θεοδωρίδη. Με μια μικρή τροποποίηση, μπορούμε επίσης να πούμε πως πρόκειται για την διαδικασία η οποία λαμβάνει ως εισροή δεδομένα και αποδίδει ως εκροή συγκεκριμμένες ενέργειες βάσει της κατηγορίας του προτύπου, σύμφωνα με τον ορισμό του Duda. Η επιστήμη της Αναγνώρισης Προτύπων, εισάγει την έννοια της μηχανικής αντίληψης (machine perception) καθώς στα πλαίσια της γίνεται ανάπτυξη μηχανών – συστημάτων που αναγνωρίζουν πρότυπα. Κάποιες από τις πιο συνήθεις εφαρμογές τέτοιων συστημάτων χρησιμοποιούν:

- φωνητικά, ηχητικά και μουσικά δεδομένα για εφαρμογές αναγνώρισης φωνής, αναγνώρισης και ταυτοποίησης ομιλητή κ.ά.
- εικόνες και βίντεο για εφαρμογές OCR (Optical Character Recognition), AVASR (Audio – Visual Automatic Speech Recognition), αναγνώρισης προσώπου, αναγνώρισης δακτυλικού αποτυπώματος και κατάτμησης βίντεο κ.ά.
- κείμενα και γλωσσικά αρχεία για εφαρμογές μηχανικής διερμηνείας, κατηγοριοποίησης εγγράφων, κατανόησης κειμένων κ.ά.
- ιατρικά και βιολογικά δεδομένα για εφαρμογές διάγνωσης ασθενειών, αποκωδικοποίησης ακολουθιών DNA, μοντέλων γενετικών ασθενειών κ.ά.
- κάθε φύσης δεδομένα για εφαρμογές όπως μοντελοποίηση χρηστών σε εμπορικές εφαρμογές, ανάλυση ιστοτόπων διαδυκτίου, παιχνίδια κ.ά.

Ένα σύστημα Αναγνώρισης Προτύπων, αποτελείται από τα μέρη της Αίσθησης (Sensing), της Κατάτμησης και Ομαδοποίησης (Segmentation and Grouping), της Εξαγωγής Χαρακτηριστικών (Feature Extraction), της Κατηγοριοποίησης ή Ταξινόμησης (Classification) και της Ύστερης Επεξεργασίας (Post Processing), όπως φαίνεται και στο σχήμα 1.1. Επιγραμματικά:

- *Αίσθηση (Sensing)*: επιτυγχάνεται με χρήση κάποιου μέσου καταγραφής, όπως ένα μικρόφωνο, μια κάμερα κτλ.. Στη σχεδίαση λαμβάνονται υπ' όψη οι παράμετροι του μέσου καταγραφής όπως το εύρος ζώνης, η ακρίβεια και η ευαισθησία καθώς και το σφάλμα και παραμόρφωση που το όργανο εισάγει.
- *Κατάτμηση και Ομαδοποίηση (Segmentation and Grouping)*: τα πρότυπα χωρίζονται σωστά και αποφεύγονται τυχόν αλληλεπικαλύψεις (overlapping).
- *Εξαγωγή Χαρακτηριστικών (Feature Extraction):* εξάγονται τα χαρακτηριστικά που επιδεικνύουν μεγάλη διακριτική ικανότητα (discriminative features)
- *Κατηγοριοποίηση ή Ταξινόμηση (Classification):* είναι η διαδικασία κατά την οποία χρησιμοποιώντας ένα διάνυσμα χαρακτηριστικών ταξινομείται το αντικειμένο σε μια κατηγορία.
- *Ύστερης Επεξεργασίας (Post Processing):* εκμεταλλευόμαστε πληροφορία (ανεξάρτητη της εισόδου) εκτός του προτύπου στόχου για την περεταίρω βελτίωση της απόδοσης.

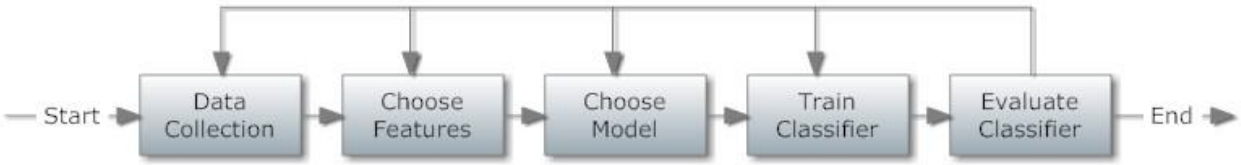
Παρακάτω φαίνεται σχηματικά η γενική δομή ενός συστήματος Αναγνώρισης Προτύπων.



Σχήμα 1.1 – Η δομή ενός συστήματος Αναγνώρισης Προτύπων

Τέλος, σε αυτή την ενότητα, χρήσιμο θα ήταν να αναφερθεί η διαδικασία σχεδιασμού ενός τέτοιου συστήματος σε λίγα απλά βήματα. Τα βήματα του σχεδιασμού, θα μπορούσαν να συνοψιστούν στην Συλλογή των Δεδομένων (Data Collection), την Επιλογή των Χαρακτηριστικών (Feature Selection), την Επιλογή του Μοντέλου (Model Choice), την Εκπαίδευση (Training), την Αξιολόγηση (Evaluation) και την Υπολογιστική Πολυπλοκότητα (Computational Complexity). Εποπτικά παρουσιάζονται και στο σχήμα 1.2. Συνοπτικά:

- *Συλλογή των Δεδομένων (Data Collection):* συγκέντρωση σχετικά μεγάλου και αντιπροσωπευτικού πλήθους δεδομένων, ικανών για την εκπαίδευση και αξιολόγηση του συστήματος.
- *Επιλογή των Χαρακτηριστικών (Feature Selection):* επιλέγονται τα χαρακτηριστικά εκείνα που είναι εύκολο να εξαχθούν, αμετάβλητα σε μετασχηματισμούς και ανθεκτικά στο θόρυβο. Μια καλή επιλογή χαρακτηριστικών επιβεβαιώνει τον κανόνα: «Χαρακτηριστικά των ίδιων κλάσεων έχουν παρόμοιες τιμές και διαφορετικές από αυτά των άλλων κλάσεων»
- *Επιλογή του Μοντέλου (Model Choice):* κατάλληλη επιλογή του μοντέλου σύμφωνα πάντα με τη μορφή των ορίων απόφασης (decision boundaries).
- *Εκπαίδευση (Training):* χρησιμοποιούνται κάποια από τα δεδομένα για τον ορισμό του ταξινομητή (εκπαίδευση παραμέτρων μοντέλου).
- *Αξιολόγηση (Evaluation):* μέτρηση του ποσοστού επιτυχίας (ή αποτυχίας)
- *Υπολογιστική Πολυπλοκότητα (Computational Complexity):* η αποσαφήνιση ενός αλγορίθμου ταξινόμησης σαν συνάρτηση του αριθμού των χαρακτηριστικών, των προτύπων και των κατηγοριών.



Σχήμα 1.2 – Τα στάδια σχεδιασμού ενός συστήματος Αναγνώρισης Προτύπων

1.3 Τα Φυσήματα στα Παιδιά και η Κατηγοριοποίηση τους

Το καρδιακό φύσημα (heart murmur), απότελεi ένα από τα ευρήματα κατά την διάρκεια μιας εξέτασης καρδιακής ακρόασης. Σημειώνεται εδώ, πως ως καρδιακή ακρόαση είναι η διαδικασία κατά την οποία ο Ιατρός εξετάζει τον ασθενή, χρησιμοποιώντας το στηθοσκόπιο. Το φύσημα είναι ένας επιπρόσθετος και «άχρηστος» ήχος που παρουσιάζεται κατά την διάρκεια του καρδιακού κύκλου και δεν αποτελεί έναν από τους προσδοκώμενους ήχους. Περισσότερες λεπτομέρειες σχετικά με τα χαρακτηριστικά του καρδιακού κύκλου, θα δίνονται στο επόμενο κεφάλαιο.

Το φύσημα παράγεται από την ροή του αίματος στην καρδιά και στο κυκλοφορικό σύστημα. Ουσιαστικά, το φύσημα είναι το ηχητικό αποτέλεσμα που παράγεται από την τριβή ενός ρευστού, του αίματος, που ρέει σε ένα αγγείο, διαμέσου της καρδιάς ή μιας βαλβίδας. Η τριβή αυτή δημιουργεί δονήσεις οι οποίες μεταφράζονται σε ήχους σε διάφορες συχνότητες. Έτσι δημιουργούνται τα λεγόμενα φυσιολογικά φυσήματα ή αθώα ή λειτουργικά (normal or innocent or functional murmurs), από την φυσιολογική λειτουργία της καρδιάς.

Αντίθετα, τα λεγόμενα παθολογικά φυσήματα (abnormal murmurs), συνδέονται με σημαντικές επιταχύνσεις της ροής του αίματος. Αίτια όπως στενώσεις, ελλείμματα, ενδοκαρδιακή επικοινωνία, βαλβιδικές παθήσεις και άλλα προκαλούν σημαντική επιτάχυνση της ροής του αίματος, η οποία είναι τυρβώδης. Η τύρβη που εμφανίζεται στη ροή του αίματος, παράγει δονήσεις σε καρδιαγγειακές δομές ικανής έντασης, ώστε να μεταδίδονται διαμέσου του θωρακικού τοιχώματος και άρα να γίνονται αντιληπτά κατά τη διάρκεια της καρδιακής ακρόασης. Υπάρχουν πολλοί παράγοντες που χαρακτηρίζουν τέτοια φυσήματα όπως η ένταση, ο χρονισμός, η συχνότητα, ο περιστασιακός εντοπισμός και άλλοι.

Η ακρόαση ενός φυσήματος στο νεογνό, στο παιδί ή στον έφηβο ευαισθητοποιεί τον Παιδίατρο ή τον γενικό Ιατρό για την πιθανότητα ύπαρξης κάποιου οργανικού προβλήματος, στο καρδιαγγειακό σύστημα. Σε πολλές περιπτώσεις, ο εξετάζων Ιατρός αποφασίζει την παραπομπή του ασθενή σε ειδικούς Ιατρούς – συνήθως Παιδοκαρδιολόγους ή Καρδιολόγους – για περαιτέρω εξετάσεις, με ηλεκτροκαρδιογράφημα και υπερηχοκαρδιογράφημα. Απαιτείται μια προσεκτική ακρόαση παράλληλα με μια λεπτομερή εξέταση του ιστορικού του ασθενή, προκειμένου να διαπιστωθεί η ύπαρξη ή μη κάποιας παθολογίας.

Μια τέτοια ακρόαση είναι πολύ συχνή στον παιδιατρικό κυρίως πληθυσμό, δηλαδή σε παιδιά μέχρι και 15 ετών. Υπολογίζεται ότι μέχρι και 72% του σχολικού πληθυσμού, παρουσιάζει κάποιας μορφής φυσιολογικό φύσημα, το οποίο συνήθως εξαλείφεται κατά την ανάπτυξη του. Η ελαστικότητα του θωρακικού τοιχώματος, η απουσία αλλοιώσεων του αέρα στους πνεύμονες, η λεπτότητα των στοιχείων μεταξύ της καρδιάς του παιδιού και του στηθοσκοπίου, η μειωμένη πηκτικότητα του αίματος λόγω της φυσιολογικης αναιμίας, είναι

μόνο μερικοί από τους παράγοντες, που καθιστούν την εμφάνιση λεπτών ήχων και φυσιολογικών φυσημάτων στα παιδιά αρκετά συχνά μέσα από τους ήπιους στροβιλισμούς της ροής του αίματος που δημιουργούν.

Έτσι λοιπόν τίθεται το ερώτημα της κατηγοριοποίησης, του χαρακτηρισμού με άλλα λόγια των φυσημάτων που παρουσίαζονται. Την αποστολή αυτή αναλαμβάνουν οι Παιδίατροι και οι γενικοί Ιατροί που πολλές φορές μάλιστα δεν είναι εξειδικευμένοι στην καρδιακή ακρόαση και σύμφωνα με αρκετές μελέτες μόνον το 50% τέτοιων Ιατρών είναι ικανό να διακρίνει ένα φυσιολογικό φύσημα από ένα παθολογικό, ενώ σύμφωνα με άλλες μελέτες ο μέσος Παιδίατρος ή γενικός Ιατρός, αποτυγχάνει να εντοπίσει 2 στα 10 παιδιά με καρδιαγγειακή πάθηση, ενώ παραπέμπει εσφαλμένα για περαιτέρω εξέτασεις 3 με 4 στα 10 παιδιά με φυσιολογικό φύσημα. Ως εκ τούτου, είναι ξεκάθαρο γιατί η κατηγοριοποίηση των φυσημάτων είναι πάρα πολύ σημαντική. Τα παιδιά στα οποία δεν αναγνωρίζεται ο κίνδυνος μέσα από την καρδιακή ακρόαση, θα αντιμετωπίσουν επιδεινώσεις και σε πολλές περιπτώσεις και κατάληξη από την καρδιακή νόσο που πάσχουν. Στην αντίθετη περίπτωση, η εσφαλμένη παραπομπή παιδιών με αθώα φυσήματα για καρδιολογικές εξετάσεις, δημιουργεί άγχος στην οικογένεια αλλά και στο ίδιο το παιδί ενώ επιβαρύνει οικονομικά το σύστημα υγείας.

Από ιατρικής πλευράς για την αντιμετώπιση αυτού του προβλήματος, της ορθής διάκρισης μεταξύ των φυσιολογικών και των παθολογικών φυσημάτων, ενταντικοποιείται η εκπαίδευση των γενικών Ιατρών γύρω από θέματα καρδιακής ακρόασης και φυσικά των υπό ειδίκευση Παιδιάτρων, ενώ μάλιστα τα τελευταία χρόνια, δίνεται η ευκαιρία σε Ιατρούς να εξειδικευτούν ως Παιδοκαρδιολόγοι. Ωστόσο και ο κλάδος της Μηχανικής και συγκεκριμμένα της Στατιστικής Μοντελοποίησης και Αναγνώρισης Προτύπων, της Μηχανικής Μάθησης, των Τηλεπικοινωνιών και άλλων έχουν καταπιαστεί με αυτό το πρόβλημα, πράγμα που καταμαρτηρούν οι πολυάριθμες εργασίες γύρω από αυτό. Στις διάφορες έρευνες έχουν χρησιμοποιηθεί μέθοδοι όπως στατικοί ταξινομητές, νευρωνικά δίκτυα, ανάλυση σπεκτρογραφημάτων και πολλές άλλες.

1.4 Σχετικές Εργασίες και Έρευνες

Όπως προαναφέρθηκε προηγουμένως, πλήθος ερευνών και εργασιών έχουν καταπιαστεί με το αντικείμενο της αυτοματοποιημένης καρδιακής αξιολόγησης και ακόμα πιο συγκεκριμμένα και με την κατηγοριοποίηση των φυσημάτων της παιδικής ηλικίας. Οι

δεσπόζουσες σε αυτές τις έρευνες μέθοδοι και μοντέλα, είναι τα κρυφά μαρκοβιανά μοντέλα (Hidden Markov Models - HMMs), τα τεχνητά νευρονικά δίκτυα (artificial neural networks - ANNs), Extreme Learning Machines, Support Vector Machines (SVMs), πολλές άλλες και φυσικά υβριδικά μοντέλα συνδυασμών τους. Αντίστοιχη πληθώρα υπάρχει και στα διάφορα χαρακτηριστικά που επιλέγονται για την μοντελοποίηση, με τα πιο συχνά να είναι οι MFCCs, το φασματογράφημα των σημάτων, το Energy Signature, το ανακατανεμημένο φασματογραφήμα (reassigned spectrogram) και άλλα.

Μέχρι και μόλις 7.07+/-3.65% φτάνει το ποσοστό σφάλματος (Error Rate) στην εργασία των Μαρκάκη Μαρίας, Γερμανάκη Ιωάννη και Στυλιανού Ιωάννη με τίτλο «Αυτόματη Ανίχνευση Φυσημάτων σε Παιδικά Φωνοκαρδιογραφήματα και Ανάπτυξη Λογισμικού Εκπαίδευσης στην Παιδιατρική Ακρόαση». Στην εργασία αυτή, η προεπεξεργασία και κατάτμηση των σημάτων ακολουθήθηκε από την εξαγωγή των χαρακτηριστικών από το ανακατανεμημένο φασματογράφημα (reassigned spectrogram), ενώ η αυτόματη ταξινόμηση σε παθολογικά και φυσιολογικά φυσήματα έγινε με τη χρήση SVM σε 25 πειράματα με 5-fold stratified cross-validation.

Σύγκριση διαφόρων ταξινομητών κάνουν στην εργασία με τίτλο “Cardiac Disorder Classification Based on Extreme Learning Machine (ELM)” οι Chul Kwak και Oh-Wook Kwon. Η αυτόματη κατάτμηση των σημάτων ακολουθείται από την εξαγωγή των χαρακτηριστικών, κατά την οποία μετά τον υπολογισμό του βραχέως μετασχηματισμού Fourier (Short Fourier Transform - SFT), υπολογίζονται οι mel-scale filter bank συντελεστές γνωστοί και ως MELSPEC που χρησιμοποιούνται. Ο ELM ταξινομητής, επιτυγχάνει επιτυχία αναγνώρισης φυσιολογικών καρδιακών σημάτων της τάξης του 97%, με αντίστοιχη επιτυχία από 44% ως 81% στην αναγνώριση καρδιακών σημάτων με διάφορες παθήσεις και μέση ακρίβεια του 81.6%. Στα πλαίσια της έρευνας αυτής, συγκρίθηκαν ως προς την ακρίβεια ταξινόμησης και άλλοι ταξινομητές όπως με χρήση HMMs, MLP και SVM με αντίστοιχες ακρίβειες 75.6%, 73.1% και 76.6%.

Χρησιμοποιώντας μεγάλο πλήθος δεδομένων ο Yong-Joo Chung στην εργασία του με τίτλο “A Classification Approach for the Heart Sound Signals Using Hidden Markov Models”, επιτυγχάνει πολύ μεγάλες ακρίβειες ταξινόμησης. Σκοπός της εργασίας του, ήταν να δείξει την καταλληλότητα των HMMs σαν μοντέλο για τέτοιου είδους ταξινόμηση αντί των ANNs, που ευραίως χρησιμοποιούνταν. Η συνολική ακρίβεια που πετυχαίνει ξεπερνά το 99% και πιο συγκεκριμμένα για τα φυσιολογικά σήματα είναι 100%, ενώ για τα σήματα με διάφορα ευρήματα η ακρίβεια είναι 97.48% και 98.11% για MFCC και Filterbank χαρακτηριστικά αντίστοιχα.

Μεγάλο ενδιαφέρον ως προς την χρησιμότητα και επεκτασιμότητα της παρουσιάζει η μελέτη των Matias Brusco και Homer Nazeran με τίτλο “Digital Phonocardiography: A PDA-

Based Approach”. Στην εργασία αυτή, παρουσιάζεται παραστατικά μια μέθοδος για την επεξεργασία και κατάλληλη απεικόνιση των φωνοκαρδιογραφηματων σε PDA που επιτρέπει την αναλυτική και απεικονιστική διάγνωση. Τέτοιες απλές τηλεϊατρικές εφαρμογές, λέγεται πως θα αποτελέσουν τη βάση για μια νέα εποχή στις τηλεματικές, τηλεϊατρικές και βιοιατρικές τεχνολογίες που συνεχώς αναπτύσσονται.

1.5 Συνεισφορά και Σκοπός της Εργασίας

Όπως αναπτύχθηκε και στην προηγούμενη ενότητα, η έγκαιρη και ακριβής διάκριση μεταξύ των φυσιολογικών και των παθολογικών φυσημάτων, είναι μεγάλης σημασίας αρχικά για το ίδιο το παιδί και την υγεία του και στη συνέχεια για ψυχολογικούς και οικονομικούς λόγους. Η σωστή αντίληψη ενός οργανικού φυσήματος που ακολουθείται και από την αντίστοιχη διάγνωση κάποιας καρδιαγγειακής ή βαλβιδικής πάθησης, θα δώσει την δυνατότητα στον θεράποντα Ιατρό έγκαιρα να αντιμετωπίσει την κατάσταση με τον καλύτερο τρόπο κι έτσι το παιδί με τη σειρά του να απαλλαχτεί από το πρόβλημα ή να εξασφαλίσει μια υψηλή ποιότητα ζωής. Η σωστή αντίληψη ενός φυσιολογικού φυσήματος δεν θα ανησυχίσει το παιδί και την οικογένεια του και δεν θα επιβαρύνει με επιπλέον κόστη επισκέψεων και εξετάσεων το ασφαλιστικό σύστημα. Βέβαια, όπως αναφέρθηκε στην προηγούμενη ενότητα η διάκριση αυτή δεν γίνεται με την επιθυμητή ακρίβεια, κύριως γιατί οι Παιδίατροι και οι γενικοί Ιατροί δεν διαθέτουν ικανοποιητική εμπειρία στην καρδιακή ακρόαση, ούτε απέκτησαν τέτοια εξειδίκευση κατά την εκπαίδευσή τους.

Αυτό το απαιτητικό, ενδιαφέρον και πολύ σημαντικό πρόβλημα, έχει προσελκύσει το ενδιαφέρον της Μηχανικής, η οποία προσπαθεί με διάφορα εργαλεία να το πραγματευτεί και να έρθει σε λύσεις μέσα από το σχεδιασμό κατάλληλων συστημάτων. Ο σκοπός αυτής της εργασίας, είναι η παρουσίαση και σύγκριση διαφόρων μεθόδων κατηγοριοποίησης παιδικών φυσημάτων. Η παρούσα εργασία είχε ως στόχο να καταλήξει στην πρόταση ενός τελικού συστήματος κατηγοριοποίησης, που να επιτυγχάνει ένα ικανοποιητικό ποσοστό ακρίβειας, χρησιμοποιώντας εργαλεία κυρίως από τον χώρο της Αναγνώρισης Φωνής και της Επεξεργασίας Σημάτων. Φυσικά η εργασία αυτή δεν έχει ως σκοπό να αντικαταστήσει το ιατρικό έργο, αλλά αποτελεί:

- Μια πρώτη προσέγγιση για το σχεδιασμό ενός υποστηρικτικού συστήματος για την λήψη απόφασης κατηγοριοποίησης.

- Μια προσέγγιση, μια βάση για το σχεδιασμό ενός συστήματος μαζικής εξέτασης σχολικού πληθυσμού παραμεθωρίων περιοχών.
- Μια βάση για το σχεδιασμό ενός συστήματος για εκπαιδευτική χρήση από φοιτητές Ιατρικής, εκπαιδευόμενους Ιατρούς κ.ά.
- Την βασική ιδέα για τον σχεδιασμό ενός πολύ απλού συστήματος κατηγοριοποίησης φυσημάτων για εφαρμογή ανδροειδούς συσκευής ή PDA κτλ.
- Την βάση για οποιαδήποτε εργασία και έρευνα που σκοπό έχει την επεξεργασία και κατηγοριοποίηση ηχητικών καρδιακών σημάτων, με χρήση εργαλείων αναγνώρισης φωνής, επεξεργασίας σημάτων κ.ά.

1.6 Οργάνωση της Αναφοράς της Διπλωματικής Εργασίας

Στο επόμενο κεφάλαιο γίνεται εκτενής ανάλυση και παρουσίαση του ιατρικού προβλήματος, που αυτή η εργασία πραγματεύεται. Γίνεται μάλιστα μια επεξηγηματική εισαγωγή στον καρδιακό ήχο, στα παρεμβαλόμενα σήματα και παρουσιάζονται τα επιμέρους τμήματα ενός ηχοκαρδιογραφήματος. Αναφέρονται τα παθολογικά και φυσιολογικά του χαρακτηριστικά, καθώς επίσης και στοιχεία για την διάγνωση αυτών και τα ενδεχόμενα προβλήματά της.

Στο 3^{ο} Κεφάλαιο αναλύονται οι διαδικασίες συλλογής και προεπεξεργασίας των δεδομένων (front-end) και πιο ειδικά ακόμα αναφέρεται στην κατάτμηση των ηχοκαρδιογραφημάτων καθώς επίσης και στην εξαγωγή των χαρακτηριστικών.

Το 4^{ο} Κεφάλαιο καταπιάνεται με την ανάπτυξη των ταξινομητών. Στο πρώτο μέρος του, αναλύεται η μέθοδος αναγνώρισης με χρήση στατικού ταξινομητή και συγκεκριμμένα βασισμένου στη μοντελοποίηση με μίγματα γκαουσιανών κατανομών ή Gaussian Mixture Models (GMMs), ενώ γίνεται μια εκτενής σύγκριση της χρήσης διάφορων χαρακτηριστικών. Στο δεύτερο μέρο,ς γίνεται σύγκριση μέσα από τη χρήση διάφορων ταξινομητών και μεθόδων με χρήση των εργαλείων του WEKA.

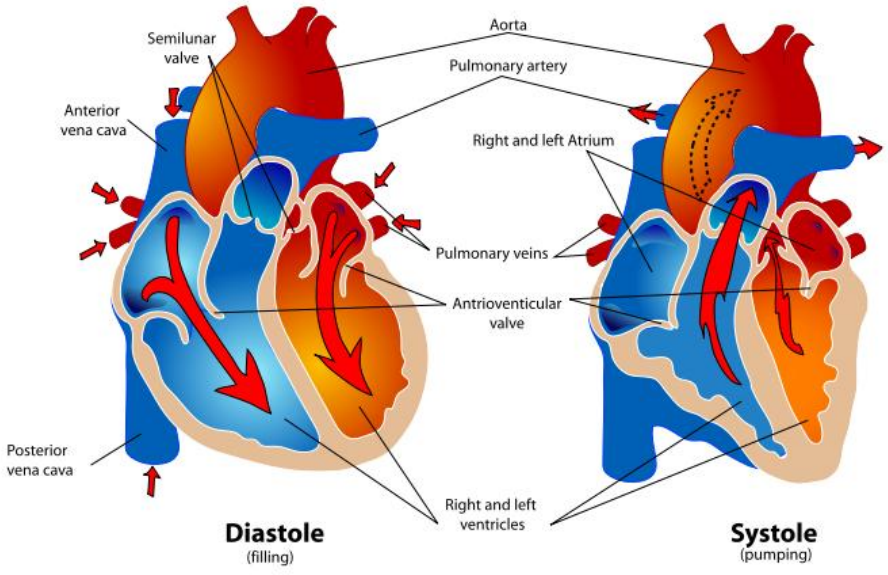
Το 5^{ο} Κεφάλαιο, εστιάζει στην πειραματική διαδικασία εν γένει, τον τρόπο που σχεδιάστηκαν και εκτελέστηκαν τα πειράματα και φυσικά παρατίθενται τα πειραματικά αποτελέσματα που προέκυψαν με τη χρήση των παραπάνω μεθόδων.

Στο 6^{ο} Κεφάλαιο, βρίσκεται ένας συμπερασματικός επίλογος στον οποίο παρουσιάζονται τα ωφέλη αυτής της εργασίας, καθώς και η ενδεχόμενη χρήση των αποτελεσμάτων της και τέλος προτάσεις για μελλοντικές βελτιώσεις και αναβαθμίσεις της παρούσας εργασίας.

2 Ανίχνευση Καρδιοπαθειών μέσω Ηχητικών Σημάτων Καρδιακής Λειτουργίας

2.1 Το Καρδιακό Ηχητικό Σήμα και τα Φυσιολογικά του Χαρακτηριστικά

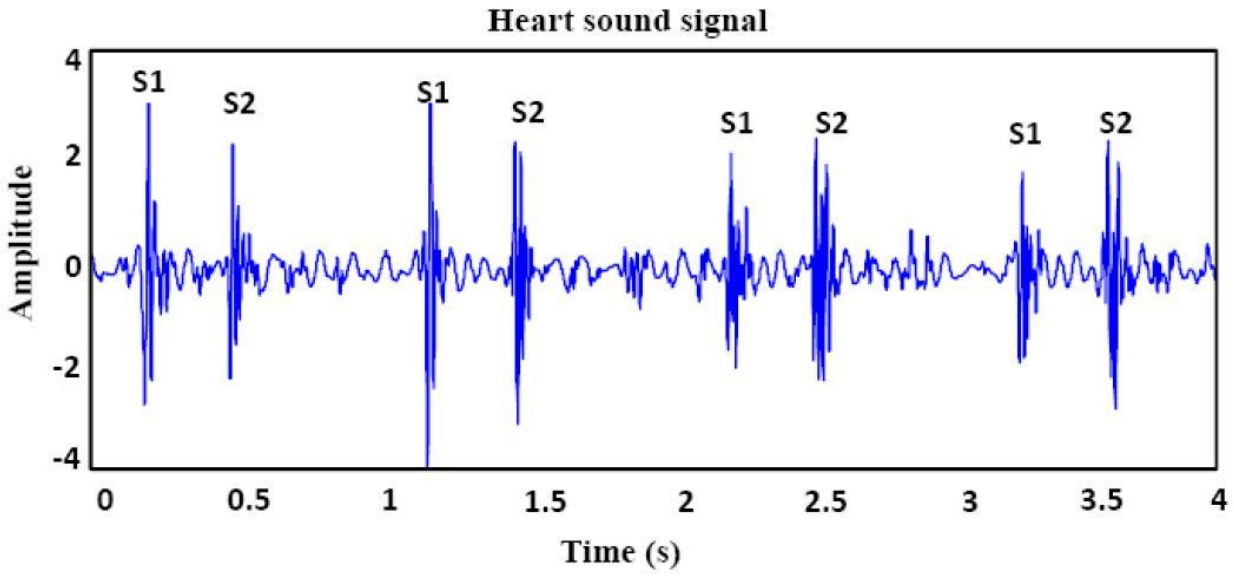
Αρκετά εύκολα μπορεί κανείς να αντιληφθεί το μηχανικό ανάλογο της καρδιάς, σαν μια δεξαμενή διασυνδεμένη σε ένα σύστημα σωληνώσεων μέσω βαλβίδων ελέγχου της ροής. Το καρδιακό ηχητικό σήμα, δεν είναι τίποτα άλλο από το ηχητικό αποτέλεσμα της μηχανικής λειτουργίας της καρδιάς. Η λειτουργία της καρδιάς είναι περιοδική, δηλαδή επαναλαμβάνει περιοδικά στο χρόνο ένα συγκεκριμμένο σύνολο ενεργειών. Τέτοιες λειτουργίες είναι απαραίτητες για να επιτραπεί η κυκλοφορία του αίματος προς και από την καρδιά και κυρίως πρόκειται για το άνοιγμα και κλείσιμο βαλβίδων και την εισροή και εξώθηση του αίματος. Σε ένα κύκλο τεσσάρων χρόνων η καρδιά καταφέρνει να οξυγονώσει το αίμα που δέχεται από την κυκλοφορία, στέλνοντας το στους πνεύμονες μέσω της πνευμονικής αρτηρίας και να το διοχετεύσει πίσω οξυγονωμένο πια στην κυκλοφορία μέσω της αορτής. Στο σχήμα 2.1, φαίνονται ξεκάθαρα οι λειτουργίες αυτές, καθώς με κόκκινο παρουσιάζονται οι περιοχές που περιέχουν οξυγονωμένο αίμα και με μπλέ το αντίθετο.



Σχήμα 2.1 – Οι δύο φάσεις τις καρδιακής λειτουργίας

Επιγραμματικά, όπως φαίνεται και στο σχήμα 2.1, η καρδιά προσλαμβάνει το αίμα από την κυκλοφορία μέσω της οπίσθιας κοίλης φλέβας και της ημισεληνοειδούς βαλβίδας και το εξωθεί προς τους πνεύμονες μέσω της πνευμονικής αρτηρίας, προκειμένου αυτό να οξυγονωθεί. Στη συνέχεια προσλαμβάνει το αίμα από τους πνεύμονες μέσω των πνευμονικών φλεβών και το διοχετεύει πίσω στην κυκλοφορία εξωθώντας το στην αορτή. Το αίμα εισέρχεται τόσο από την κυκλοφορία όσο και από τους πνεύμονες κατά τη διάρκεια της συστολής, όπου οι κοιλίες γεμίζουν με αίμα καθώς ανοίγουν οι βαλβίδες και εξωθείται προς τους πνεύμονες και την κυκλοφορία κατά την διάρκεια της συστολής.

Έχοντας κατανοήσει σαφώς την καρδιακή λειτουργία, μπορεί κανείς να αντιληφθεί πια την φυσική σημασία των ήχων που παράγονται. Έτσι λοιπόν ένας φυσιολογικός καρδιακός κύκλος αποτελείται από τον τόνο S1 που είναι ο ήχος που παράγεται από το κλείσιμο των κολποκοιλιακών βαλβίδων κατά το τέλος της διαστολής και από τον τόνο S2 που οφείλεται στο κλείσιμο των αρτηριακών βαλβίδων κατά το τέλος της συστολής. Το μεσοδιάστημα των τόνων S1 και S2 είναι η συστολή όπου η καρδιά συμπιέζεται εξωθώντας το αίμα τόσο στην πνευμονική, όσο και στην συστηματική κυκλοφορία ενώ το διάστημα που μεσολαβεί του S2 τόνου του τρέχοντος κύκλου και του S1 τόνου του επόμενου κύκλου είναι η διαστολή και κατά τη διάρκεια του η καρδιά εκπτύσεται καθώς γεμίζει με αίμα. Παρακάτω, φαίνεται το ηχητικό καρδιακό σήμα που παράγει μια υγειής καρδιά:



Σχήμα 2.2 – Το φωνοκαρδιογράφημα (PCG) 4 κύκλων του καρδιακού ήχου

Στο παραπάνω σχήμα διακρίνονται 4 κύκλοι του καρδιακού σήματος. Σημειώνονται στο σχήμα οι καρδιακοί τόνοι S1 και S2 που σηματοδοτούν και τις διάφορες καρδιακές φάσεις.

Κατά τη διάρκεια των μεσοδιαστημάτων των καρδιακών τόνων, δηλαδή της συστολής και της διαστολής δεν υπάρχει κάποιος ήχος, όπως άλλωστε φαίνεται και στο σχήμα. Αυτό γιατί η φυσιολογική ροή του αίματος μέσα στα αγγεία αλλά και στην καρδιά, είναι ομαλή και με φυσιολογικό ρυθμό και έτσι δεν παράγει κανένα θόρυβο. Στον πίνακα που ακολουθεί, αντιστοιχίζονται οι μηχανικές λειτουργίες της καρδιάς με τους αντίστοιχους ήχους ή απουσίες ήχων που προβλέπονται για μια φυσιολογική καρδιακή λειτουργία:

Α/Α	**Φάση**	**Μηχανικό συμβάν**	**Εύρημα ακρόασης**
1	Συστολή	Κλείσιμο κολποκοιλιακών βαλβίδων	Πρώτος καρδιακός τόνος (S1)
2	Συστολή	Άνοιγμα αρτηριακών βαλβίδων	Απουσία ήχου
3	Συστολή	Αποβολή αίματος	Απουσία ήχου
4	Διαστολή	Κλείσιμο αρτηριακών βαλβίδων	Δεύτερος καρδιακός τόνος (S2)
5	Διαστολή	Άνοιγμα κολποκοιλιακών βαλβίδων	Απουσία ήχου
6	Διαστολή	Παθητική εισροή αίματος	Απουσία τρίτου καρδ. τόνου (S3)
7	Διαστολή	Ενεργητική εισροή αίματος	Απουσία τέταρτου καρδ. τόνου (S4)

Πίνακας 2.3 – Ένας φυσιολογικός καρδιακός κύκλος σε λίστα γεγονότων

2.2 Μή Φυσιολογικά Χαρακτηριστικά του Καρδιακού Σήματος

Τα διάφορα παθολογικά ή ακριβέστερα μή φυσιολογικά ευρήματα, σχετίζονται είτε με ανωμαλίες των ήχων που φυσιολογικά υπάρχουν στον καρδιακό κύκλο είτε με την παρουσία επιπρόσθετων ήχων, ήχων δηλαδή που φυσιολογικά θα έπρεπε να απουσιάζουν από τον καρδιακό κύκλο. Στην περίπτωση όπου παρουσιάζονται αλλοιώσεις και ανωμαλίες σε κάποιο ήχο, αναφέρονται παρακάτω συνοπτικά τα πιο συχνά ευρήματα στην παιδιατρική ακρόαση:

Μεταβλητή ένταση των καρδιακών ήχων. Σχετίζεται κυρίως με την μεταβολή της έντασης του τόνου S1. Δεν είναι ιδιαίτερα συχνό εύρημα, ωστόσο αν η ύπαρξη χαμηλής συχνότητας τόνου S1 (βραδυκαρδία) συνδυάζεται με κατά τα άλλα φυσιολογικό καρδιακό ήχο είναι μια ένδειξη απουσίας συγχρονισμού των κολποκοιλιακών βαλβίδων, ενώ όταν το χαρακτηριστικό αυτό συνδέεται με γενικότερη ανομαλία στην καρδιακή λειτουργία, είναι μια ένδειξη για extra – συστολική αρρυθμία.

Η ομοιόμορφη ενίσχυση των καρδιακών ήχων. Πρόκειται για ένα χαρακτηριστικό που οφείλεται στις μικρές διαστάσεις του παιδικού θώρακα. Αυτό το ανατομικό χαρακτηριστικό, ενισχύει την μετάδοση του καρδιακού σήματος, με αποτέλεσμα οι καρδιακοί ήχοι να ακούγονται εντονότεροι και να διακρίνονται ευκολότερα από τους αντίστοιχους των ενηλίκων.

Η ομοιόμορφη εξασθένηση των καρδιακών ήχων. Το συγκεκριμμένο χαρακτηριστικό, κάνει την εμφάνισή του κυρίως σε περίπτωση εμφυσήματος, συγκέντρωσης περικαρδιακών υγρών και λίπους. Επίσης, μια τέτοια εξασθένηση γίνεται αντιληπτή κατά την εκτέλεση ακρόασης αριστερού προκαρδίου όταν το παιδί έχει δεξιοκαρδία.

Ο υψηλής έντασης καρδιακός τόνος S2. Το χαρακτηριστικό αυτό, αποτελεί ένδειξη πνευμονικής υπέρτασης όταν εντοπίζεται στη θέση της πνευμονικής ακρόασης.

Ο διαχωρισμός (splitting) του S2 τόνου. Αποτελεί ένα σύνηθες παθολογικό εύρημα στην παιδιατρική ακρόαση. Αποτελεί όπως και η λέξη περιγράφει, τον διαχωρισμό του τόνου σε δύο μέρη και ουσιαστικά μοιάζει με δυο διαφορετικούς ήχους πολύ κοντά χρονικά. Συνδέεται με την αρκετά επιταχυνμένη ροή του αίματος και με την ενδοκαρδιακή επικοινωνία.

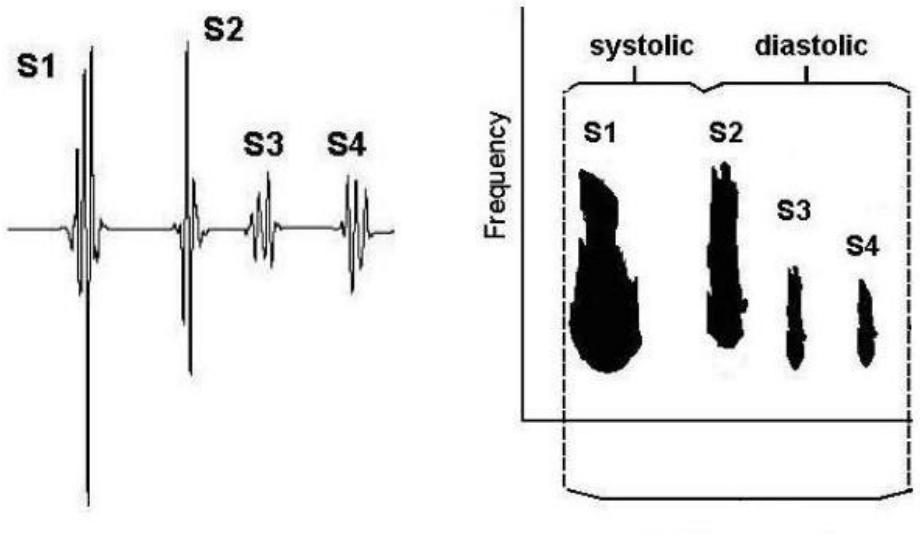
Ο διαχωρισμός του S1 τόνου. Αρκετά σπάνιος είναι αυτός ο διαχωρισμός. Στα παιδιά συνήθως δεν είναι τίποτα άλλο, από ένα φυσιολογικό τόνο S1 ακολουθούμενο από κάποιο πρώιμο συστολικό φύσημα.

Η φλεβοκομβική ταχυκαρδία. Πρόκειται για την επιταγχυμένη κατά τα άλλα φυσιολογική καρδιακή λειτουργία. Στα βρέφη χωρίς την παρουσία κάποιας παθολογίας, ο καρδιακός ρυθμός μπορεί να φτάνει και το 180, ενώ στα παιδιά το 120 με 130 κατά τη διάρκεια της καρδιολογικής εξέτασης, ιδίως όταν βρίσκονται σε κατάσταση άγχους. Το χαρακτηριστικό αυτής της ταχυκαρδίας, είναι η μεταβλητή της συχνότητα.

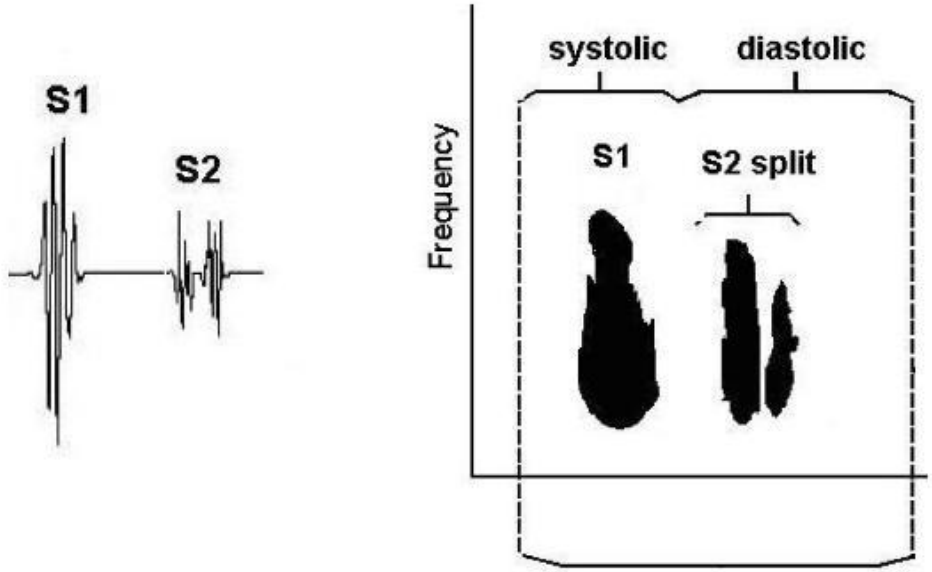
Η αναπνευστική αρρυθμία. Ο καρδιακός ρυθμός προσαρμόζεται έτσι, ώστε να κρατήσει σταθερή ροή αίματος στην έξοδο της καρδιάς. Μεταβάλλεται γιατί κατά την εισπνοή και την εκπνοή συγκεντρώνεται μεγαλύτερη ποσότητα αίματος στην δεξιά και αριστερή κοιλία αντίστοιχα. Ο ασταθής ρυθμός αναπνοής στα παιδιά, δημιουργεί τέτοιες μεταβολλές στο καρδιακό ρυθμό που παραπέμπουν σε παθολογική αρρυθμία.

Οι δυνατοί S3 – S4 τόνοι. Εντοπίζονται μόνο με προσεκτική ακρόαση και είναι συνήθως χαμηλής συχνότητας. Όταν το εύρημα αυτό συνδυάζεται με την ταχυκαρδία, χαρακτηρίζεται ως ρυθμός Gallop και είναι από τα συνήθη ευρήματα στην παιδιατρική ακρόαση.

Ακολουθούν μερικά σχήματα που απεικονίζουν κάποια από τα πιο συνηθισμένα ευρήματα.

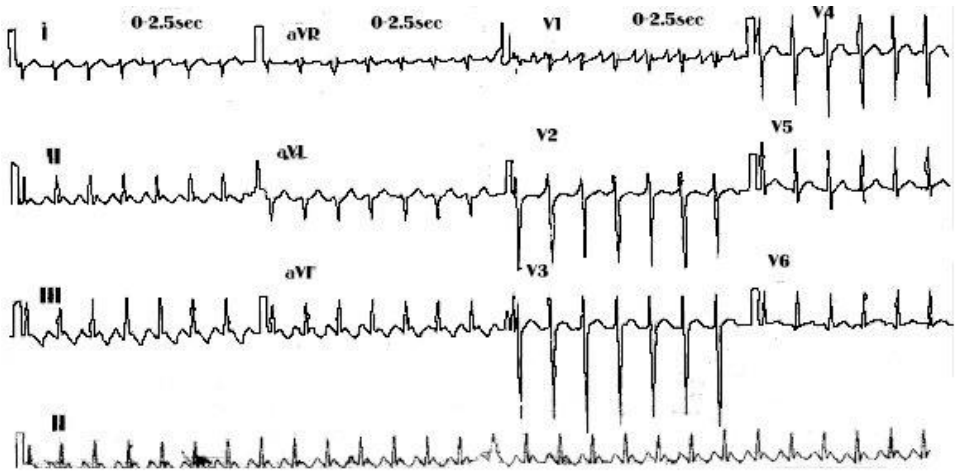


Σχήμα 2.4 – Καρδιακός κύκλος με επιπρόσθετους ήχους S3 και S4



Σχήμα 2.6 – Καρδιακός κύκλος με διαχωρισμό του τόνου S2

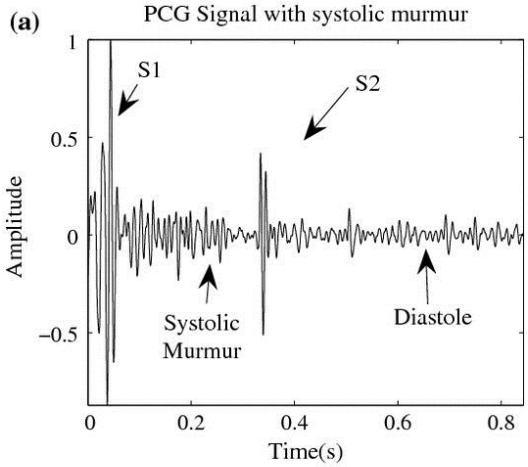
Παραπάνω απεικονίζονται τόσο τα PCGs, δηλαδή το ηχητικό σήμα, όσο και ο μετασχηματισμός στο πεδίο της συχνότητας σε αυτή την απεικόνιση συχνότητας – χρόνου που βρίσκεται στο δεξί τμήμα των σχημάτων 2.5 και 2.6.



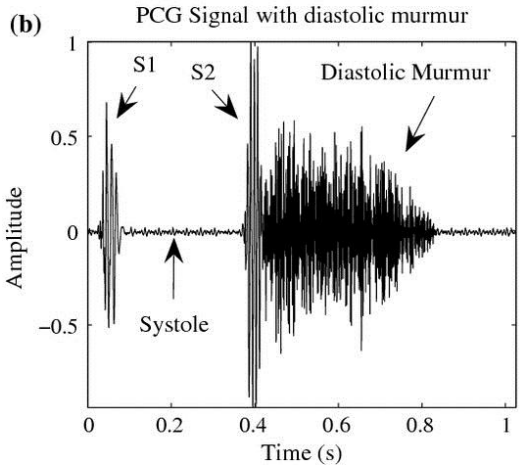
Σχήμα 2.5 – ECG καρδιακών κύκλων που παρουσιάζουν φλεβοκομβική ταχυκαρδία

Στην δεύτερη κατηγορία, εντάσσονται οι επιπρόσθετοι ήχοι που ακούγονται κατά τη διάρκεια του καρδιακού κύκλου (π.χ. φυσήματα). Για τον χαρακτηρισμό τους (παθολογικότητα ή μη) και φυσικά για την διάγνωση πρέπει να απαντηθούν αρκετά ερωτήματα, που κυριώς αφορούν στην ύπαρξή του αρχικά και στην συνέχεια τις διάφορες ιδιότητες των ήχων αυτών. Τα σημαντικότερα ερωτήματα είναι:

- Αν υπάρχει ένας ή περισσότεροι διακριτοί επιπρόσθετοι ήχοι.
- Αν είναι σύντομου χρόνου όπως ο τόνος S3, η κλαγγή (click) και άλλα.
- Αν είναι μεγαλύτερης διάρκειας όπως το φύσημα.
- Από ποιά θέση ακρόασης ακούγεται καλύτερα. Οι θέσεις της παιδιατρικής ακρόασης αναλύονται στην επόμενη ενότητα.
- Σε ποιο διάστημα του καρδιακού κύκλου τοποθετείται χρονικά. Συστολή, διαστολή, και τα δύο, είναι οι πιθανές περιπτώσεις.
- Ποια η διάρκεια του σε σχέση με το διάστημα που βρίσκεται. Πρώιμη συστολή, ολοσυστολή, μέση συστολή είναι μερικά παραδείγματα τέτοιων χαρακτηρισμών.
- Ποια η συχνότητά του. Χαμηλής, μέσης και υψηλής συχνότητας.
- Ποια η μεταβλητότητα της έντασης του στο χρόνο. Αυξανόμενη ένταση (crescendo), μειούμενη (decrescendo), αυξομειούμενης (crescendo – decrescendo) και σταθερής.
- Τέλος, ποιά είναι η κατάλληλη διάγνωση που αντιστοιχεί στο εκάστοτε εύρημα. Δηλαδή, μια φυσική εξήγηση για το που οφείλεται η ύπαρξη και τα επιμέρους χαρακτηριστικά του διακριτού επιπρόσθετου ήχου που παρατηρήθηκε.



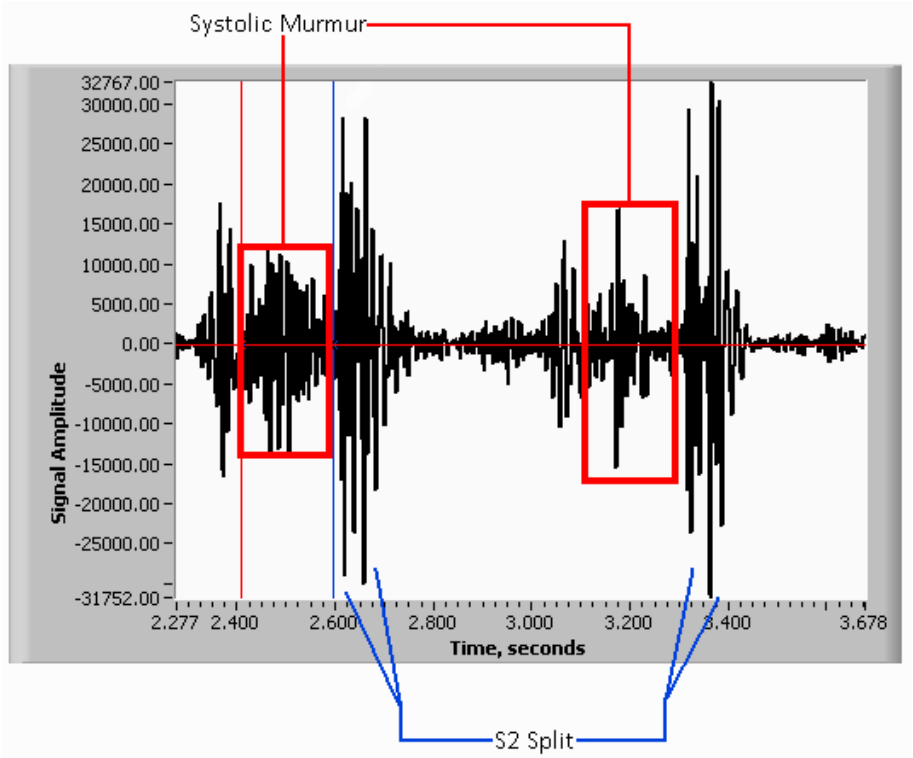
Φαίνεται στο πάνω μέρος του σχήματος (2.3.a), ένας καρδιακός κύκλος στον οποίο εντοπίζεται φύσημα. Το φύσημα αυτό χαρακτηρίζεται συστολικό και μάλιστα είναι σχετικά χαμηλής decrescendo έντασης. Στη διάρκεια της διαστολής δεν ακούγεται κάποιος ήχος, όπως φυσιολογικά. Τα συστολικά φυσήματα συνήθως οφείλονται στο μή κλείσιμο των κολποκοιλιακών βαλβίδων, στο μή άνοιγμα των αρτηριακών, σε κάποια ενδοκαρδιακή επικοινωνία ή σε αγγειακές στενώσεις.



Στο κάτω μέρος του ίδιου σχήματος (2.3.b) φαίνεται ένα έντονο διαστολικό φύσημα, crescendo – decrescendo υψηλής έντασης. Στη διάρκεια της συστολής δεν παρατηρείται κανένας ήχος. Πιθανό αίτιο το να μην κλείνουν οι αρτηριακές βαλβίδες ή να μην ανοίγουν (στένωση) οι κολποκοιλιακές.

Σχήμα 2.6 –Συστολικό και διαστολικό φύσημα

Τέλος, στο σχήμα 2.7 που ακολουθεί απεικονίζεται το PCG ενός καρδιακού ήχου δύο κύκλων στο οποίο παρατηρείται ολοσυστολικό φύσημα, crescendo – decrescendo, ισχυρής έντασης παράλληλα με ευρύ διχασμό του δεύτερου τόνου (S2).



Σχήμα 2.7 – Δυο καρδιακοί κύκλοι με συστολικό φύσημα και διαχωρισμό του S2

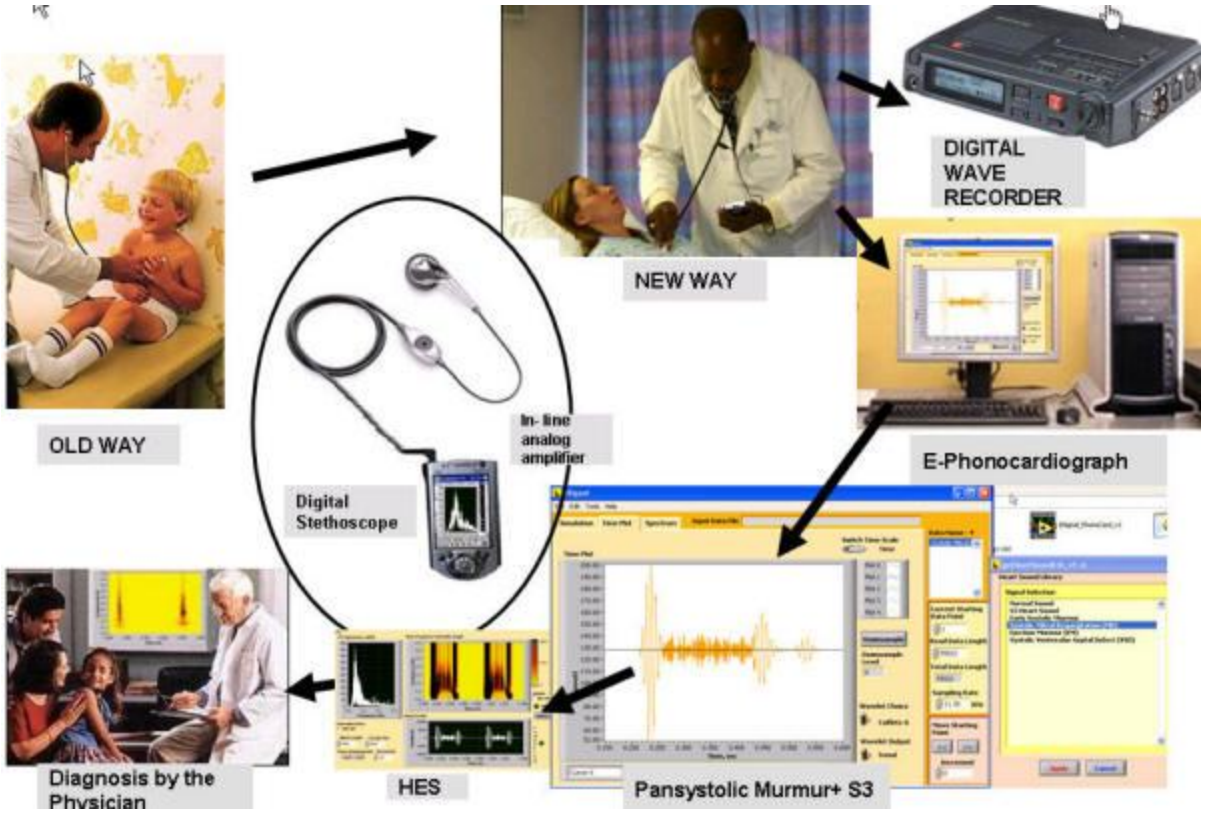
2.3 Η Καρδιακή Ακρόαση

Η μέθοδος που ενδείκνυται για την αξιολόγηση της καρδιακής κατάστασης, είναι αυτή της καρδιακής ακρόασης. Ακούγοντας την καρδιά, μπορεί κανείς να συγκεντρώσει πληροφορίες σχετικά με τη συχνότητα και το ρυθμό του καρδιακού ήχου, με τυχόν δυσλειτουργίες του καρδιαγγειακού συστήματος και με τυχόν ανατομικές ανωμαλίες του. Μέσα από την αναγνώριση των φυσιολογικών και παθολογικών χαρακτηριστικών του ηχητικού σήματος της καρδιάς, μπορεί ο Ιατρός να αξιολογήσει την κατάσταση και να συγκεντρώσει τις παραπάνω πληροφορίες που αναφέρθηκαν. Πρόκειται για μια πολύ απλή, σύντομη και μηδαμινού κόστους διαδικασία που χρησιμοποιείται εδώ και πάρα πολλά χρόνια από τους γιατρούς.

Ο τύπος αυτός εξέτασης, διενεργείται με την χρήση του στηθοσκόπιου. Καθώς τα στηθοσκόπια χρησιμοποιούνται για αυτό το έργο, είναι κατασκευασμένα βάσει συγκεκριμμένων προδιαγραφών. Το διάφραγμα του στηθοσκοπίου ειναι καταλληλότερο για την ανίχνευση ήχων υψηλής συχνότητας, σε αντίθεση με τον κώδωνα, που είναι καταλληλότερος για ήχους χαμηλών συχνοτήτων. Τα τελευταία χρόνια, χρησιμοποιούνται σύγχρονα ψηφιακά στηθοσκόπια τα οποία διαθέτουν και διατάξεις για την ενίσχυση της έντασης. Το μεγαλύτερο προτέρημα ωστόσο των ψηφιακών στηθοσκοπίων, είναι η συνδεσιμότητά τους με συστήματα υπολογιστών, πράγμα που επιτρέπει την εύκολη αποθήκευση των καρδιογραφημάτων, για ύστερη ή απομακρυσμένη αξιολόγηση (remote evaluation) που επιτρέπει και την ανάπτυξη συστημάτων τηλεϊατρικής (telemedical systems). Τα τελευταία ψηφιακά στηθοσκόπια, παρέχουν την δυνατότητα ταυτόχρονης καταγραφής τόσο του φωνοκαρδιογραφήματος (PCG), όσο και του ηλεκτροκαρδιογραφήματος (ECG) σε δύο διαφορετικά κανάλια.

Σχήμα 2.8 – Καρδιακή ακρόαση με καρδιολογικό στηθοσκόπιο

Στα παιδιά, η καρδιακή ακρόαση γίνεται είτε με το παιδικό είτε με το συμβατικό ή καρδιολογικό στηθοσκόπιο. Το παιδικό στηθοσκόπιο επιτρέπει την ανίχνευση ήχων πολύ υψηλών συχνοτήτων, ενώ το καρδιολογικό, ήχων χαμηλής συχνότητας. Κάθε είδος στηθοσκοπίων, έχει την δική του περιοχή ανίχνευσης και μόνο γνωρίζοντας τα θετικά και αρνητικά χρήσης του, μπορεί κανείς να επιλέξει. Συνιστάται δε κανείς, να προτιμά το είδος στηθοσκοπίου που χρησιμοποίησε στην καρδιολογική του εκπαίδευση.



Σχήμα 2.9 – Σύστημα καρδιολογικής εξέτασης με χρήση ηλεκτρονικού στηθοσκοπίου και συστήματος καταγραφής

Στο σχήμα 2.9, φαίνεται ένα απλό τηλεϊατρικό σύστημα, το οποίο βασίζεται στη χρήση ψηφιακού στηθοσκοπίου. Η διάταξη του καταγραφέα κυματομορφών, επιτρέπει την αποθήκευση των καρδιογραφημάτων σε Η/Υ όπου με κατάλληλο πρόγραμμα απεικόνισης και επεξεργασίας, δίνεται η δυνατότητα να γίνει αξιολόγηση του PCG, του ECG αλλά και του σπεκτρογράμματος. Τέτοιες διατάξεις επιτρέπουν την εξ' αποστάσεως αξιολόγηση, την επανεξέταση ανά πάσα στιγμή, την διευκόλυνση εκπαιδευτικών διαδικασιών μέσω μελετών υποθέσεων (case studies).

Η καρδιακή ακρόαση όπως αναφέρθηκε εξ' αρχής, έχει ως σκοπό τον εντοπισμό και την αξιολόγηση του καρδιακού ηχητικού σήματος. Ο εντοπισμός, αλλά ακόμα περισσότερο η σωστή αξιολόγησή του, έχει ως απαραίτητη προϋπόθεση τη διενέργεια της ακρόασης σε συγκεκριμμένες θέσεις (cardiac auscultation sites). Συγκεκριμμένοι ήχοι σε συγκεκριμμένη θέση ακρόασης, σημαίνουν πολύ συγκεκριμμένα πράγματα για την καρδιακή λειτουργία και μόνο η προσεκτική ακρόαση σε συνδυασμό με την επίγνωση της ανατομίας της θέσης της και

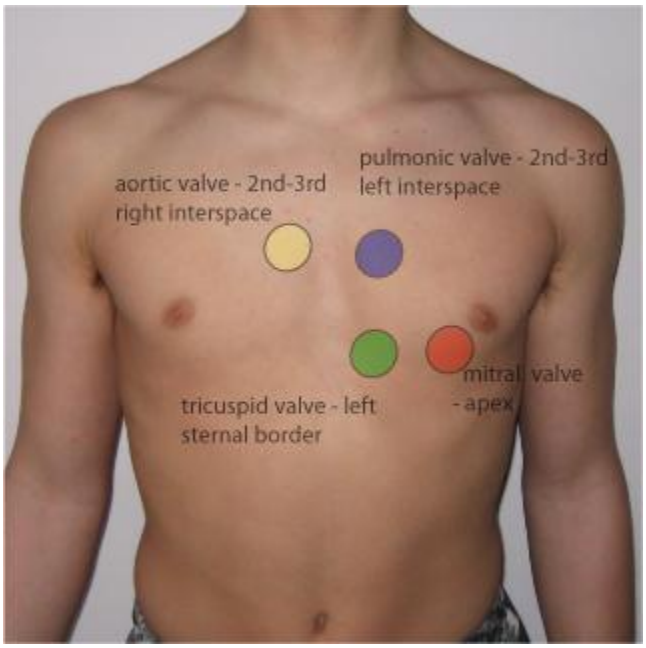
την γενικότερη κλινική εικόνα του παιδιού, επιτρέπουν την ορθή αξιολόγηση και διάγνωση. Ακολουθεί ο πίνακας 2.10 στον οποίο αναφέρονται οι θέσεις καρδιακής ακρόασης και τα αντίστοιχα ευρήματα, φυσιολογικά και μή, που συνδέονται με τις θέσεις αυτές.

Θέση Ακρόασης	**Φυσιολογικό εύρημα**	**Μή φυσιολογικό εύρημα**
Κορυφή	S1 (μητροειδής βαλβίδα)	Μεσοσυστολική κλαγγή Πρώιμη συστολική κλαγγή Παλινδρόμηση Μητροειδούς Βαλβίδας Στένωση Μητρ. Βαλβίδας S3, S4
Κάτω αριστερή παραστερνική	S1 (τριγλώχινα βαλβιδα)	Παλινδρόμηση Τριγλώχινας Βαλβίδας Αθώο φύσημα
Μέση αριστερή παραστερνική		Παλινδρόμηση Αορτής Παλινδρόμηση Πνευμονικής Βαλβίδας Έλλειμμα Μεσοκοιλιακού Διαφράγματος (VSD)
Άνω αριστερή παραστερνική	S2 (πνευμονική βαλβίδα)	Στένωση Πνευμονικής Βαλβίδας Αθώο φύσημα
Άνω δεξιά παραστερνική	S2 (αορτική βαλβίδα)	Στένωση βαλβίδας αορτής
Σφαγιτιδικός βόθρος		Στένωση βαλβίδας αορτής Στένωση Αορτής Φλεβικό βουητό
Υποκλείδια περιοχή		Αρτηριακή στένωση
Μασχαλιαία		Περιφερειακή πνευμονική στένωση
Περιοχή ωμοπλάτης		Στένωση Αορτής

Σχήμα 2.10 – Πίνακας των καρδιακών αρκοαστικών ευρημάτων ανά θέση ακρόασης

Η θέση ακρόασης στην οποία η ένταση της καρδιακής ανωμαλίας (για παράδειγμα φύσημα) μεγιστοποιείται, καλείται punctum maximum (p.m.) και προσδιορίζεται έπειτα από διεξοδική ακρόαση όλων των θέσεων ακρόασης, στην προκαρδιακή περιοχή. Η γνώση του punctum maximum, είναι απαραίτητη για την ακριβή διάγνωση. Σε περίπτωση που παρατηρείται και ροίζος (thrill), κάτι που γίνεται αντιληπτό με το άγγιγμα της παλάμης στο θώρακα του παιδιού και είναι ουσιαστικά κάποια δόνηση, η γνώση για το punctum maximum γίνεται αρκετά πιο εύκολα, με τον εντοπισμό του σημείου που ο ροίζος γίνεται αισθητός.

Συχνά στη παιδιατρική ακρόαση, πολύ μικρά VSDs είναι η αιτία για φυσήματα που γίνονται αντιληπτά, μόνο σε ορισμένες μικρές περιοχές του προκαρδίου. Αν το διάφραγμα δεν τοποθετηθεί με ακρίβεια πάνω στις περιοχές αυτές, τότε θα αποτύχει ο εντοπισμός των φυσημάτων αυτών. Οι σημαντικότερες θέσεις ακρόασης σε παιδιά σε κάθε καρδιακή ακρόαση φαίνονται στο σχήμα 2.11. Πρόκειται για τις θέσεις των βαλβίδων, που εξετάζονται σχολαστικά σε συνδυασμό με την κορυφαία θέση (apex) και επιπρόσθετη εξέταση με την παλάμη για ανίχνευση τυχόν ροίζου.



Σχήμα 2.12 – Οι θέσεις των βαλβίδων στον παιδικό θώρακα

Τέλος, να αναφερθεί ότι στην ακρόαση γίνονται όλες οι απαραίτητες ενέργειες για τον προσδιορισμό όλων των χαρακτηριστικών, που διέπουν τους μη φυσιολογικούς ήχους που ανιχνεύονται, όπως αναφέρονται στην προηγούμενη ενότητα. Για τον πολύ σημαντικό παράγοντα του χρονισμού, χρησιμοποιούνται διάφορες τεχνικές, όπως για παράδειγμα το χτύπημα του δακτύλου στο άκουσμα του τόνου S2. Συνεπώς, ότι ακούγεται πρωτύτερα ανήκει στην συστολική φάση, ενώ ότι έπεται στην διαστολική, έτσι μπορεί ο εξετάζων Ιατρός, να ταυτοποιήσει τους φυσιολογικούς ήχους που ακούει (σύμφωνα με τον πίνακα 2.3) και να εντοπίσει τυχόν ανωμαλίες τοποθετώντας τες χρονικά ως προς τον καρδιακό κύκλο.

Στην ενότητα που ακολουθεί, αναλύεται το ερώτημα, γιατί η μέθοδος της καρδιακής ακρόασης ενέχει πολλά προβλήματα και το πως τα προβλήματα αυτά δυσχεραίνουν την καρδιακή αξιολόγηση και διάγνωση του παιδιού;

2.4 Προβλήματα και Δυσκολίες της Καρδιακής Ακρόασης

Παρότι όπως προαναφέρθηκε, η ακρόαση είναι η προβλεπόμενη μέθοδος για τον έλεγχο της καρδιαγγειακής υγείας, όπως καταμαρτηρείται από πολλές έρευνες η ακρίβεια της παρουσιάζει μεγάλες μεταβολές και είναι συνάρτηση της εμπειρίας και ειδίκευσης του εξετάζοντα Ιατρού. Μόνο έμπειροι Καρδιολόγοι και Παιδοκαρδιολόγοι μπορούν με μεγάλη ακρίβεια, συνήθως μεγαλύτερη του 90%, να διακρίνουν τα παθολογικά φυσήματα από τα φυσιολογικά. Μέτρια ως χαμηλή ακρίβεια χαρακτηρίζει την διαχωριστική ικανότητα Παιδιάτρων και γενικών Ιατρών, ενώ για τους φοιτητές Ιατρικής και τους ειδικευόμενους Ιατρούς η διαχωριστική τους ικανότητα δεν ξεπερνά την ακρίβεια του 50%. Εύκολα καταλαβαίνει κανείς, ότι τα παιδιά υπό ορισμένες συνθήκες κινδυνεύουν να μην τους διαγνωστεί κάποια ασυμπτωματική καρδιακή πάθηση, ή να παραπεμπφθούν για εξειδικευμένες καρδιολογικές εξετάσεις, με τα αντίστοιχα συναισθηματικά και χρηματικά κόστη που αυτό συνεπάγεται.

Πέρα όμως από την δυσκολία που προκύπτει στην καρδιακή ακρόαση λόγω απειρίας ή μή εξειδίκευσης, υπάρχουν πολλοί άλλοι παράγοντες που την δυσχεραίνουν. Το καρδιακό σήμα όπως έχει ήδη τονιστεί,, είναι και ηχητικό σήμα. Σαν ηχητικό σήμα υπόκειται σε αλλοιώσεις, εξασθένηση, ενίσχυση κι επίδραση θορύβου και αυτό μπορεί να οδηγήσει σε επισφαλή αξιολόγηση και συμπεράσματα. Το καρδιακό ηχητικό σήμα αλλοιώνεται σε ειδικές περιπτώσεις, όπου παρεμβάλλεται της καρδιάς και του στηθοσκοπίου αέρας, λόγω πνευμονοθώρακα ή εμφυσήματος, υγρό λόγω περικαρδίτιδας ή λίπος. Αυτή η παραμόρφωση και εξασθένηση που επιδέχεται το σήμα στις περιπτώσεις αυτές, λειτουργεί αρνητικά ως προς την ευκολία και ακρίβεια της αξιολόγησης.

Όσο μικρότερης ηλικίας είναι οι ασθενείς, τόσο πιο εξειδικευμένη ακρόαση απαιτείται. Λόγω της ανατομίας τους, στα παιδιά παρατηρείται αυξημένη ένταση των καρδιακών τους ήχων, λόγω του μικρού τους θώρακα και αυτό διευκολύνει την διάδοση του καρδιακού ηχητικού σήματος. Οι διάφοροι ήχοι του καρδιακού σήματος, παρατηρούνται με πολύ μεγαλύτερη ένταση από το φυσιολογικό. Το αντίθετο συμβαίνει στην περίπτωση της δεξιοκαρδίας, όπου κατά την εκτέλεση ακρόασης στις αριστερές θέσεις παρουσιάζεται

ιδιαίτερη εξασθένηση του καρδιακού σήματος και τότε θα πρέπει η καρδιακή ακρόαση να γίνεται σε θέσεις συμμετρικές προς τις προβλεπόμενες, αλλά από την δεξιά πλευρά του θώρακα του παιδιού.

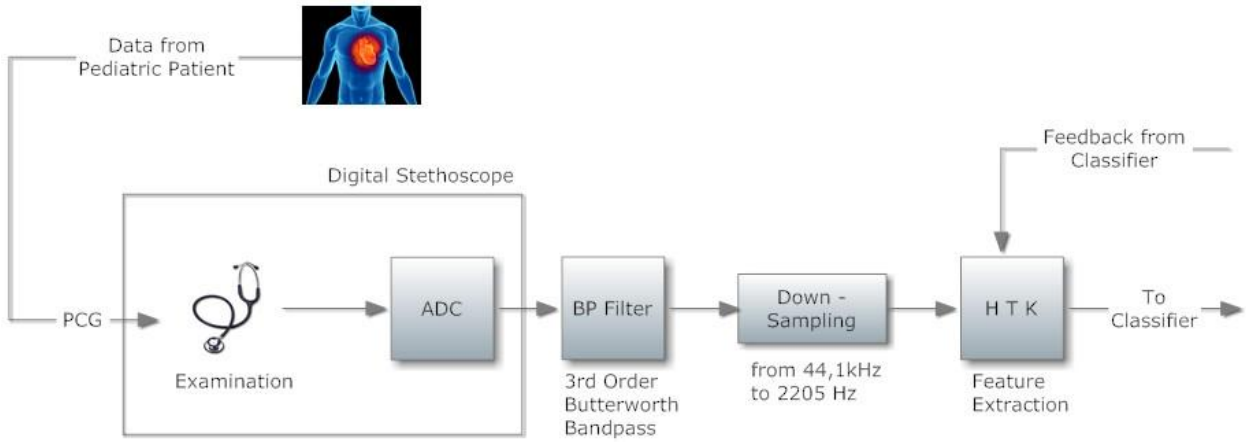
Αντιλαμβανόμενοι την διαδικασία της καρδιακής ακρόασης σαν ένα τηλεπικοινωνιακό σύστημα, θα λέγαμε ότι το καρδιακό σήμα εκτίθεται σε ένα σύνολο θορύβων που αλλοιώνουν το περιεχόμενο του. Εμπλεκόμενα σήματα προστίθενται στο καρδιακό σήμα ήδη από το σώμα του παιδιού, καθώς λαμβάνονται από το στηθοσκόπιο, με αποτέλεσμα να καταλήγουν στο αυτί του εξετάζοντα Ιατρού. Αυτά τα λεγόμενα εμπλεκόμενα σήματα, δεν είναι τίποτα άλλο από εσωτερικούς ήχους, όπως για παράδειγμα η αναπνοή, οι στομαχικοί και εντερικοί θόρυβοι, οι θόρυβοι από την λειτουργία διαφόρων οργάνων του παιδιού και σε περίπτωση παιδιών μικρής ηλικίας, το κλάμα κατά τη διάρκεια της εξέτασης. Επιπλέον, ο εξωτερικός θόρυβος, είναι άλλος ένας παράγοντας που επιδρά στην ποιότητα του καρδιακού ηχητικού σήματος. Είναι αρκετά συνηθισμένο πρόβλημα, ιδίως όταν η εξέταση γίνεται σε κάποιο ιατρείο νοσοκομείου ή κλινικής ή σε παιδιατρεία, όπου ο θόρυβος από άλλους ασθενείς, από ομιλίες, από ενέργειες άλλων Ιατρών και νοσηλευτών, επιδρούν στο τελικό ηχητικό αποτέλεσμα που ο εξετάζων Ιατρός αντιλαμβάνεται. Τέλος, το γεγονός ότι τα παιδιά, ιδίως δε τα μικρά σε ηλικία, δεν είναι ιδιαίτερα συνεργάσιμα, είναι γνωστό στο σύνολο των αντίστοιχων Ιατρών από τα προβλήματα που δημιουργούνται στις εξετάσεις. Η επιτάχυνση του καρδιακού ρυθμού λόγω φόβου, ο ασταθής ρυθμός της αναπνοής που επηρεάζει τον ρυθμό ροής του αίματος, το κλάμα, οι κινήσεις και άλλα είναι πολύ συνηθισμένα προβλήματα κατά την εξέταση παιδιών.

Εν κατακλείδι, αξίζει να αναφερθεί ότι ο εντοπισμός του φυσήματος δεν συνεπάγεται και σωστή κατηγοριοποίηση του, διάκριση μεταξύ φυσιολογικού και παθολογικού, και άρα διάγνωση. Ο εντοπισμός ενός φυσήματος, είναι μόνο η αρχή για μια ενδελεχή ακρόαση, με σκοπό να γίνουν αντιληπτά συγκεκριμμένα χαρακτηριστικά που θα επιτρέψουν την σωστή τελικά διάγνωση. Η αποσαφήνιση όλων των χαρακτηριστικών που περιγράφονται στην ενότητα 2.2, απαιτεί τεχνική γνώση και εμπειρία γύρω από την καρδιακή ακρόαση, μια διαδικασία που εμπεριέχει όλες τις δυσκολίες που προαναφέρθηκαν. Έτσι λοιπόν, όλοι αυτοί οι παράγοντες καθιστούν την διάγνωση με βάση την ακρόαση, αρκετά δύσκολη και ανεπαρκώς αποτελεσματική, όταν αυτή διενεργείται από μή ειδικούς Καρδιολόγους ή Παιδοκαρδιολόγους ή από γενικούς Ιατρούς με ελάχιστη εμπειρία στην ακρόαση.

3 Συλλογή και Προεπεξεργασία των Δεδομένων

3.1 Γενικά

Σε αυτό το κεφάλαιο θα αναλυθούν όλα τα στάδια λήψης, ανάλυσης και επεξεργασίας των δειγμάτων που προηγούνται της ταξινόμησης. Αφού τέθηκε και αναγνωρίστηκε τόσο το γενικό πλαίσιο του προβλήματος, όσο και το ιατρικό υπόβαθρό του, το επόμενο στάδιο προς υλοποίηση είναι αυτό της συλλογής και αποθήκευσης των δεδομένων. Ο αντικειμενικός σκοπός του προβλήματος που η εργασία αυτή πραγματεύεται, είναι η εκπαίδευση στατιστικών μοντέλων, ικανών να αποφαίνονται την φυσιολογική ή παθολογική παρουσία φυσημάτων σε παιδιά. Ως εκ τούτου, δεν χρησιμοποιήθηκε κάποια μέθοδος αυτόματης συσταδοποίησης (clustering) των δεδομένων, αλλά η Ιατρική ομάδα του κ. Γερμανάκη σημείωσε τις αντίστοιχες γνωματεύσεις για κάθε καταγραφή κι έτσι μπορούσαμε στο στάδιο της υλοποίησης, να χρησιμοποιήσουμε την πληροφορία αυτή. Όπως προαναφέρθηκε στο προηγούμενο κεφάλαιο, κατά την ακρόαση και καταγραφή των καρδιογραφημάτων, παρατηρούνται πολλά εμπλεκόμενα σήματα. Αυτό το πρόβλημα, λύνεται με ανάλυση των συχνοτικών περιεχομένων όλων των εμπλεκόμενων σημάτων και κατάλληλο φιλτράρισμα του καρδιακού σήματος.



Σχήμα 3.1 – Σχεδιάγραμμα της φάσης προεπεξεργασίας σήματος και εξαγωγής χαρακτηριστικών

Τέλος, μετά και από την κατάλληλη υποδειγματοληψία, ακολουθεί και η διαδικασία της εξαγωγής χαρακτηριστικών, που ολοκληρώνει το στάδιο αυτό της προ-επεξεργασίας των

σημάτων, γνωστό και ως front – end. Έπειτα, τα χαρακτηριστικά είναι έτοιμα προς χρήση από τον ταξινομητή. Η διαδικασία αυτή όπως φαίνεται και στο σχήμα 3.1, επιδέχεται την ανατροφοδότηση από τον ίδιο τον ταξινομητή κι έτσι το στάδιο της εξαγωγής των χαρακτηριστικών, βελτιστοποιείται πειραματικά με κριτήριο το ποσοστό ορθής κατηγοριοποίησης. Επιμέρους θα αναλυθούν διεξοδικά όλα τα στάδια αυτά που προαναφέρθηκαν, από τη συλλογή των δεδομένων μέχρι και την εξαγωγή των χαρακτηριστικών σε αυτό το κεφάλαιο.

3.2 Συλλογή και Αξιολόγηση των Δεδομένων

Η συλλογή των δεδομένων είναι η αρχή της υλοποίησης κάθε συστήματος αναγνώρισης προτύπων. Στο πείραμά μας, η συλλογή των δεδομένων ήταν μια χρονοβόρα διαδικασία, που διενεργήθηκε από την 1^{η} Ιανουαρίου 2008 μέχρι την 30^{η} Ιουνίου 2009. Μέσα σε αυτό το χρονικό διάστημα, η ομάδα με επικεφαλής τον Παιδοκαρδιολόγο Ιατρό Ιωάννη Γερμανάκη, συνέλεξαν το σύνολο των καταγραφών που χρησιμοποιήθηκε με χρήση ηλεκτρονικού στηθοσκόπιου, που με την τοποθέτηση 3 ηλεκτροδίων έγινε εφικτή η ταυτόχρονη καταγραφή τόσο του ψηφιακού φωνοκαρδιογραφήματος (PCG) όσο και του ηλεκτροκαρδιογραφήματος (ECG). Τα παιδιά των οποίων η εξέταση καταγράφηκε, είτε έχουν παραπεμφθεί στην Καρδιολογική Κλινική του ΠαΓΝΗ από Παιδιάτρους για περεταίρω ανίχνευση των φυσημάτων τους ή παρακολουθούνται τακτικά από την ίδια Κλινική λόγω ήδη διαγνωσμένων συγγενών καρδιοπαθειών. Πέραν της καταγραφής, έγινε συνολική εξέταση των παιδιών, εκτίμηση της κλινικής τους εικόνας και πλήρης διάγνωση που κοινοποιήθηκε στους γονείς τους προς ενημέρωση. Για λόγους διαφάλισης του Ιατρικού απορρήτου, όλες οι καταγραφές φέρουν συγκεκριμένο αριθμό για την κατάλληλη χρήση τους και φυσικά είναι ανώνυμες.

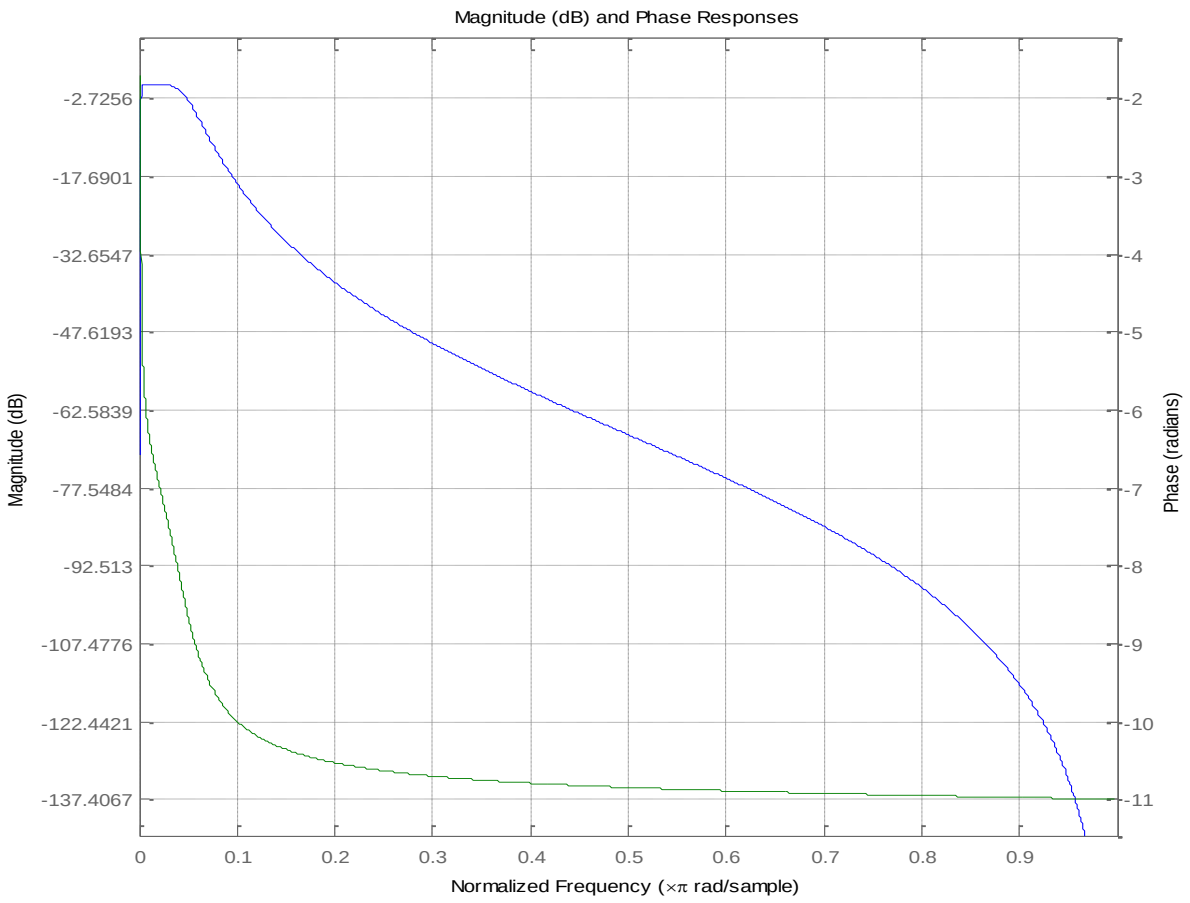
Έγιναν καταγαφές για κάθε ασθενή από διαφορετικά σημεία ακρόασης, έτσι ώστε να ανιχνευτούν όλα τα σημεία της καρδιάς. Διευκρινίζεται εδώ και πάλι, ότι οι θέσεις ακρόασης είναι η κάτω παραστερνική, ή δεξιά και αριστερή άνω παραστερνική, ο σφαγιτιδικός βόθρος και η κορυφαία. Οι καταγραφές έγιναν με συχνότητα δειγματοληψίας των 44,1 kHz. Δεδομένα των οποίων η ποιότητα καταγραφής δεν ήταν ικανοποιητική, λόγω υπερβολικού θορύβου για παράδειγμα, εξαιρέθηκαν από την διαδικασία. Με την ολοκλήρωση των καταγραφών, ο Παιδοκαρδιολόγος Ι. Γερμανάκης εκτίμησε τα ηχητικά χαρακτηριστικά – ευρήματά τους, χρησιμοποιώντας ειδικό λογισμικό απεικόνισης και αναπαραγωγής του PCG. Για κάθε καταγραφή έγινε ανάλυση των ψηφιακών φωνοκαρδιογραφημάτων, για τα οποία καταγράφηκε η ύπαρξη ή μη φυσήματος, ο λεπτομερής χαρακτηρισμός του φυσήματος με

βάση τον τύπο του, με βάση τον χρονισμό του (συστολικό, διαστολικό, συνεχές), το είδος του (αθώο, εξώθησης, κτλπ.), την ένταση του (σε εξαβάθμια κλίμακα), την θέση ακρόασης στην οποία μεγιστοποιείται η έντασή του, καθώς επίσης και την παρουσία τυχόν επιπρόσθετων ήχων όπως για παράδειγμα επιπρόσθετων τόνων, κλαγγές κτλ. Επιλέχθηκαν συνολικά 114 καταγραφές από 38 παιδιατρικούς ασθενείς, από τους οποίους οι 19 παρουσίαζαν φυσιολογικά φυσήματα και οι υπόλοιποι 19 παθολογικά.

3.3 Προεπεξεργασία και Κατάτμηση των Δεδομένων

3.3.1 Προεπεξεργασία των Δεδομένων

Η προεπεξεργασία των σημάτων που περιγράφεται σε αυτήν την υποενότητα, υλοποιήθηκε από τον κώδικα του script HSP.m, που αναπτύχθηκε στα πλαίσια της εργασίας αυτής. Αρχικά, με χρήση της συνάρτησης wavread του εργαλείου MATLAB, ανακτήθηκαν επαναληπτικά οι κυματομορφές μία προς μία. Στη συνέχεια, απαραίτητη ήταν η διαδικασία του φιλτραρίσματος τους. Αυτό γιατί όπως προαναφέρθηκε στην αναφορά αυτή, εμπλέκονται διάφορα σήματα στις καταγραφές. Τέτοια σήματα μπορεί να είναι άλλα βιολογικά σήματα, όπως η αναπνοή, στομαχικοί και εντερικοί θόρυβοι, κλάμα και δυσαρέσκεια ασθενούς, εξωτερικός θόρυβος, καθώς επίσης και ο θόρυβος που εν γένει εισάγεται από το ψηφιακό στηθοσκόπιο. Δεν πρόκειται για την ακρόαση με κάποιο συμβατικό καρδιολογικό στηθοσκόπιο, το καρδιακό σήμα διέρχεται μέσα από μια ηλεκτρονική διάταξη ηχογράφησης, ενίσχυσης και καταγραφής κι έτσι την παρουσία τους κάνουν και τα κοινά σφάλματα των ψηφιακών συστημάτων, ο θόρυβος, τα σφάλματα κβάντισης και άλλα. Αναλύοντας το φάσμα των κυματομορφών και γνωρίζοντας και από άλλες μελέτες στο αντικείμενο τις συχνότητες αυτών των εμπλεκόμενων σημάτων και θορύβων, καταλήξαμε στην εφαρμογή ενός ζωνοπερατού (Band-Pass) Butterworth φίλτρου με ζώνη συχνότητων διέλευσης 40 – 1100 Hz. Στο επίπεδο της υλοποίησης σχεδιάστηκε και κατόπιν εφαρμόστηκε ένα $3^{ης}$ τάξης Band-Pass Butterworth φίλτρο με τις συχνότητες αποκοπής, που προαναφέρονται με χρήση των συναρτήσεων butter και filtfilt του MATLAB αντίστοιχα.



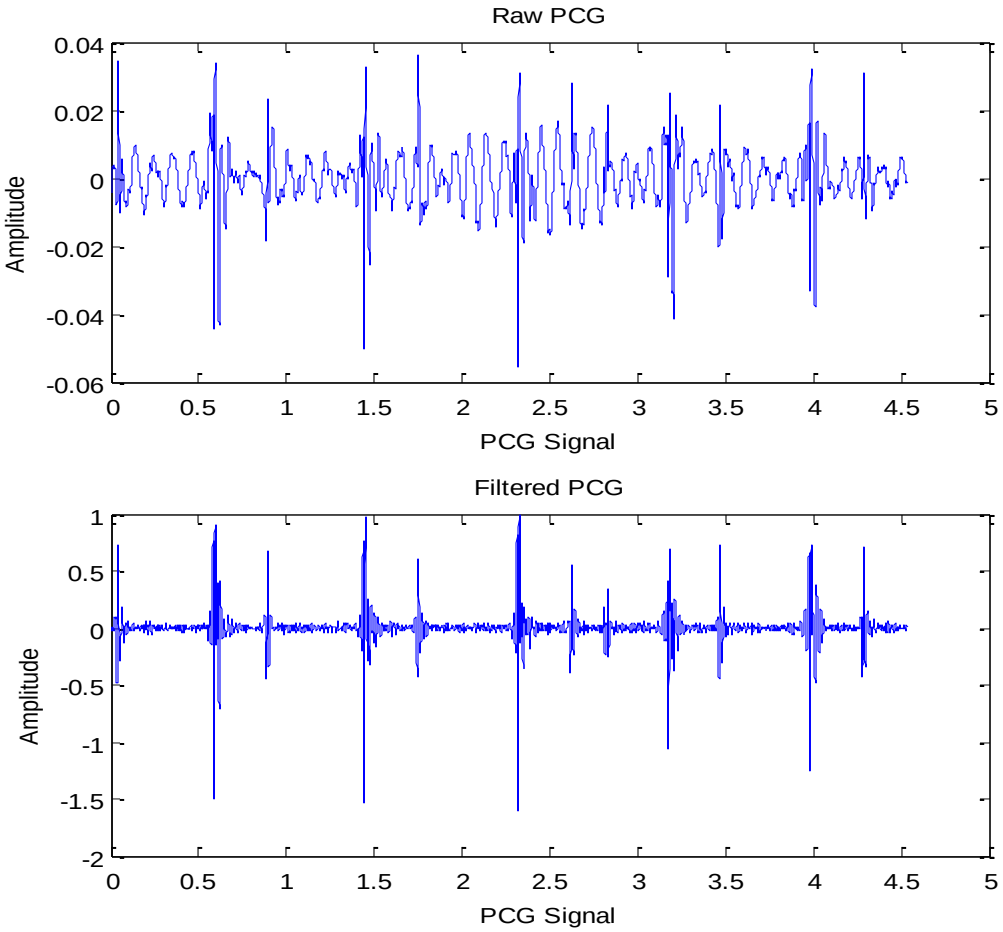
Σχήμα 3.2 – Απόκριση πλάτους (μπλε) και φάσης (πράσινο) του φίλτρου

Ως αποτέλεσμα του φιλτραρίσματος, προφανώς στοιχεία συχνότητας μέχρι και 1100 Hz επιτρέπεται να διαβαίνουν το φίλτρο κι έτσι σύμφωνα με τον τύπο του Nyquist, θα πρέπει τα δείγματα μας να έχουν συχνότητα δειγματοληψίας στα 2200 Hz.

$$f_{Nyquist} = 2 \cdot f_{max}$$, όπου εδώ: $f_{max} = f_{high} = 1100\ Hz$

Έγινε υποδειγματοληψία των δειγμάτων στη συχνότητα 2205 Hz, για λόγους ευκολίας χρήσης του αντίστοιχου εργαλείου της MATLAB, από την αρχική των 44100 Hz με χρήση της συνάρτησης downsample. Στο τέλος, και πριν εισέλθουμε στο στάδιο υπολογισμού της ενέργειας και κατάτμησης, έγινε κανονικοποίηση των σηματων προκειμένου να προσομειώνουν έτσι την σχετική ένταση που αντιλαμβάνεται το ανθρώπινο αυτί.

Παρακάτω, ακολουθεί ένα γράφημα με το παράδειγμα επεξεργασίας μιας καταγραφής από τον αλγόριθμό μας, όπου κανείς μπορεί να δει εποπτικά τα αποτελέσματά της, σχετικά με την σαφή ποιοτική διαφορά που προκύπτει.



Σχήμα 3.3 –Το αρχικό και το επεξεργασμένο PCG σήμα

3.3.2 Κατάτμηση σε Κύκλους και Τμήματα

Για την ορθή χρήση των δεδομένων κάτα την εκπαίδευση και έλεγχο των μοντέλων ταξινόμησης που αναπτύχθηκαν στα πλαίσια της εργασίας, κρίθηκε απαραίτητη η διαδικασία της κατάτμησης των κυματομορφών. Οι πρώτοι 3 καρδιακοί κύκλοι κάθε ηχογράφησης διασπάστηκαν, στα επιμέρους δομικά τμήματά τους, δηλαδή σε τόνους και συστολές-διαστολές. Για την εκτέλεση αυτής της λειτουργίας, αναπτύχθηκε στα πλαίσια της εργασίας

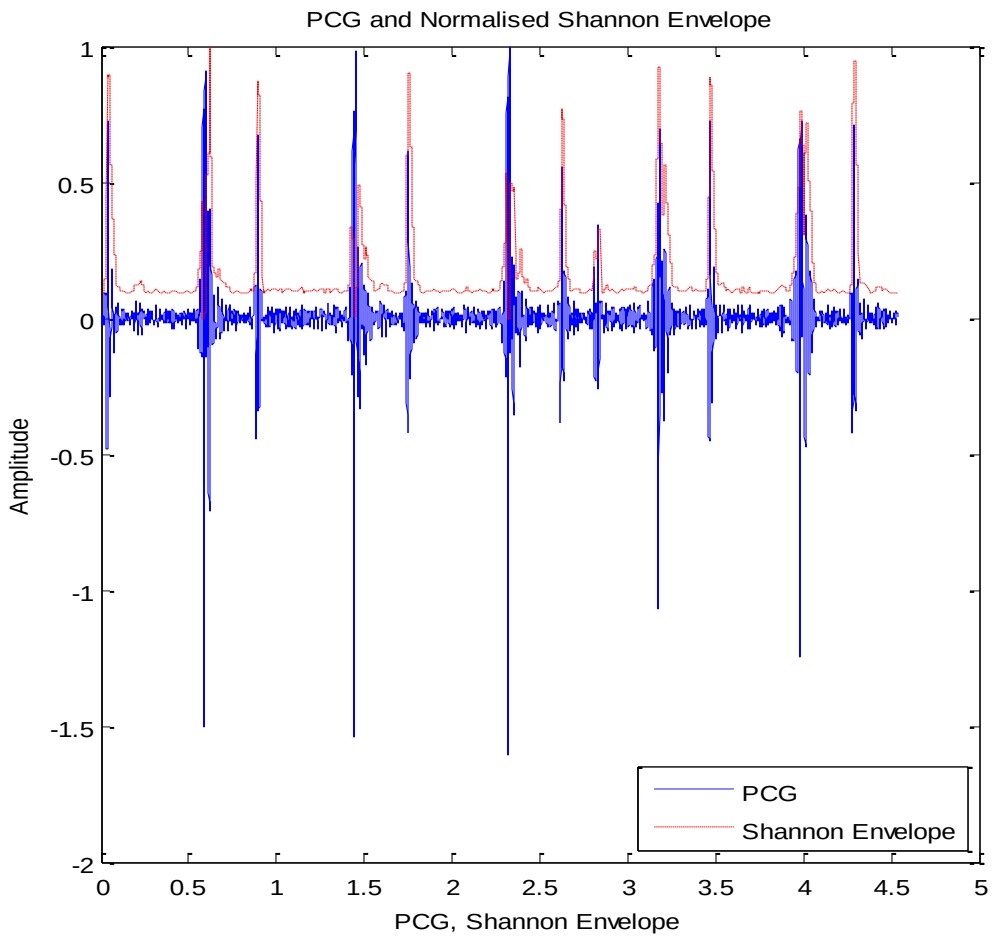
αυτής ειδικός αλγόριθμος κατάτμησης των κυματομορφών. Ο αλγόριθμος αυτός βασίστηκε στην εξαγωγή της περιβάλλουσας ενέργειας Shannon του PCG (Shannon Envelope). Ο υπολογισμός της ενέργειας Shannon, έγινε σε συνεχή παράθυρα (frames) των 0.02 sec, με επικάλυψη 50%, δηλαδή 0.01 sec, σύμφωνα με την σχέση 3.1 :

$$E_s = -\frac{1}{N}\sum_{i=1}^{N} x(i)^2 \log x(i)^2$$

Όπου N ο αριθμός των παραθύρων, x το κανονικοποιημένο PCG σήμα και E_s η κανονικοποιημένη ενέργεια Shannon. Στη συνέχεια, ο τελικός υπολογισμός της περιβάλουσας Shannon (Shannon Envelope) γίνεται ως εξής (σχέση 3.2):

$$P(t) = \frac{E_s(t) - M(E_s(t))}{S(E_s(t))}$$

Όπου $E_s(t)$ η μέση ενέργεια Shannon για το παράθυρο t, $M(E_s(t))$ η μέση τιμή των $E_s(t)$ και $S\big(E_s(t)\big)$ η τυπική απόκλιση των $E_s(t)$. Κι έτσι έχουμε υπολογίσει την κανονικοποιημένη μέση ενέργεια Shannon, $P(t)$, που αλλιώς καλείται και περιβάλουσα Shannon. Στο γράφημα 3.4, φαίνεται το ψηφιακό φωνοκαρδιογράφημα (PCG) και με κόκκινη διακεκομμένη η περιβάλουσα του. Όπως φαίνεται και στο σχήμα, κάθε καρδιακός τόνος του PCG αντιστοιχεί σε τοπικό μέγιστο της γραφικής παράστασης της περιβάλουσας του. Από αυτή την παρατήρηση, μπορεί κανείς να υποθέσει ότι την χρονική στιγμή που κορυφώνεται τοπικά (τοπικό μέγιστο), η περιβάλουσα σηματοδοτεί την ύπαρξη καρδιακού τόνου στο PCG. Για την ακρίβεια, *στην γειτονιά της χρονικής αυτής στιγμής* οριοθετείται ο αντίστοιχος τόνος. Από την στιγμή που μια τέτοια ακμή εντοπίζεται παράλληλα εντοπίζεται και ο χρόνος που αυτή συμβαίνει, μπορούμε αξιοποιώντας ιδιότητες των χρονισμών της καρδιάς, να εντοπίσουμε τα επιμέρους τμήματα του καρδιακού κύκλου. Πιο συγκεκριμμένα, γνωρίζουμε ότι η συστολή είναι βραχύτερη της διαστολής, καθώς επίσης και ότι η συστολή έπεται του τόνου S1 ενώ η διαστολή του τόνου S2. Έτσι λοιπόν, ένα τοπικό μέγιστο της περιβάλουσας θα αντιστοιχεί στον τόνο S1, αν το χρονικό διάστημα που μεσολαβεί μεταξύ του τρέχοντος τοπικού μεγίστου και του επόμενου είναι μικρότερο από αυτό των δυο επόμενων τοπικών μεγίστων, ενώ θα αστιστοιχεί στον S2, αν συμβαίνει το αντίθετο. Προφανώς, μετά τον εντοπισμό και κατάλληλο χαρακτηρισμό των τόνων που αντιστοιχούν στα τοπικά μέγιστα της περιβάλουσας, είναι αρκετά προφανές ότι έχουμε ήδη βρει ταυτόχρονα και τις περιοχές του PCG που αντιστοιχούν στις συστολές και διαστολές που μεσολαβούν.



Σχήμα 3.4 – Ψηφιακό PCG και η περιβάλουσα του (ανηψωμένη ελάχιστα για λόγους απεικόνισης)

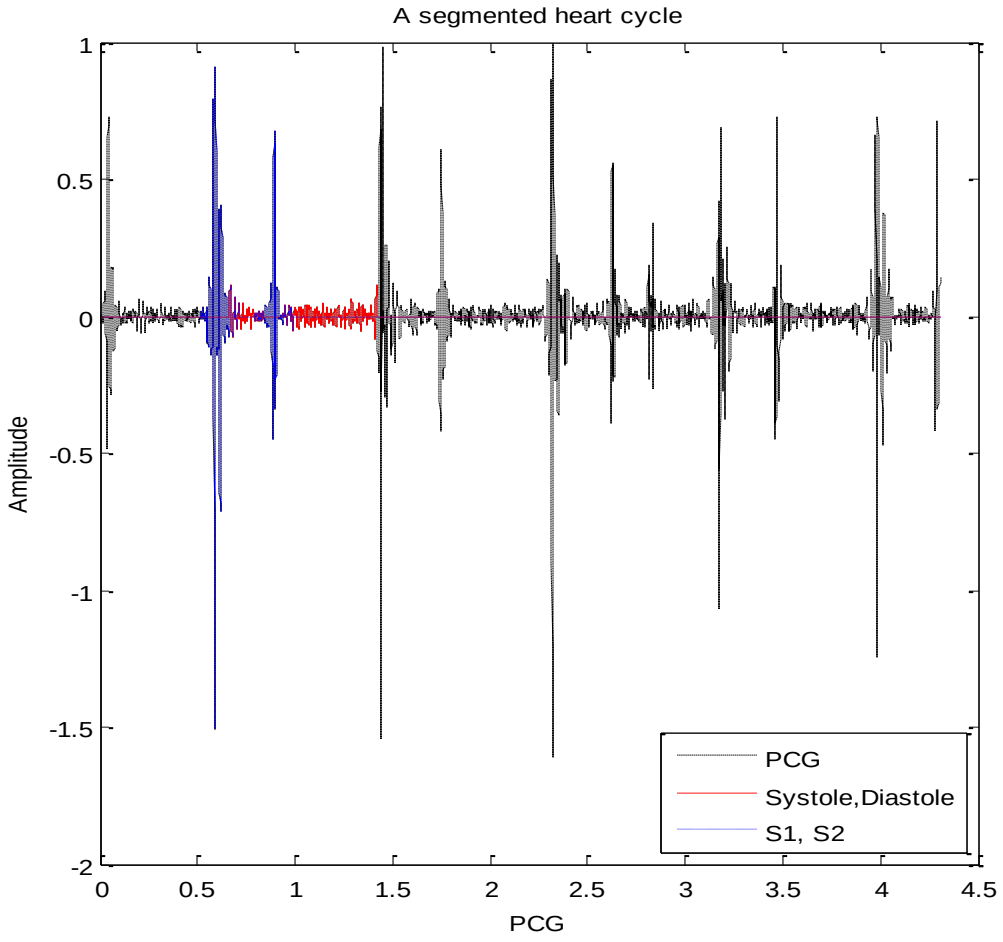
Ωστόσο, από κάθε χρονική στιγμή μέχρι και το πέρας της καταγραφής, θα υπάρχουν περισσότερα από ένα τοπικά μέγιστα της περιβάλουσας. Αυτό βεβαίως αποτελεί πρόβλημα, γιατί δεν θα επιθυμούσαμε να χάσουμε κάποιο κύκλο από την καταγραφή. Έτσι λοιπόν, η αναζητήση του τοπικού μεγίστου στην περιβάλουσα, θα έπρεπε να διεξάγεται εντός ενός αυστηρού χρονικά περιθωρίου. Βέβαια, αυτό το χρονικό περιθώριο δεν μπορεί να αποτελεί μια μεμονωμένη αριθμητική επιλογή, καθότι κάθε καταγραφή κι όχι μόνο κάθε ασθενής διαφέρει από τις υπόλοιπες (λόγω μεταβολής του καρδιακού ρυθμού για παράδειγμα).

Παρόλα ταύτα από σχετικές μελέτες γνωρίζουμε ότι αν Y είναι η διάρκεια της συστολής (σε ms) και X ο καρδιακός ρυθμός (σε παλμούς ανα λεπτό) τότε θα ισχύει η σχέση 3.3:

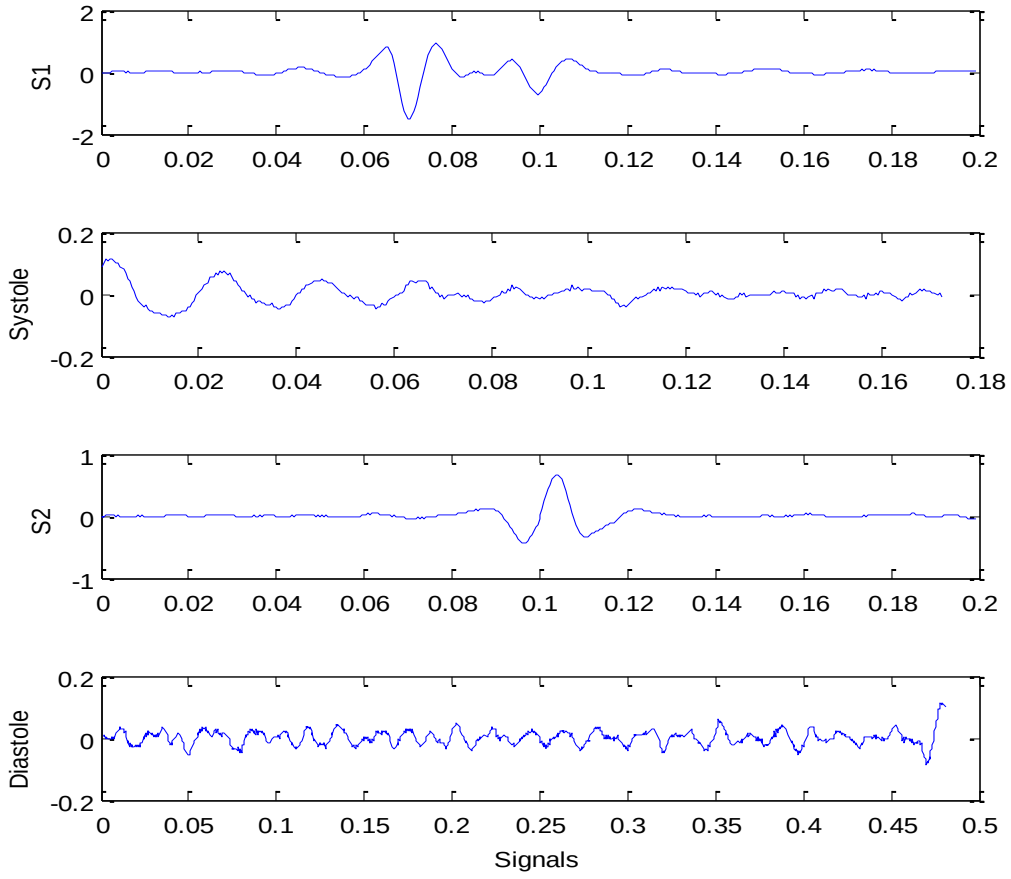
$$Y = 0{,}351 - 0{,}001X$$

Οπότε, γνωρίζοντας προσεγγιστικά την διάρκεια της συστολής σύμφωνα με την παραπάνω σχέση, καθώς επίσης και την διάρκεια της διαστολής, αφού γνωρίζουμε τον καρδιακό ρυθμό, τον αριθμό των κύκλων και την διάρκεια της συστολής, κι έτσι μπορούμε να οριοθετήσουμε χρονικά το χρονικό κομμάτι στο οποίο θα αναζητήσουμε το τοπικό μέγιστο της περιβάλουσας.

Με βάση όλα τα παραπάνω, σχεδιάστηκε και υλοποιήθηκε ο αλγόριθμος κατάτμησης. Στο παρακάτω σχήμα φαίνεται το PCG, καθώς επίσης με ειδική σήμανση τα επιμέρους τμήματα του πρώτου ολόκληρου καρδιακού κύκλου, όπως ο αλγόριθμος υπολόγισε.

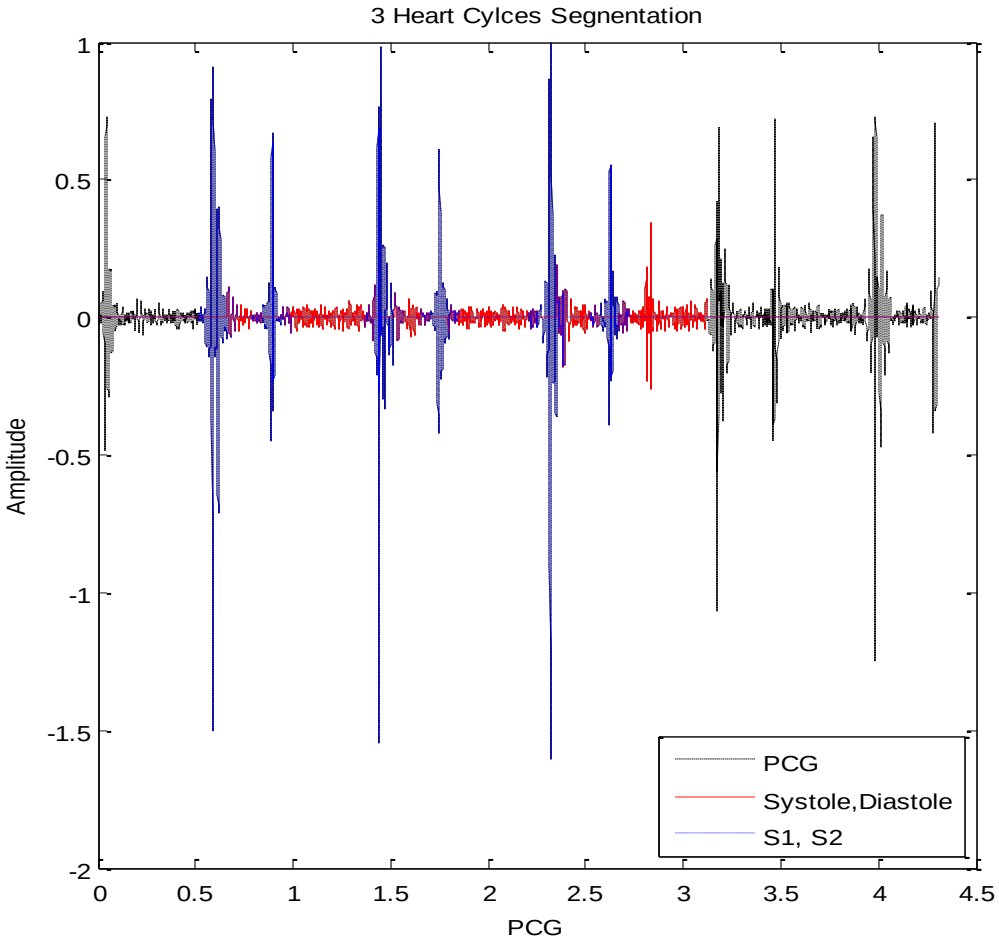


Σχήμα 3.5 – Ο πρώτος ολόκληρος καρδιακός κύκλος κατατμημένος στα επιμέρους τμήματά του



Σχήμα 3.6 – Τα επιμέρους τμήματα του καρδιακού κύκλου εξηγμένα από το PCG

Ο αλγόριθμος αυτός γενικεύτηκε, έτσι ώστε να εξάγει τα τμήματα των τριών πρώτων κύκλων της καταγραφής. Δεν θα ήταν Ιατρικά ορθό να ασχοληθούμε αποκλειστικά με τον πρώτο κύκλο, καθώς αυτός θα μπορούσε να παρουσιάζει μια «φαινομενική ανωμαλία»,, η οποία να οφείλεται σε φυσιολογικά αίτια, όπως αναπνοή, ή σε εξωγενείς και τυχαίους παράγοντες, όπως ξαφνικός θόρυβος. Έτσι το σύστημα μας γίνεται πιο ανεκτικό σε τέτοιου είδους συμβάντα, με το να λαμβανει υπ'όψιν του όχι μόνο τα στοιχεία ενός, αλλά τριών διαφορετικών καρδιακών κύκλων. Παρακάτω φαίνεται η γενίκευση αυτή στην ίδια καταγραφή.



Σχήμα 3.7 – Οι πρώτοι τρεις κατατμημένοι καρδιακοί κύκλοι της καταγραφής

Εποπτικά, φαίνεται μάλιστα η ανάγκη της γενίκευσης του αλγορίθμου στην κατάτμηση των τριών καρδιακών κύκλων, καθώς αν για παράδειγμα λαμβάναμε υπ'όψιν μας τον τρίτο ολόκληρο καρδιακό κύκλο, ίσως να τον χαρακτηρίζαμε ως παθογενή μιας και ο θόρυβος κατά τη διάρκεια της συστολής του, θα μπορούσε να θεωρηθεί συστολικό φύσημα. Κάτι που προφανώς δεν ισχύει στην περίπτωση αυτή, δεν αποτελεί δηλαδή τίποτα παραπάνω από ένα τυχαίο συμβάν με πιθανότερο αίτιο τον θόρυβο.

Ο αλγόριθμος κατάτμησης είχε ποσοστό επιτυχίας **91.9%**. Το 8.1% του ποσοστιαίου σφάλματος, οφείλεται κυρίως στο θόρυβο (εξωτερικό, εμπλεκόμενα σήματα), σφάλματα του οργάνου λήψης των καταγραφών και αστοχίες του ίδιου του αλγορίθμου κατάτμησης.

3.4 Εξαγωγή των Χαρακτηριστικών

Στα συστήματα ταξινόμησης, όπως φαίνεται άλλωστε και στο σχήμα 1.1, το στάδιο της προεπεξεργασίας και κατάτμησης ακολουθείται από αυτό της εξαγωγής χαρακτηριστικών (feature extraction).

Στα πλαίσια της εργασίας αυτής, εξάχθηκαν από τα δείγματα χαρακτηριστικά όπως Mel Frequency Cepstral Coefficients (MFCCs) με ή χωρίς εφαρμογή Cepstral Mean Normalization (CMN), Melspec, Filter Banks (FBANK) και τέλος Perceptual Linear Prediction (PLPs). Για κάθε είδος χαρακτηριστικών, δοκιμάστηκε επίσης η εξαγωγή των συντελεστών ενέργειας, οι πρώτες και δεύτερες Δ – διαφορές, διάφορα μήκη παραθύρων διαφορών και τα λοιπά. Το στάδιο αυτό της εξαγωγής χαρακτηριστικών έγινε με τη χρήση των εργαλείων του Hidden Markov model Tool Kit (HTK) και πιο συγκεκριμμένα με το εργαλείο HCopy. Η εντολή αυτή, πλην των άλλων, χρησιμοποιείται για την μετατροπή κυματομορφών (waveforms), σε διανύσματα χαρακτηριστικών σύμφωνα με τον ακόλουθο πίνακα συμβατών μετατροπών:

	Outputs										
Inputs	WAVEFORM	LPC	LPREFC	LPCEPSTRA	IREFC	MFCC	FBANK	MELSPEC	USER	DISCRETE	PLP
WAVEFORM	√	√	√	√	√	√	√	√		√	
LPC		√	√	√	√					√	
LPREFC		√	√	√	√					√	
LPCEPSTRA		√	√	√	√					√	
IREFC		√	√	√	√					√	
MFCC						√				√	
FBANK						√	√			√	
MELSPEC						√	√	√		√	
USER									√	√	
DISCRETE										√	
PLP										√	√

Σχήμα 3.8 – Πίνακας συμβατών μετασχηματισμών με χρήση της HCopy

Η γενική σύνταξη του παραπάνω εργαλείου είναι η εξής:

HCopy [options] [input files] [output files],

όπου στο πεδίο [options] προσδιορίζονται οι παράμετροι της επιθυμητής μετατροπής, συνήθως δίδονται σε ένα αρχείο παραμέτρων ή configuration file με το πρόθεμα –C και στα πεδία [input files] και [output files], δίδονται λίστες των αρχείων εισόδου και εξόδου αντίστοιχα, συνήθως υπό την μορφή αρχείων .txt, .scp, .list κ.ά. με το πρόθεμα –S. Πληκτρολογώντας τα παραπάνω σε κάποια τερματικό περιβάλλοντος UNIX, έχοντας εγκαταστήσει επιτυχώς τα αρχεία της εφαρμογής HTK, μας έδωσε σαν αποτέλεσμα τα αντίστοιχα διανύσματα χαρακτηριστικών, σύμφωνα με ότι έχουμε προκαθορίσει στο αντίστοιχο configuration file.

3.4.1 Εξαγωγή των MFCCs με ή χωρίς CMN

Πολύ συχνά στην επεξεργασία και αναγνώριση φωνής, χρησιμοποιούνται οι λεγόμενοι Mel-Frequency Cepstral Coefficients (MFCCs), λόγω κυρίως της πολύ καλής διακριτικότητας που χαρίζουν στο σύστημα, καθώς και του γεγονότος ότι είναι είναι αρκετά εύκολα διαχειρίσιμοι. Πιο συγκεκριμμένα, υπολογίζονται από τα log filterbank πλάτη $\{m_j\}$ με την χρήση του διακριτού μετασχηματισμού συνημιτόνου:

$$c_i = \sqrt{\frac{2}{N}} \cdot \sum_{j=1}^{N} m_j \cdot \cos(\frac{\pi i}{N} \cdot (j - 0.5))$$

όπου N ο αριθμός των filterbank καναλιών και καθορίζεται από την παράμετρο NUMCHANS, ο αριθμός των απαραίτητων συντελεστών ανά διάνυσμα καθορίζεται από την παράμετρο NUMCEPS, ενώ μάλιστα δυνατή είναι η ανέλκυση ή liftering σύμφωνα με την παράμετρο CEPLIFTER. Η παραγωγή τέτοιων χαρακτηριστικών, γίνεται με τον ορισμό της παραμέτρου TARGETKIND σε MFCC.

Σε ορισμένες περιπτώσεις, απαραίτητη κρίνεται η αφαίρεση της επίδρασης της συνάρτησης μεταφοράς του καναλιού, κάτι που επιτυγχάνεται κατόπιν του υπολογισμού της μέσης τιμής της συγκεκριμμένης παραμέτρου χαρακτηριστικών του αρχείου. Πρακτικά, αρκεί να προστεθεί στην παράμετρο TARGETKIND το επίθεμα _Z, δηλαδή αν το αρχικό ήταν MFCC και θέλαμε να εφαρμόσουμε CMN θα τροποποιούσαμε την τιμή της παραμέτρου σε MFCC_Z.

Υπάρχουν πολλές ακόμα επιλογές, όπως η λήψη διαφορών πρώτης, δευτέρας κτλ. τάξης, η λήψη ενέργειας, η προσθήκη/μή προσθήκη του συντελεστή c_0 κ.ά. που θα σχολιαστούν παρακάτω.

Παρακάτω παρατίθεται το αρχείο παραμέτρων για την παραγωγή τέτοιων χαρακτηριστικών:

```
SOURCEKIND = WAVEFORM
SOURCEFORMAT = WAV
TARGETKIND = MFCC_0
TARGETRATE = 100000.0
WINDOWSIZE = 4000000.0
USEHAMMING = T
PREEMCOEF = 0.97
NUMCHANS = 24
CEPLIFTER = 22
NUMCEPS = 12
ENORMALISE = FALSE
LOFREQ = 40
HIFREQ = 1100
```

Με την τροποποίηση TARGETKIND = MFCC_0_Z εκτελείται και η διαδικασία CMN.

3.4.2 Εξαγωγή των PLPs

Εναλλακτικά στους MFCCs χρησιμοποιούνται οι PLPs. Για τον υπολογισμό τους, όπως αυτός υλοποιείται στο HTK χρησιμοποιείται η mel-frequency filterbank. Οι mel-filterbank συντελεστές, ισοζυγίζονται πάνω σε μια καμπύλη και στη συνέχεια συμπιέζονται με χρήση της κυβικής ρίζας. Από το ηχητικό spectrum που προκύπτει, υπολογίζονται οι Linear Prediction (LP) συντελεστές οι οποίοι μετατρέπονται σε cepstral coefficients. Ακολουθούν ορισμένοι παράμετροι απαραίτητοι για την εξαγωγή των PLPs.

```
TARGETKIND = PLP_E_D
DELTAWINDOW = 2
TARGETRATE = 100000.0
WINDOWSIZE = 4000000.0
COMPRESSFACT = 0.33
```

Μάλιστα, είναι δυνατό να προσαρμοστεί η τιμή της παραμέτρου COMPRESSFACT σε οποιαδήποτε επιθυμητή τιμή.

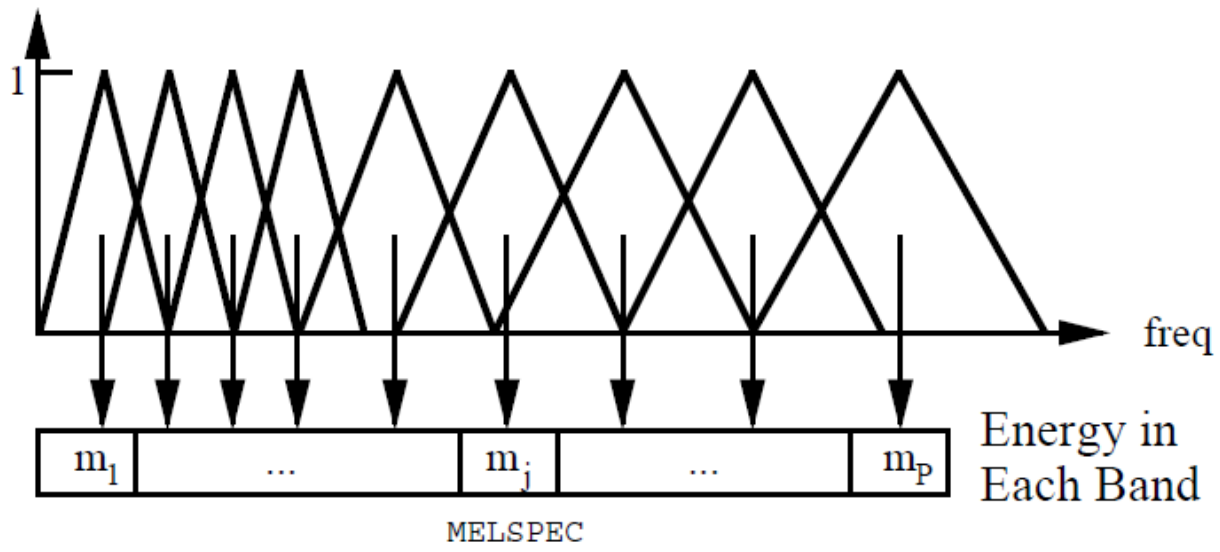
3.4.3 Εξαγωγή των MELSPEC και FBANK

Σύμφωνα με την Φυσιολογία, το ανθρώπινο αυτί αντιλαμβάνεται μή – γραμμικά τις συχνότητες του ηχητικού φάσματος και ως εκ τούτου υπάρχει εμπειρική θεώρηση σύμφωνα με την οποία, αν ένα σύστημα σχεδιαστεί έτσι ώστε το front – end του να «αντιλαμβάνεται» με

τον ίδιο μή – γραμμικό τρόπο τις συχνότητες, θα έχει και βελτιωμένη απόδοση. Εναλλακτική λοιπόν μέθοδος, είναι αυτή της Filterbank Ανάλυσης, καθώς είναι μια μέθοδος που προσφέρει έναν άμεσο τρόπο ανάκτησης τέτοιας μή – γραμμικής συχνότητας. Το HTK παρέχει έναν απλό τρόπο υπολογισμού του μετασχηματισμού Fourier. Η mel – scale προσδιορίζεται σύμφωνα με τη σχέση:

$$Mel(f) = 2595 \log_{10}(1 + \frac{f}{700})$$

ενώ τα φίλτρα είναι τριγωνικά όπως φαίνεται και παρακάτω:



Σχήμα 3.9 – Mel – Scale Filter Bank

Πρακτικά, κάθε συντελεστής του πλάτους (magnitude) μετασχηματισμού Fourier, πολλαπλασιάζεται με το αντίστοιχο κέρδος φίλτρου και τελικά παίρνουμε ένα ζυγισμένο άθροισμα που αναπαριστά το φασματικό πλάτος για το συγκεκριμμένο filterbank κανάλι. Θέτοντας την παράμετρο USEPOWER στην τιμή TRUE χρησιμοποιείται η ισχύς (power) αντί του πλάτους (magnitude).

Χρησιμοποιώντας τις παραμέτρους LOFREQ και HIFREQ, μπορούμε στην περίπτωση της ανάλυσης αυτής, να προσδιορίσουμε τις επιθυμητές συχνότητες αποκοπής έτσι ώστε να αποφύγουμε ανεπιθύμητες συχνότητες, περιοχές συχνοτήτων όπου το σήμα δεν έχει ουσιαστική ενέργεια κτλπ. Διαφορετικά, η επιτρεπτή ζώνη συχνοτήτων είναι μεταξύ του 0 και της συχνότητας Nyquist.

Τέλος, σε ότι αφορά την παράμετρο TARGETKIND, θέτοντας την “MELSPEC” παράγονται απευθείας οι συντελεστές όπως παρουσιάστηκε πιο πάνω, ενώ θέτοντας την “FBANK” οι συντελεστές μετατρέπονται σε λογαριθμική κλίμακα (log – scale). Παρακάτω, φαίνεται το σύνολο των παραμέτρων που προσδιορίζονται για την παραγωγή τέτοιων χαρακτηριστικών:

SOURCEKIND = WAVEFORM
SOURCEFORMAT = WAV
TARGETKIND = FBANK_D_A
DELTAWINDOW = 14
ACCWINDOW = 14
TARGETRATE = 100000.0
WINDOWSIZE = 2000000.0
ZMEANSOURCE = T
USEHAMMING = T
PREEMCOEF = 0.97
NUMCHANS = 24
CEPLIFTER = 22
NUMCEPS = 12
ENORMALISE = FALSE
LOFREQ = 40
HIFREQ = 1100
NATURALREADORDER = TRUE
NATURALWRITEORDER = FALSE

3.4.4 Μετρικές Ενέργειας, Διαφορών και λοιπές Παράμετροι

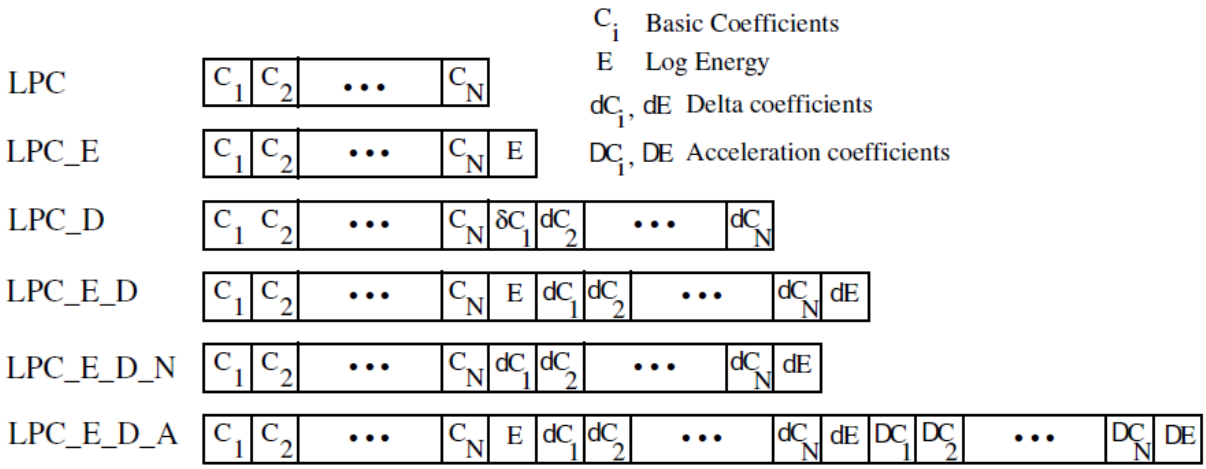
Υπάρχει η δυνατότητα να ανακτηθεί και ο συντελεστής της ενέργειας, χρησιμοποιώντας το επίθεμα _E στην παράμετρο TARGETKIND, η οποία υπολογίζεται ως ο λογάριθμος της ενέργειας σήματος με βάση την σχέση:

$$E = \log \sum_{n=1}^{N} {s_n}^2$$

Με το επίθεμα _D και παράλληλο προσδιορισμό της παραμέτρου DELTAWINDOW εφαρμόστηκε στα εκάστοτε χαρακτηριστικά η τεχνική των πρώτων διαφορών. Τέτοιες χρονικές παράγωγοι, μπορούν να βλτιώσουν την απόδοση ενός συστήματος όταν εφαρμοστούν πάνω στα στατικά χαρακτηριστικά. Στις πρώτες διαφορές μπορεί αναδρομικά να εφαρμοστεί η ίδια ανάλυση κι έτσι να πάρουμε τις δεύτερες διαφορές, πράγμα που χρησιμοποιήθηκε με την προσθήκη του επιθέματος _A στην παράμετρο TARGETKIND και τον προσδιορισμό της τιμής της παραμέτρου ACCWINDOW. Ομοίως και για τις τρίτες διαφορές, όπου το επίθεμα είναι _T. Ο υπολογισμός των διαφορών γίνεται με βάση τη σχέση:

$$d_t = \frac{\sum_{\theta=1}^{\Theta} \theta(c_t + \theta - c_t - \theta)}{2 \cdot \sum_{\theta=1}^{\Theta} \theta^2}$$

Παρακάτω, φαίνεται πως ακριβώς αποθηκεύονται τα χαρακτηριστικά ανάλογα με τι πληροφορία φέρουν. Αρχικά συντελεστές γραμμικής πρόβλεψης (LPC) και στη συνέχεια με προσθήκη της ενέργειας, $1^{ων}$ διαφορών, ενέργειας και πρώτων διαφορών, απόλυτη ενέργεια και τέλος $2^{ων}$ διαφορών.



Σχήμα 3.10 – Αποθήκευση των χαρακτηριστικών

Τέλος, να αναφερθεί ότι στην διαδικασία εξαγωγής των χαρακτηριστικών, χρησιμοποιήθηκαν και πολλές ακόμα παράμετροι των οποίων μελετήθηκε η επίδρασή τους στο αποτέλεσμα της ταξινόμησης. Αναφέρονται συνοπτικά οι σημαντικότερες παράμετροι:

WINDOWSIZE: Χρησιμοποιήθηκαν πολλές διαφορετικές τιμές για την συγκεκριμμένη παράμετρο, μιας και το μήκος παραθύρου όπου εξάγονται τα χαρακτηριστικά επηρεάζει άμεσα την απόδοση του συστήματος.

ZMEANSOURCE: Δίνει την δυνατότητα να αγνοηθεί η επίδραση του DC σφάλματος.

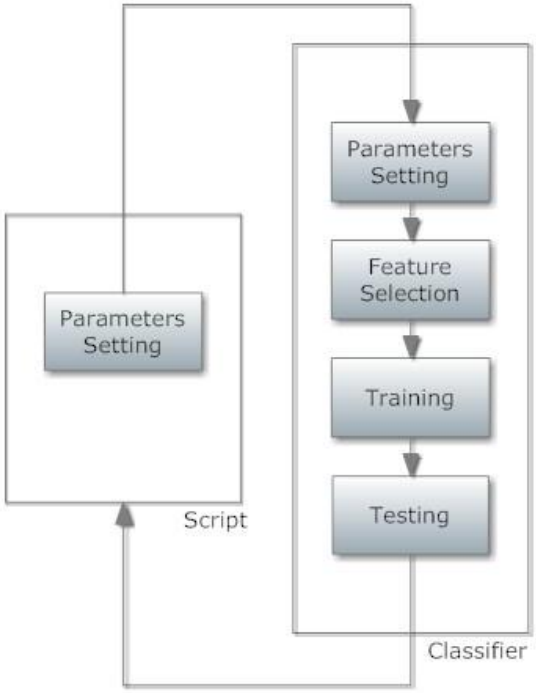
USEHAMMING: Αυτή η παράμετρος δίνει την επιλογή χρήσης παραθύρου Hamming, κατά τη διαδικασία παραθύρωσης των δεδομένων. Ένα τέτοιο παράθυρο, συνήθως βοηθά στην ομαλότερη αποκοπή τμημάτων του σήματος.

ENORMALISE: Πρόκειται για την κανονικοποίηση της ενέργειας.

4 Ανάπτυξη των Ταξινομητών

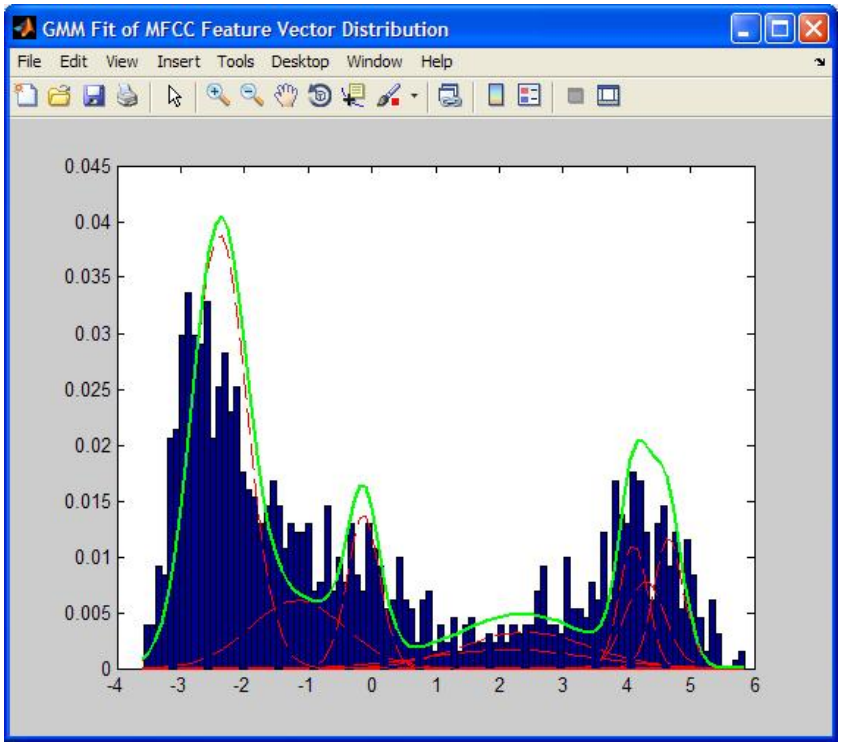
4.1 Ανάπτυξη του GMM ταξινομητή

Στο πρώτο μέρος της παρούσας εργασίας, αναπτύχθηκε και χρησιμοποιήθηκε ένας στατικός ταξινομητής με χρήση GMMs, χρησιμοποιώντας τα εργαλεία του MATLAB. Σκοπός αυτού του πρώτου μέρους των πειραμάτων, είναι η σύγκριση της απόδοσης του ταξινομητή για διαφόρων τύπων χαρακτηριστικά. Για την παραμετροποίηση του συστήματος μας που θα επέτρεπε την εύκολη εναλλαγή μεταξύ των διαφόρων πειραμάτων, ο ταξινομητής αναπτύχθηκε σαν κώδικας εντός μιας συνάρτησης (function) σε κώδικα MATLAB και καλούνταν από ένα script που διοχέτευε τις κατάλληλες παραμέτρους για το κάθε πείραμα. Η δομή του ταξινομητή είναι ιδιαίτερα απλή και αποτελείται από τα μέρη του ορισμού παραμέτρων, της εκπαίδευσης των μοντέλων (training) και της αξιολόγησης της επίδοσης (testing). Να αναφερθεί εδώ, ότι επιλέχθηκαν συγκεκριμμένα χαρακτηριστικά βάσει του χρόνου, αντιπροσωπεύοντας έτσι συγκεκριμμένο τμήμα του καρδιακού σήματος, για παράδειγμα συστολές μόνο και έγιναν και συγκρίσεις μεταξύ των διαφόρων επιλογών αυτών. Περισσότερα για την πειραματική διαδικασία θα αναφερθούν στο επόμενο κεφάλαιο.



Σχήμα 4.1 – Η δομή της ανάπτυξης του GMM ταξινομητή

Το πρώτο βήμα για το σχεδιασμό ενός ταξινομητή, είναι φυσικά η μελέτη των χαρακτηριστικών και οι στατιστικές τους ιδιότητες. Κάτι τέτοιο διευκολύνει την επιλογή του μοντέλου που ταιριάζει στα χαρακτηριστικά και αποτελεί φυσικά το κύριο δομικό μέρος του ταξινομητή. Γενικά η διαδικασία της ανάλυσης των δεδομένων και της επιλογής κατάλληλου μοντέλου, είναι αρκετά δύσκολη και απαιτεί εμπειρία για να είναι αποτελεσματική, καθώς δεν βασίζεται τόσο σε συγκεκριμμένους κανόνες και τεχνικές, αλλά περισσότερο στη διαισθητική προσέγγιση. Παρόλα αυτά, τα τελευταία χρόνια έχουν αναπτυχθεί και χρησιμοποιούνται ευρέως στατιστικά εργαλεία και εφαρμογές εξόρυξης δεδομένων (Data Mining), που με κατάλληλες εντολές και διαδικασίες διευκολύνουν κατά πολύ την δύσκολη αυτή απόφαση σχεδιασμού. Στα πλαίσια της παρούσας εργασίας, χρησιμοποιήθηκαν κυρίως τα εργαλεία του στατιστικού πακέτου του MATLAB (MATLAB Statistical Toolbox), καθώς και τα εργαλεία οπτικοποίησης που παρέχει ο Explorer του WEKA. Πιο συγκεκριμμένα, με το πέρας της εξαγωγής των χαρακτηριστικών από τις κυματομορφές, είχαμε στη διάθεση μας όλα τα δεδομένα του πειράματος, «στοιβαγμένα» κατάλληλα σε στήλες διανυσμάτων χαρακτηριστικών. Με απλές εντολές στη κονσόλα του MATLAB, λήφθηκαν συγκεκριμμένες στήλες των διανυσμάτων και απεικονίστηκαν τα αντίστοιχα ιστογράμματα. Με την εντολή hist και με το εργαλείο gmdistribution.fit του στατιστικού πακέτου του MATLAB, εμφανίζεται το συνταίριασμα των ιστογραμμάτων των χαρακτηριστικών με τις γκαουσσιανές κατανομές του μίγματος.



Σχήμα 4.1 – Τα χαρακτηριστικά και οι κατανομές που τα αντιπροσωπεύουν

Ο αριθμός των γκαουσσιανών κατανομών στο μίγμα ή αλλιώς η τάξη μίγματος, είναι μια παράμετρος που μπορεί κανείς ελεύθερα να επιλέξει και έπειτα να τροποποιήσει μέχρι να πετύχει το καλύτερο συνταίριασμα. Στην παραπάνω εικόνα, επιλέχθηκε το M=8 ως η τάξη μίγματος. Με τις παρακάτω εντολές πετυχαίνουμε το λεγόμενο "fit" που περιγράφηκε:

```
M = 8 ;
model = gmdistribution.fit(MFCCdata,M) ;
```

όπου MFCCdata το σύνολο των χαρακτηριστικών και M η τάξη του μίγματος. Με αυτή την απλή τεχνική, έγινε η επιλογή του μοντέλου, ξεκινώντας δηλαδή από την παρατήρηση των ίδιων μας των χαρακτηριστικών. Η απλή παρατήρηση και τελικά διερεύνηση για το αν τα δεδομένα μοντελοποιούνται κατάλληλα από ένα μοντέλο τύπου GMM, έγινε καθώς τα ιστογράμματα έδειχναν να ακολουθούν κανονική ή γκαουσσιανή κατανομή, με το χαρακτηριστικό κωδωνοειδές σχήμα κατά τόπους, με άλλα λόγια, διάφορα υποσύνολα των χαρακτηριστικών ακολουθούσαν τέτοια κατανομή, με αποτέλεσμα το σύνολό τους να μοντελοποιείται με ένα σύνολο γκαουσσιανών κατανομών με τα αντίστοιχα βάρη, τα οποία είναι ανάλογα με την πληθυκότητα των υποσυνόλων. Με την αποσαφήνιση πια του μοντέλου που θα χρησιμοποιηθεί, η φάση της ανάπτυξης του κώδικα είναι το αμέσως επόμενο βήμα.

Στο σώμα του κώδικα του ταξινομητή και στο πρώτο τμήμα του, έγινε ο καθορισμός όλων των πειραματικών παραμέτρων, όπως η μέθοδος πειραματισμού, ενδεικτικά 5 – cross fold validation, αρχικές τάξεις μιγμάτων των μοντέλων, a-priori πιθανότητες κλάσεων σε λογαριθμική κλίμακα και άλλα. Σε αυτό το στάδιο της ανάπτυξης, δόθηκε έμφαση ως προς την παραμετροποίηση του συστήματος, πράγμα που το κατέστησε ευέλικτο κι έτσι υπήρχε η δυνατότητα να γίνουν όσο το δυνατό περισσότερα πειράματα χωρίς αλλαγές στον ταξινομητή, αλλά μόνο με αλλαγές στις τιμές των παραμέτρων.

Στο δεύτερο τμήμα, έγινε όπως προαναφέρθηκε, μια επιλογή χαρακτηριστικών βάσει του τμήματος του καρδιακού σήματος στο οποίο βρίσκονταν. Το στάδιο της κατάτμησης του καρδιακού σήματος, που αναλύθηκε διεξοδικά στο 3^{ο} κεφάλαιο, μας επέτρεψε να αποθηκεύσουμε σε μια δομή όλα τα σημαντικά χρονικά χαρακτηριστικά του σήματος (time stamps), όπως το σημείο αρχής και τέλους της συστολής, της διαστολής, των τόνων κ.ο.κ. για κάθε κυματομορφή. Με αυτό τον τρόπο και εφόσον είχαμε ακρίβεια των 10ms κατά την εξαγωγή των χαρακτηριστικών, είχαμε την δυνατότητα στα πλαίσια των δοκιμών, να επιλέγουμε κατάλληλα όποια τμήματα του καρδιακού σήματος ήταν επιθυμητά με την αντίστοιχη επιθυμητή επικάλυψη (overlapping). Οι διάφοροι τρόποι χρονικής επιλογής των χαρακτηριστικών, παρουσιάζονται στο επόμενο κεφάλαιο στις πειραματικές διαδικασίες.

Στη συνέχεια, ακολουθεί η εκπαίδευση των μοντέλων (training). Χρησιμοποιώντας τα δεδομένα εκπαίδευσης τα μοντελοποιούμε κατάλληλα, κατασκευάζοντας ένα μίγμα

γκαουσιανών κατανομών για κάθε μια από τις δύο κατηγορίες. Αυτό σημαίνει πως εκπαιδεύονταν ένα μίγμα γκαουσιανών για τα καρδιακά σήματα που φέρουν φυσιολογικά φυσήματα και ένα για αυτά που φέρουν παθολογικά. Εναλλακτικά, εκπαιδεύτηκαν και χρησιμοποιήθηκαν τέσσερα τέτοια μοντέλα, για φυσιολογικές και παθολογικές συστολές και διαστολές. Η εκπαίδευση των GMMs, έγινε με τη χρήση της εντολής gaussmix του MATLAB. Στην είσοδό της δέχεται πίνακα με τα δεδομένα εκπαίδευσης του μοντέλου, καθώς επίσης και τις αρχικοποιήσεις και επιστρέφει τις παραμέτρους του μοντέλου, δηλαδή τις μέσες τιμές, τις διασπορές και τα βάρη των γκαουσιανών κατανομών, αφού έχει πρώτα εκπαιδεύσει τα μοντέλα με τη χρήση του επαναληπτικού Expectation Maximization (EM) αλγορίθμου.

```
[MU1,Sigma1,W1,g1,f1,pp1,gg1] = gaussmix(trainCat1,[],300,m0,v0,w0) ;
[MU2,Sigma2,W2,g2,f2,pp2,gg2] = gaussmix(trainCat2,[],300,m0,v0,w0) ;
```

Με τις παραπάνω γραμμές κώδικα, επιτυγχάνεται η εκπαίδευση δύο μοντέλων τύπου GMM, ένα για την πρώτη κατηγορία (φυσιολογικά φυσήματα) κι ένα για την δεύτερη κατηγορία (παθολογικά φυσήματα). Για την εκπαίδευση, απαιτούνται πίνακες των δεδομένων, εδώ τα trainCat1 και trainCat2, αρχικές τιμές για τις μέσες τιμές, τις διασπορές και τα βάρη και τέλος τον μέγιστο αριθμό επαναλήψεων για τη σύγκλιση του EM αλγορίθμου, εδώ 300, ενώ δίνεται και η δυνατότητα να οριστεί κάποιο κατώτερο όριο μεταβολής των παραμέτρων. Οι παραπάνω συναρτήσεις επιστρέφουν τις μέσες τιμές, τις διασπορές και τα βάρη των μοντέλων όπως επίσης και τις λογαριθμικές πιθανότητες και τις τιμές fisher.

Τέλος, με παρομοίως απλό τρόπο, έγινε και η αξιολόγηση της ταξινόμησης, χρησιμοποιώντας την συνάρτηση gaussmixp, η οποία δεχόμενη ένα δείγμα ελέγχου (testing sample) επιστρέφει πίνακα με τις αντίστοιχες a-posteriori πιθανότητες σε λογαριθμική κλίμακα το δείγμα να ανήκει στην κατηγορία του μοντέλου. Υλοποιείται έτσι με άλλα λόγια, η διαδικασία της ταυτοποίησης των προτύπων (pattern matching) και στο τέλος κατόπιν κάποιας επεξεργασίας (post processing), όπως για παράδειγμα ο συνυπολογισμός των a-priori πιθανοτήτων και άλλα, λαμβάνεται η απόφαση γύρω από το σε πια κατηγορία ανήκει το δείγμα ελέγχου.

```
[lp1,rp1,kh1,kp1] = gaussmixp(temp,MU1,Sigma1,W1) ;
[lp2,rp2,kh2,kp2] = gaussmixp(temp,MU2,Sigma2,W2) ;
l1 = length(lp1) ;
l2 = length(lp1) ;
logp1 = sum(lp1) ;
logp2 = sum(lp2) ;
post1 = logp1 + lapriori1*l1 ;
post2 = logp2 + lapriori2*l2 ;
```

Με τις παραπάνω εντολές, μπορεί εύκολα να λάβει τις τελικές posterior πιθανότητες και επειδή ο παραπάνω έλεγχος γίνεται σε διαφορετικά τμήματα ανάλογα με την κατηγορία που τα δείγματα ανήκουν, με μια απλή σύγκριση ελέγχουμε την ορθότητα ή μη της ταξινόμησης, θεωρώντας φυσικά ότι η απόφαση του συστήματος είναι η κατηγορία για την οποία το δείγμα έχει μεγαλύτερη posterior πιθανότητα.

```
if post1>post2
      corr1 = corr1 + 1 ;
else
      fault1 = fault1 + 1 ;
end
```

Έτσι λοιπόν με τον πιο πάνω απλό έλεγχο, εφόσον στο συγκεκριμμένο τμήμα ελέγχονται φυσιολογικά καρδιακά σήματα, όταν η posterior πιθανότητα το δείγμα να ανήκει στην κατηγορία 1 (φυσιολογικά), δηλαδή η post1, είναι μεγαλύτερη από την αντίστοιχη για την κατηγορία 2 (παθολογικά), δηλαδή την post2, τότε μιλάμε για σωστή ταξινόμηση και για λανθασμένη αντίθετα. Τελικά, γίνεται μια τελευταία επεξεργασία, όπου τα αθροίσματα των μετρητών ορθής και λανθασμένης ταξινόμησης, δίνουν τα τελικά ποσοστά επιτυχίας, αλλά και τα επιμέρους ποσοστά missed detection και false alarm και φυσικά παρουσιάζονται τα αντίστοιχα αποτελέσματα με αναλυτικό τρόπο, όπως παρακάτω:

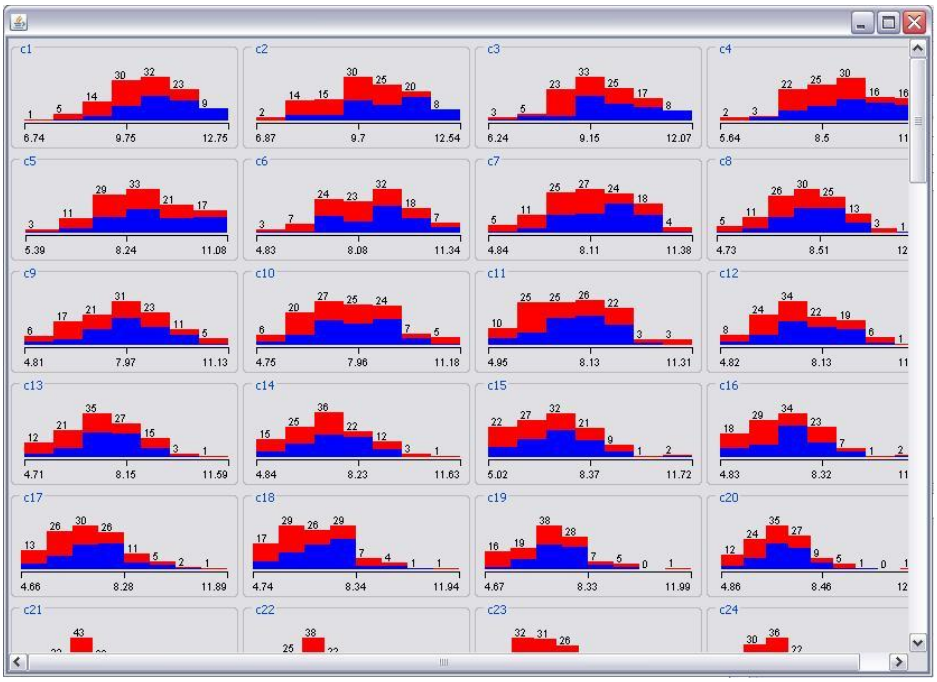
```
=================================================================
====================== EXPERIMENT PARAMETERS ======================
Mixture Order (M): 2
A-priori Probability of Category I (Normal Mrm.): 0.500
A-priori Probability of Category II (Pathological Mrm.): 0.500
====================== FINAL RESULTS ==============================
From 57 normal murmurs 45 were classified correctly and 12 incorrectly
From 57 pathological murmurs 41 were classified correctly and 16 incorrectly
Missed Detection Probability of: 28.070 percent
False Alarm Probability of: 21.053 percent
Success Rate of: 75.439 percent
=================================================================
```

4.2 Παρουσίαση και χρήση των εργαλείων του WEKA

Το WEKA είναι ένα ελεύθερο εργαλείο, που περιέχει μια συλλογή αλγορίθμων μηχανικής μάθησης (machine learning). Οι αλγόριθμοι αυτοί μπορούν να εφαρμόζονται απευθείας σε δεδομένα μέσω της ίδιας της διαπροσωπίας του WEKA ,ή να καλούνται μέσα από κώδικα Java. Το WEKA περιέχει επιπλέον ένα σύνολο εργαλείων για προ-επεξεργασία δεδομένων, για ταξινόμηση, συσταδοποίηση (clustering), για κανόνες συσχέτισης και

οπτικοποίηση, ενώ παράλληλα είναι κατάλληλο για την ανάπτυξη νέων συστημάτων μηχανικής μάθησης.

Στα πλαίσια της εργασίας αυτής, χρησιμοποιήθηκε το γραφικό περιβάλλον διάδρασης που παρέχει το WEKA, προκειμένου να ελεγχθεί η αποδοτικότητα διαφόρων ειδών ταξινομητών, χρησιμοποιώντας τα χαρακτηριστικά που στο προηγούμενο πειραματικό μέρος αποδείχθηκαν πιο αποδοτικά. Τα δεδομένα αυτά, που συγκεκριμμένα είναι τύπου FBANK με πρώτες και δεύτερες διαφορές με μήκος παραθύρου διαφορών 14 και εξαγμένα με παράθυρο των 200 ms, τοποθετήθηκαν σε αρχείο τύπου CSV (Comma Separated Values) με κατάλληλη ένδειξη της κλάσης τους. Έπειτα, χρησιμοποιώντας τον Explorer του WEKA, διευκολύνθηκε η οπτικοποίηση των ιδιοτήτων των χαρακτηριστικών ανά στήλη των διανυσμάτων, ενώ με τη χρήση του σχεδιαστή πειράματος του Experimenter του WEKA, σχεδιάστηκαν τα πειράματα και «έτρεξαν» δίνοντας αποτελέσματα που αποθηκεύτηκαν σε κατάλληλα αρχεία.



Σχήμα 4.2 – Οι κατανομές των χαρακτηριστικών χρησιμοποιώντας εργαλεία Visualization του WEKA

Ο σχεδιαστής πειράματος δίνει την δυνατότητα να σχεδιαστεί ένα πείραμα και να αποθηκευτούν όλες οι παράμετροι σχεδιασμού σε αρχείο, προκειμένου να εκτελεστεί ή τροποποιηθεί αργότερα. Απαραίτητες παράμετροι για το σχεδιασμό, είναι ο προσδιορισμός του αρχείου εισόδου με τα δεδομένα σε κατάλληλη μορφοποίηση (.ARFF, .CSV, κτλ.), ο τύπος του πειράματος (ταξινόμηση ή regression), η μέθοδος (cross fold validation, folds) καθώς επίσης και ο αριθμός επαναλήψεων του πειράματος για πιο στατιστικά ακριβή αποτελέσματα.

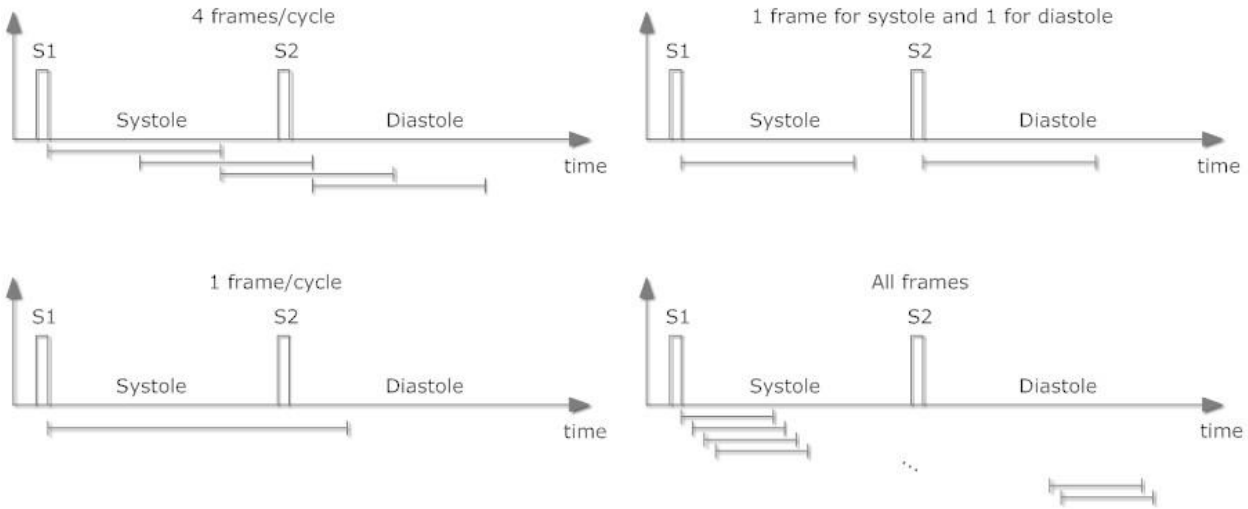
Τέλος, προσδιορίζεται ο αλγόριθμος ταξινόμησης. Έπειτα, με την επιλογή “Run” της ομωνυμης καρτέλας, δίνεται η επιλογή να «τρέξουν» τα πειράματα και τέλος στην καρτέλα “Analyze”, είναι εφικτή η αναπαράσταση των δεδομένων, με επιλογή και απόκρυψη όποιων στατιστικών στοιχείων επιθυμούμε. Στα πλαίσια του πειραματικού μέρους αυτού, αξιολογήθηκαν πειράματα με τη χρήση ταξινομητών Naïve Bayes, NNGE, Bayesian Δικτύων, Simple CART και Hyperpipes. Κάθε πείραμα «έτρεξε» 10 ανεξάρτητες φορές και κάθε φορά το σύστημα χώριζε τα δεδομένα για εκπαίδευση και έλεγχο βάσει της 10 cross fold validation τεχνικής. Τα αποτελέσματα αποθηκεύτηκαν σε κατάλληλα .CSV αρχεία και αναγνώσκονται εύκολα με τη χρήση του εργαλείου “Analyze” του Experimenter.

5 Πειραματική Διαδικασία και Αξιολόγηση των Συστημάτων

5.1 Διαδικασίες και Αποτελέσματα 1^{ου} Πειραματικού Μέρους

Όπως αναφέρθηκε και προηγουμένως το πειραματικό κομμάτι της παρούσας εργασίας, χωρίζεται σε δύο κύρια μέρη. Το πρώτο μέρος, του οποίου τα αποτελέσματα παρουσιάζονται σε αυτήν την ενότητα, αφορά στην σύγκριση διαφόρων ειδών χαρακτηριστικών, με βάση την απόδοση της ταξινόμησης με χρήση του στατικού GMM ταξινομητή που αναπτύχθηκε. Προκειμένου να επιλεγεί με σχετική αξιοπιστία ο καταλληλότερος τύπος χαρακτηριστικών, έγιναν έλεγχοι σε πολλές και διάφορες παραμέτρους. Κατ'αρχήν, επιλέχθηκε ο αλγόριθμος επιλογής χαρακτηριστικών που ήταν πιο αποδοτικός. Ελέχθηκαν τέσσερεις τέτοιοι αλγόριθμοι:

- Ο πρώτος (I) αλγόριθμος επιλέγει διανύσματα χαρακτηριστικών από τέσσερα παράθυρα για κάθε καρδιακό κύκλο, όπως φαίνεται πάνω αριστερά στο σχήμα 5.1.
- Ο δεύτερος (II) αλγόριθμος επιλέγει χαρακτηριστικά από ένα παράθυρο ανά καρδιακό κύκλο, όπως φαίνεται κάτω αριστερά στο σχήμα 5.1.
- Ο τρίτος (III) αλγόριθμος επιλέγει όλα τα διανύσματα χαρακτηριστικών του καρδιακού κύκλου με μεγάλη αλληλεπικάλυψη, κάτω δεξιά στο σχήμα 5.1.
- Ο τέταρτος (IV) αλγόριθμος επιλέγει ένα παράθυρο στη συστολική περίοδο και ένα στη διαστολική, όπως φαίνεται πάνω δεξιά στο σχήμα 5.1.



Σχήμα 5.1 – Οι διάφοροι τρόποι επιλογής χαρακτηριστικών που ελέγχθηκαν

Σε καθέναν από τους παραπάνω αλγορίθμους που αναφέρθηκαν, έγιναν οι αντίστοιχες δοκιμές και βελτιστοποιήσεις των παραμέτρων. Τέτοιες παράμετροι σε αυτό το στάδιο είναι το μήκος του παραθύρου (100ms, 200ms, 300ms, κ.ο.κ.) και το ποσοστό αλληλεπικάλυψης (overlapping) (25%, 50%, 75%).

Να σημειωθεί εδώ, ότι οι διάφορες τιμές που ελέγχθηκαν στα πειράματα, διαφέρουν κατά περίπτωση κι αυτό γιατί πάντα τα ίδια τα αποτελέσματα οδηγούσαν την πειραματική διαδικασία κι έτσι ο σχεδιασμός τους γινόταν “on the fly”. Στη συνέχεια, έγιναν μεταβολές στις διάφορες παραμέτρους χαρακτηριστικών, όπως για παράδειγμα, η προσθήκη επιθεμάτων ενέργειας, διαφορών και άλλα. Τέλος, έγιναν διάφορες δοκιμές σχετικά με την τάξη του μίγματος των μοντέλων. Παρουσιάζονται στους ακόλουθους πίνακες παραστατικά τα επιμέρους αποτελέσματα που προέκυψαν, καθώς επίσης και καμπύλες Receiver Operating Characteristic (ROC) με τις αντίστοιχες Area Under Curve (AUC) μετρικές για τα καλύτερα αποτελέσματα ανά είδος χαρακτηριστικών. Η καμπύλη ROC αποτελεί ένα μέτρο της διαχωριστικής ικανότητας για το χαρακτηριστικό μας στις δυο κλάσεις, ενώ η ικανότητα αυτή ποσοτικοποιείται ως η περιοχή που ορίζεται μεταξύ των καμπυλών (AUC).

5.1.1 Αποτελέσματα με χρήση των MFCCs με και χωρίς CMN

Σύγκριση αλγορίθμων:

Αλγόριθμος	**I**	**II**	**III**	**IV**
Ποσοστό επιτυχίας	**70.125%**	62.281%	67.544%	54.386%
Missed Detection	33.333%	40.351%	35.088%	15.789%
False Alarm	26.316%	35.088%	29.825%	75.439%

Σύγκριση διαφόρων μηκών παραθύρου:

Μήκος Παραθύρου	**100ms**	**200ms**	**300ms**	**400ms**	**500ms**
Ποσοστό επιτυχίας	57.895%	62.281%	66.667%	**70.175%**	63.158%
Missed Detection	40.351%	40.351%	38.596%	33.333%	42.105%
False Alarm	43.860%	35.088%	28.070%	26.316%	31.579%

Σύγκριση διαφόρων τύπων παραμέτων:

Τύποι	MFCC_0	MFCC_0_E	MFCC_E_D	MFCC_E_D_A	MFCC_0_D	MFCC_0_D_A
Ποσοστό επιτυχίας	**70.125%**	63.158%	63.158%	63.158%	57.895%	57.018%
Missed Detection	33.333%	40.351%	29.825%	31.579%	35.088%	38.596%
False Alarm	26.316%	33.333%	43.86%	42.105%	49.123%	47.368%

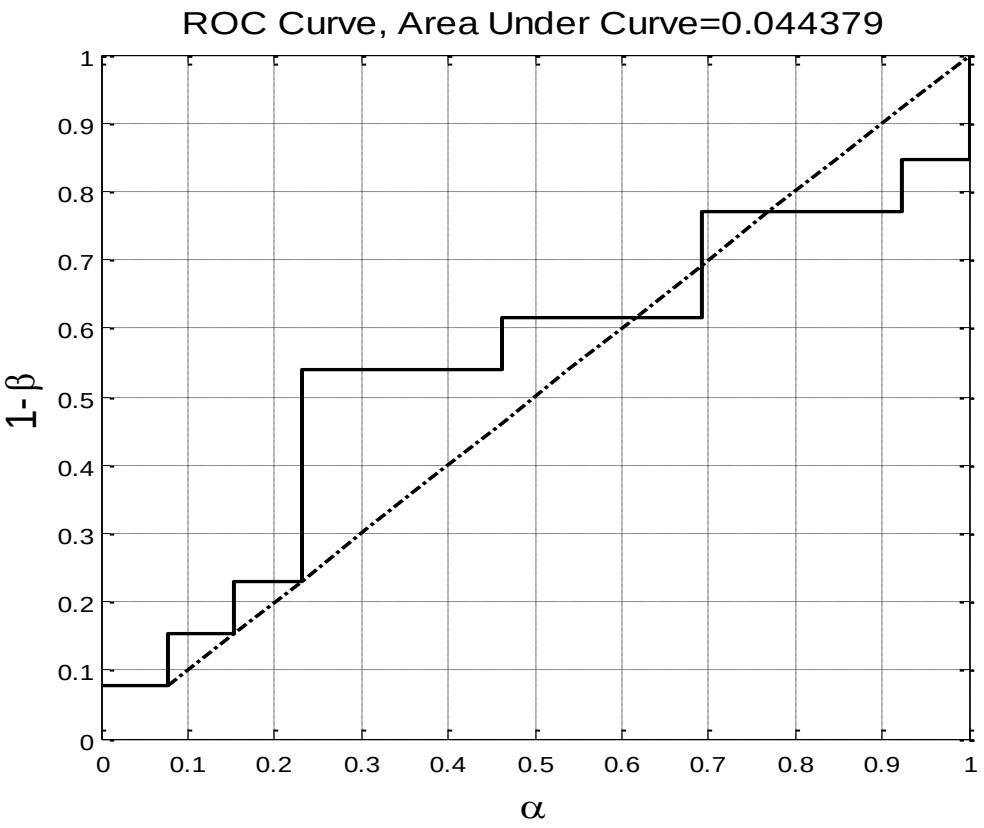
Σύγκριση τάξεων μίγματος:

Τάξη Μίγματος	1	2	3	4	5	6	7	8	9
Ποσοστό επιτυχίας(%)	61.404	**70.125**	60.526	62.281	62.281	64.912	61.404	59.649	57.895
Missed Detection(%)	40.351	33.333	38.596	40.351	38.596	29.825	35.088	43.86	40.351
False Alarm (%)	36.842	26.316	40.351	35.088	36.842	40.351	42.105	36.842	43.86

Τέλος, εφαρμόστηκε η τεχνική CMN στα δεδομένα και συγκριτικά παρατίθεται:

Τεχνική	Χωρίς CMN	Με CMN
Ποσοστό επιτυχίας	**70.125%**	68.421%
Missed Detection	33.333%	25.561%
False Alarm	26.316%	38.596%

Η αντίστοιχη καμπύλη ROC για τις βέλτιστες παραμέτρους:



5.1.2 Αποτελέσματα με χρήση των MELSPEC

Σύγκριση αλγορίθμων:

Αλγόριθμος	I	II	III	IV
Ποσοστό επιτυχίας	68,421%	68,421%	**74,561%**	63,158%
Missed Detection	38,596%	28,070%	22,807%	68,421%
False Alarm	24,561%	35,088%	28,070%	5,263%

Σύγκριση διαφόρων μηκών παραθύρου:

Μήκος Παραθύρου	**100ms**	**200ms**	**300ms**	**400ms**
Ποσοστό επιτυχίας	**66.667**	64.912	64.035	63.158
Missed Detection	35.088	38.596	36.842	36.842
False Alarm	31.579	31.579	35.088	36.842

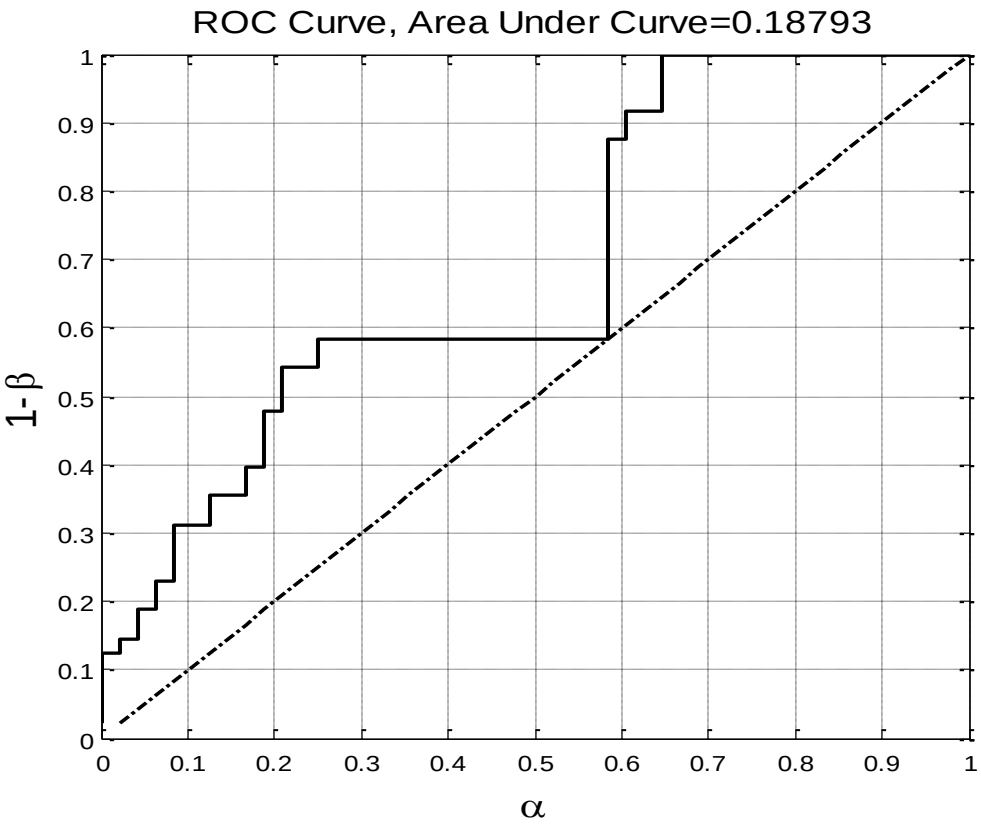
Σύγκριση διαφόρων τύπων παραμέτων:

Τύποι	**MELSPEC**	**MELSPEC_E**	**MELSPEC_D (2)**	**MELSPEC_D (4)**	**MELSPEC_D_A**
Ποσοστό επιτυχίας	66.667	65.789	**69.298**	68.421	66.667
Missed Detection	35.088	35.088	33.333	35.088	38.596
False Alarm	31.579	33.333	28.07	28.07	28.07

Σύγκριση τάξεων μίγματος:

Τάξη Μίγματος	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Ποσοστό επιτυχίας(%)	64.035	69.298	67.544	68.421	70.125	**74.561**	72.807	71.93	68.421
Missed Detection(%)	36.842	33.333	42.105	38.596	35.088	22.807	26.316	31.579	29.825
False Alarm (%)	35.088	28.07	22.807	24.561	24.561	28.07	28.07	24.561	33.333

Η αντίστοιχη καμπύλη ROC για τις βέλτιστες παραμέτρους:



5.1.3 Αποτελέσματα με χρήση των FBANK

Σύγκριση αλγορίθμων:

Αλγόριθμος	I	II	III	IV
Ποσοστό επιτυχίας	**75,439%**	65,789%	68,421%	65,789%
Missed Detection	28,070%	29,825%	42,105%	29,825%
False Alarm	21,053%	38,596%	21,053%	38,596%

Σύγκριση διαφόρων μηκών παραθύρου:

Μήκος Παραθύρου	**100ms**	**200ms**	**300ms**
Ποσοστό επιτυχίας	65,789%	**71,053%**	63,158%
Missed Detection	43,860%	42,105%	54,386%
False Alarm	24,561%	15,789%	19,298%

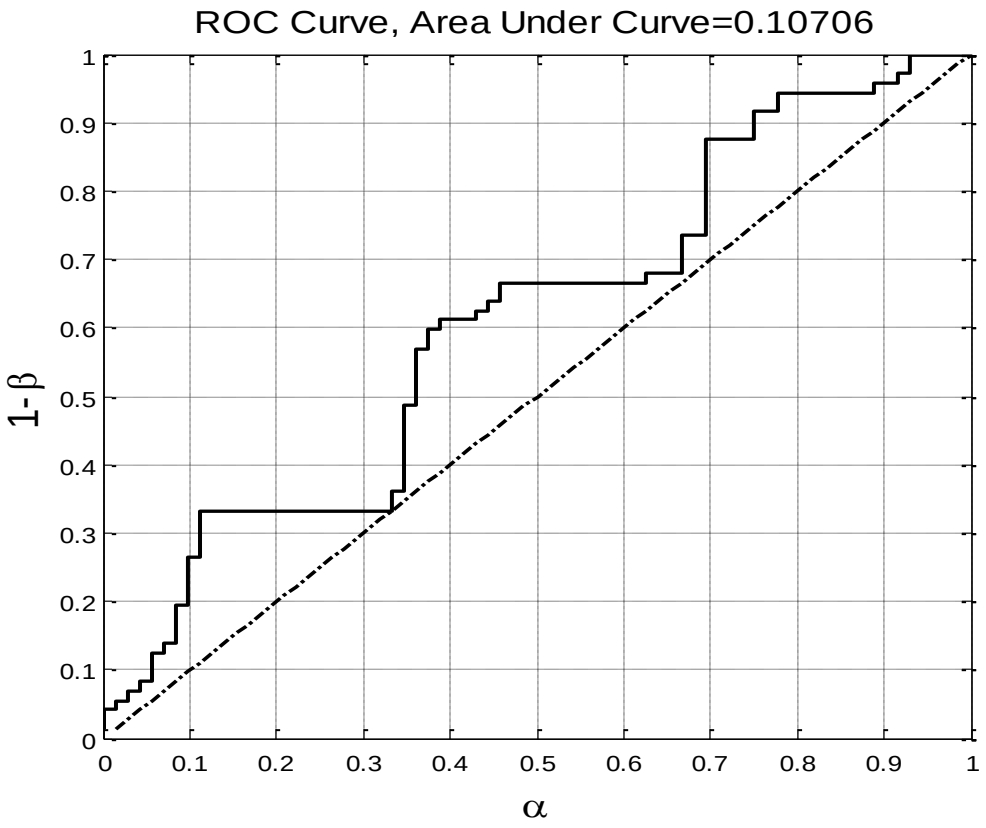
Σύγκριση διαφόρων τύπων παραμέτων:

Τύποι	**FBANK**	**FBANK_E**	**FBANK_D_A (2)**	**FBANK_D_A (12)**	**FBANK_D_A (14)**	**FBANK_D_A (16)**
Ποσοστό επιτυχίας	71.053%	70.175%	71.93%	72.807%	**75.439%**	74.561%
Missed Detection	42.105%	42.105%	33.333%	31.579%	28.07%	29.825%
False Alarm	15.789%	17.544%	22.807%	22.807%	21.053%	21.053%

Σύγκριση τάξεων μίγματος:

Τάξη Μίγματος	**1**	**2**	**3**	**4**	**5**
Ποσοστό επιτυχίας(%)	64.035	**75.439**	74.561	66.667	66.667
Missed Detection(%)	36.842	28.07	28.07	36.842	36.842
False Alarm (%)	35.088	21.053	22.807	29.825	29.825

Η αντίστοιχη καμπύλη ROC για τις βέλτιστες παραμέτρους:



5.1.4 Αποτελέσματα με χρήση των PLPs

Σύγκριση αλγορίθμων:

Αλγόριθμος	I	II	III	IV
Ποσοστό επιτυχίας	60.526%	**68.421%**	67.544%	52.632%
Missed Detection	29.825%	21.053%	29.825%	26.316%
False Alarm	49.123%	42.105%	35.088%	68.421%

Σύγκριση διαφόρων μηκών παραθύρου:

Μήκος Παραθύρου	**100ms**	**200ms**	**300ms**	**400ms**	**500ms**
Ποσοστό επιτυχίας	31.579%	44.737%	58.772%	**59.649%**	52.632%
Missed Detection	75.439%	57.895%	36.842%	43.86%	40.351%
False Alarm	61.404%	52.632%	45.614%	36.842%	54.386%

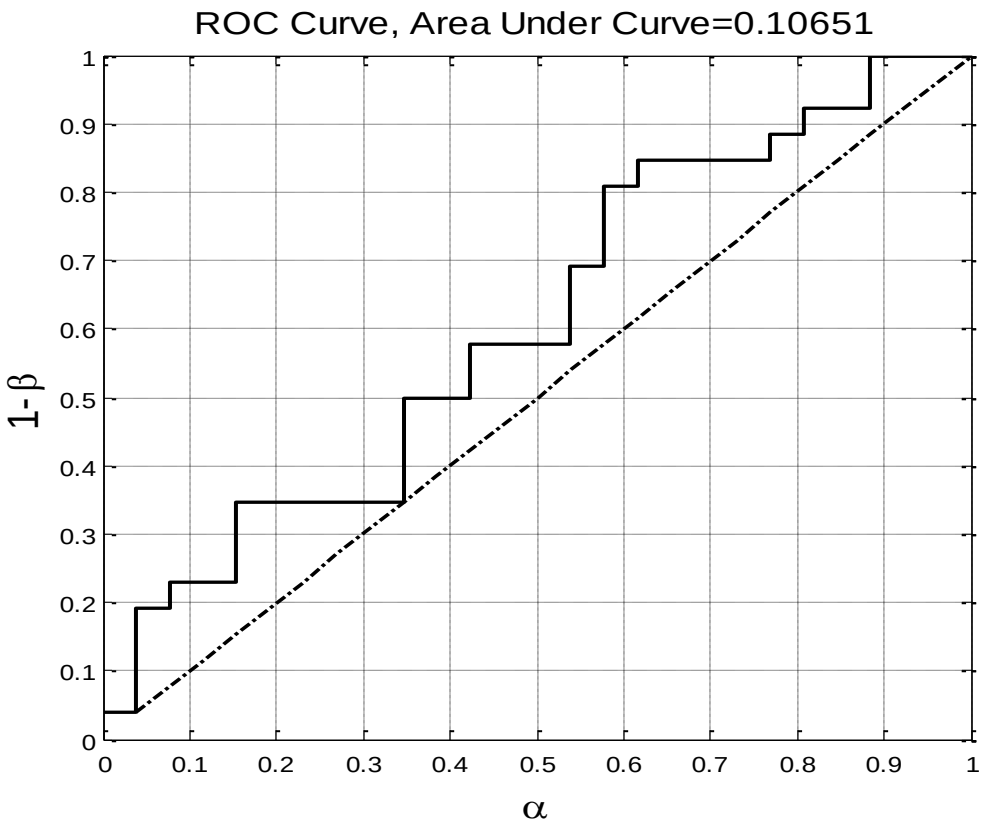
Σύγκριση διαφόρων τύπων παραμέτων:

Τύποι	**PLP**	**PLP_0**	**PLP_E**	**PLP_E_D (2)**	**PLP_E_D (4)**	**PLP_E_D_A**
Ποσοστό επιτυχίας	59.649%	57,018%	64.912%	**68.421%**	65.789%	56.140%
Missed Detection	43.86%	42.105%	35.088%	21.053%	31.579%	38.596%
False Alarm	36.842%	43.86%	35.088%	42.105%	36.842%	49.123%

Σύγκριση τάξεων μίγματος:

Τάξη Μίγματος	**1**	**2**	**3**	**4**
Ποσοστό επιτυχίας(%)	65.789	**68.421**	67.544	60.526
Missed Detection(%)	26.316	21.053	33.333	42.105
False Alarm (%)	42.105	42.105	31.579	36.842

Η αντίστοιχη καμπύλη ROC για τις βέλτιστες παραμέτρους:



Συγκεντρωτική σύγκριση χαρακτηριστικών του μέρους 5.1:

Χαρακτηριστικό	**MFCCs**	**MELSPEC**	**FBANK**	**PLPs**
Ποσοστό επιτυχίας	70.125%	74.561%	**75.439%**	68.421%
Αλγόριθμος	I	III	I	II
Μήκος Παραθύρου	400ms	100ms	200ms	400ms
Τάξη μίγματος	2	6	2	2
Τύπος Παραμέτρων	MFCC_0	MELSPEC_D (2)	FBANK_D_A (14)	PLP_E_D (2)

5.2 Διαδικασίες και Αποτελέσματα 2ου Πειραματικού Μέρους

Σε αυτό το πειραματικό κομμάτι, έγινε μια σύγκριση διάφορων ταξινομητών ως προς την απόδοσή τους. Χρησιμοποιήθηκαν τα χαρακτηριστικά, τα οποία βάση του προηγούμενου μέρους αποδείχθηκαν πειραματικά πιο αποδοτικά.

Πειραματικά αποτελέσματα με χρήση του Naïve Bayes ταξινομητή:

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          76               66.6667 %
Incorrectly Classified Instances        38               33.3333 %
Kappa statistic                          0.3333
Mean absolute error                      0.3322
Root mean squared error                  0.5754
Relative absolute error                 66.4067 %
Root relative squared error            115.0284 %
Total Number of Instances              114

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.649     0.316      0.673     0.649     0.661      0.716
a
                 0.684     0.351      0.661     0.684     0.672      0.737
b
Weighted Avg.    0.667     0.333      0.667     0.667     0.667      0.726

=== Confusion Matrix ===

  a  b   <-- classified as
 37 20 |  a = a
 18 39 |  b = b
```

Πειραματικά αποτελέσματα με χρήση του Nearest Neigbor Generalized Exemplar (NNGE) ταξινομητή:

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          79               69.2982 %
Incorrectly Classified Instances        35               30.7018 %
Kappa statistic                          0.386
Mean absolute error                      0.307
Root mean squared error                  0.5541
Relative absolute error                 61.3727 %
Root relative squared error            110.7599 %
Total Number of Instances              114
```

```
=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.632     0.246      0.72      0.632     0.673      0.693
a
                 0.754     0.368      0.672     0.754     0.711      0.693
b
Weighted Avg.    0.693     0.307      0.696     0.693     0.692      0.693

=== Confusion Matrix ===

  a  b   <-- classified as
 36 21 |  a = a
 14 43 |  b = b
```

Πειραματικά αποτελέσματα με χρήση του Bayesian Net ταξινομητή:

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          82               71.9298 %
Incorrectly Classified Instances        32               28.0702 %
Kappa statistic                          0.4386
Mean absolute error                      0.2775
Root mean squared error                  0.5248
Relative absolute error                 55.4763 %
Root relative squared error            104.908  %
Total Number of Instances              114

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.737     0.298      0.712     0.737     0.724      0.734
a
                 0.702     0.263      0.727     0.702     0.714      0.786
b
Weighted Avg.    0.719     0.281      0.72      0.719     0.719      0.76

=== Confusion Matrix ===

  a  b   <-- classified as
 42 15 |  a = a
 17 40 |  b = b
```

Πειραματικά αποτελέσματα με χρήση του Simple CART ταξινομητή:

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          72               63.1579 %
Incorrectly Classified Instances        42               36.8421 %
Kappa statistic                          0.2632
Mean absolute error                      0.3986
Root mean squared error                  0.5374
Relative absolute error                 79.68   %
Root relative squared error            107.4178 %
Total Number of Instances              114

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.561     0.298      0.653     0.561     0.604      0.632
a
                 0.702     0.439      0.615     0.702     0.656      0.632
b
Weighted Avg.    0.632     0.368      0.634     0.632     0.63       0.632

=== Confusion Matrix ===

  a  b   <-- classified as
 32 25 |  a = a
 17 40 |  b = b
```

Πειραματικά αποτελέσματα με χρήση του Hyperpipes ταξινομητή

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          75               65.7895 %
Incorrectly Classified Instances        39               34.2105 %
Kappa statistic                          0.3158
Mean absolute error                      0.4976
Root mean squared error                  0.4977
Relative absolute error                 99.4799 %
Root relative squared error             99.4873 %
Total Number of Instances              114

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.649     0.333      0.661     0.649     0.655      0.698
a
                 0.667     0.351      0.655     0.667     0.661      0.698
b
Weighted Avg.    0.658     0.342      0.658     0.658     0.658      0.698
```

```
=== Confusion Matrix ===

  a  b   <-- classified as
 37 20 |  a = a
 19 38 |  b = b
```

5.3 Σύγκριση Αποτελεσμάτων

Όπως φαίνεται και από τα πειραματικά αποτελέσματα, προκύπτουν ξεκάθαρα ορισμένα συμπεράσματα μετά από όλες αυτές τις δοκιμές.

Το πιο αποδοτικό χαρακτηριστικό που μετέφερε όλη την απαραίτητη πληροφορία και βελτίωνε περισσότερο από τα υπόλοιπα την απόδοση του ταξινομητή, είναι αυτό του FBANK. Ακόμα πιο συγκεκριμμένα στο πρόβλημά μας, αυτό δηλαδή της κατηγοριοποίησης φυσιολογικών και παθολογικών φυσημάτων, πιο αποδοτικά ήταν τα FBANK χαρακτηριστικά, που εξάχθηκαν με χρήση παραθύρων των 200ms, ενώ μάλιστα λάβαμε τις πρώτες και δεύτερες διαφορές με παράθυρο μήκους 14. Το συγκεκριμμένο χαρακτηριστικό διαχειριζόμενο από τον βασισμένο σε GMM ταξινομητή που αναπτύξαμε, επέτρεπε ποσοστό επιτυχίας **75.439%**. Ακολουθούσαν τα λιγότερο αποδοτικά χαρακτηριστικά MELSPEC, MFCCs και PLPs, με ποσοστά επιτυχούς ταξινόμησης 74.561%, 70.425% και 68.421% αντίστοιχα.

Στη συνέχεια και στο δεύτερο πειραματικό μέρος, με τη χρήση των πιο αποδοτικών χαρακτηριστικών, σύμφωνα πάντα με τα πειραματικά αποτελέσματα του πρώτου μέρους, ελέγχθηκε η αποτελεσματικότητα ορισμένων ταξινομητών. Εκτός του GMM ταξινομητή, που αναπτύχθηκε στα πλαίσια της εργασίας αυτής και ελέγχθηκε στο πρώτο μέρος, χρησιμοποιήθηκαν επίσης οι ταξινομητές Naïve Bayes, NNGE, Bayesian Net, Simple CART και HyperPipes, με ποσοστά επιτυχίας 66.667%, 69.298%, **71.929%**, 63.157% και 65.789% αντίστοιχα.

6 Συμπεράσματα

Δεδομένου ότι κατά το μεγαλύτερο ποσοστό ο σχολικός πληθυσμός θα παρουσιάσει κάποια στιγμή ένα φυσιολογικό φύσημα, αποκτά μεγάλη κλινική σημασία ο ασφαλής διαχωρισμός μεταξύ των παθολογικών φυσημάτων – φυσήματα που σηματοδοτούν καρδιαγγειακή ανωμαλία – και των φυσιολογικών φυσημάτων. Οι σύγχρονες απεικονιστικές και μη επεμβατικές μέθοδοι, βοηθούν στην ακριβή διάγνωση και έτσι η διαγνωστική ικανότητα των Παιδιάτρων και γενικών Ιατρών φθίνει στον χρόνο. Το κόστος των σύγχρονων διαγνωστικών μεθόδων που προαναφέρθηκαν ωστόσο είναι πολύ υψηλό και καθιστά αδύνατη την καθολική χρήση τους από όλα τα μέλη του σχολικού πληθυσμού, ενώ επίσης δεν προσφέρει και κάποιο τρόπο για μαζική αξιολόγηση του πληθυσμού αυτού. Από τα παραπάνω, προκύπτει ότι η ανάπτυξη ενός αυτόματου συστήματος διαχωρισμού φυσιολογικών και παθολογικών φυσημάτων, θα ήταν πολύ χρήσιμη για την εκπαίδευση του Ιατρικού προσωπικού, καθώς επίσης και για την υποστήριξη αποφάσεων ως προς την ορθότητα και ακρίβεια της Ιατρικής διάγνωσης. Αυτό εκτός των άλλων, θα βελτιώσει την διαγνωστική ικανότητα των Ιατρών κατά την διαδικασία της καρδιακής ακρόασης.

Επιλέχθηκαν δύσκολες διαγνωστικές περιπτώσεις, που συμπεριλάμβαναν έντονα φυσιολογικά φυσήματα, αντιληπτά από όλες τις θέσεις ακρόασης, καθώς επίσης και συστολικά παθολογικά φυσήματα. Τα καρδιακά σήματα με κατάλληλη ετικετοποίηση, δέχθηκαν προεπεξεργασία και συγκεκριμμένα φιλτράρισμα και υποδειγματοληψία ενώ στην συνέχεια, χώρα έλαβε η αυτόματη κατάτμισή τους από αλγόριθμο βασισμένο στην περιβάλλουσα ενέργεια Shannon που αναπτύχθηκε στα πλαίσια της εργασίας αυτής. Τα χαρακτηριστικά χρονικά σημεία (time landmarks) των καρδιακών σημάτων, που παρήγαγε ο αλγόριθμος αυτός (αρχή της συστολής, τόνος S1 και άλλα), αποθηκεύτηκαν σε κατάλληλη δομή δεδομένων, προκειμένου να χρησιμοποιηθούν για την επιλογή των διαστάσεων των παραθύρων. Στην συνέχεια, υλοποιήθηκε ένας στατικός ταξινομητής, βασισμένος σε μίγμα γκαουσιανών κατανομών (GMM), με σκοπό την κατηγοριοποίηση φυσιολογικών και παθολογικών φυσημάτων. Βελτιστοποιήθηκαν οι παράμετροι λειτουργίας του ταξινομητή, καθώς και οι παράμετροι των χαρακτηριστικών και επιτεύχθηκε απόδοση 75,44% ως προς την ακρίβεια ταξινόμησης. Στο δεύτερο μέρος, συγκρίθηκαν επιπλέον ταξινομητές με καλύτερο ανάμεσά τους τον βασισμένο στο Μπαγεσιανό δίκτυο (Bayesian Net), με απόδοση του 71.93%. Οι αποδόσεις που επιτεύχθηκαν στο πειραματικό μέρος της εργασίας αυτής, είναι αρκετά αξιόλογες, έχοντας υπ’όψιν το πολύ περιορισμένο πλήθος των δεδομένων και φυσικά το γεγονός, ότι στην παρούσα εργασία επιχειρείται η αυτόματη ταξινόμηση των καρδιακών σημάτων σε σήματα με λειτουργικά φυσήματα και σε σήματα με παθολογικά, σε αντίθεση με

προηγούμενες εργασίες του χώρου, όπου γίνεται ταξινόμηση μεταξύ φυσιολογικών σημάτων και παθολογικών.

Σαφέστατα υπάρχουν πολλά περιθώρια για την βελτίωση των αποτελεσμάτων που η εργασία αυτή φέρει. Αρχικά, μια πρώτη ελαχιστοποίηση του ποσοστού λάθους του αλγορίθμου κατάτμησης από το 8,1% σε χαμηλότερα επίπεδα, θα επιδρούσε θετικά στα τελικά αποτελέσματα, μιας και θα βελτίωνε την ακρίβεια του ορισμού των επιμέρους σταδίων του καρδιακού σήματος. Φυσικά, η αύξηση των δεδομένων θα βελτίωνε την απόδοση του συστήματος, καθώς όσο περισσότερα τα δεδομένα, τόσο πιο ακριβής θα είναι η εκπαίδευση (training) των μοντέλων και πιο λεπτομερής ο έλεγχος (testing). Τέλος, στα πλαίσια της εργασίας αυτής, μπορούν οι ταξινομητές και μέθοδοι που χρησιμοποιήθηκαν να συγκριθούν με ένα σύστημα ταξινομητή βασισμένο σε HMMs. Στο μέλλον, πολλές βελτιώσεις αναμένονται και σε αυτό το χώρο της τεχνολογίας μέσα από νέες μεθόδους, αλλά και από την χρήση γνωστών εργαλείων από την επεξεργασία φωνής – όπως άλλωστε έγινε και σε αυτή την εργασία – πάντα προσαρμοσμένων στις συνθήκες. Το μόνο βέβαιο ωστόσο είναι, ότι η προσεκτική και ενδελεχής ιατρική εξέταση, καθώς και η αξιολόγηση του ιστορικού του ασθενούς είναι αναντικατάστατα.

7 Βιβλιογραφία

1. Ι. Γερμανάκης, "*Καρδιακή Ακρόαση στην Παιδική Ηλικία*", Εκπαιδευτικό εγχειρίδιο πολυμέσων, Εκδόσεις Λύχνος, Ηράκλειο, 2010
2. Teaching Material, "*From Sound to Ultrasound: Multimedia based Pediatric Cardiology IP*", Erasmus Intensive Program, Faculty of Medicine, University of Crete, Herakleion, 2010
3. Γ. Μουστακίδης, "*Βασικές Τεχνικές Ψηφιακής Επεξεργασίας Σημάτων*", Εκδόσεις Τζιόλα, Θεσσαλονίκη, 2004
4. Σ. Θεοδωρίδης, Ά. Πικράκης, Κ. Κουτρουμπάς, Δ. Κάβουρας, "*Εισαγωγή στην Αναγνώριση Προτύπων με MATLAB*", Εκδόσεις Πασχαλίδη, Αθήνα, 2011
5. Remco R Bouckaert, Eibe Frank, Mark Hall, Richard Kirkby, Peter Reutemann, Alex Seewald, David Scuse, "*WEKA Manual for Version 3-6-4*", University of Waikato, Hamilton, New Zealand, 2010
6. Steve Young, Gunnar Evermann, Mark Gales, Thomas Hain, Dan Kershaw, Xunying Liu, Gareth Moore, Julian Odell, Dave Ollason, Dan Povey, Valtcho Valthcev, Phil Woodland, "*The HTK Book*", Cambridge University, Cambridge, United Kingdom, 2009
7. Μ. Μαρκάκη, Ι. Γερμανάκης, Ι. Στυλιανού, «*Αυτόματη ανίχνευση Φυσημάτων σε Παιδικά Φωνοκαρδιογραφήματα και Ανάπτυξη Λογισμικού Εκπαίδευσης στην Παιδιατρική Ακρόαση*»
8. Markaki M, "*Selection of Relevant Features for Audio Classification Tasks (Philosophy of Doctoral)*", Department of Computer Science, University of Crete, Herakleion, 2011
9. Anna – Leena Noponen, Sakari Lukkarinen, Anna Angerla, Raimo Sepponen, "*Phono – Spectrographic Analysis of Heart Murmur in Children*", BMC Pediatrics Journal.
10. Vladimir Kudriavtsev, Vladimir Polyshchuk, Douglas L. Roy, "*Heart Energy Signature Spectrogram for Cardiovascular Diagnosis*", BioMedical Engineering OnLine Journal.
11. Faizan Javed, P. A. Venkatachalam, Ahmad Fadzil, "*A Signal Processing Module for the Analysis of Heart Sounds and Heart Murmurs*", Institute Of Physics (IOP) Journal.
12. Yong Joo Chung, "*A Classification Approach for the Heart Sound Signals Using Hidden Markov Models*".
13. Bhatikar, DeGroff, Mahajan, "*A Classifier Based on Artificial Neural Network Approach for Cardiac Ascultation in Pediatrics*", Artificial Inteligence Med. 33 (2005) 251-260.
14. MD Morton E. Tavel, "*Cardiac Ascultation: A Glorious Past – But Does It Have a Future?*", American Heart Association, 93 (1996) 1250-1253.
15. Matias Brusco, Homer Nazeran, "*Digital Phonocardiography: A PDA-based Approach*", Proc. Of the 26th Intl. Conf. IEEE EMBS, September 2004.

16. Chul Kwak, Oh-Wook Kwon, "*Cardiac Disorder Classification Based on Extreme Learning Machine*", World Academy of Science, Engineering and Technology 48 (2008).
17. Carlos S Lima, Manuel J Cardoso, "*Phonicardiogram Segmentation by Using an Hybrid RBF-HMM Model*".
18. Ervin Sejdic, Jin Jiang, "*Pattern Recognition in Time-Frequency Domain: Selective Regional Correlation and Its Applications*".
19. Stasis A, Loukis E N, Pavlopoulos S A, Koutsouris "*A Multiple Decision Tree Architecture for Medical Diagnosis: The Differentiation of Opening Snap, Second Heart Sound, Split and Third Heart Sound*", 2004 CMS Springer Verlag 254-74.
20. DeGroff C G, Sanjay, Bhatikar, Hertzberg J, Shandas R, Valdes-Curz L, Mahajan R L, "*Artificial Neural Network – based Method of Screening Heart Murmurs in Children*", American Heart Association, 103 (2001) 2711-16.
21. Mecklendreuker W, Hlawatsch F, "*The Wigner Distribution – Theory and Applications in Signal Processing*", 1997 Netherlands Elesvier.
22. Noponen A L, Lukkarinen S, Sikio K, Angerla A, Sepponen R "*How to Recognize the Innocent Vibratory Murmur*", 27 (2000) 561-564.
23. Advani N, Menahem S, Wilkinson J L "*The Diagnosis of Innocent Murmur in Childhood*", Cardiol Young 10 (2000) 340-342.
24. Rappaport M B, Sprague H B "*Physiologic and Physical Laws that Govern Auscultation and their Clinical Application; the Acoustic Stethoscope and the Electrical Amplifying Stethoscope and Stethograph*", Am Heart J 21 (1941) 257-318.
25. Tavel M E, Brown D D, Shander D "*Enhanced Ascultation with a New Graphic Display System*", Arch Intern Med 154 (1994) 893-898.
26. Norgard G, Greve G, Rosland G A, Barg A "*Referral Practice and Clinical Assesement of Heart Murmurs in Children*" Tidsskr Nor Laegeforen 125 (2005) 996-998.
27. Van Oort A, Le Blanc-Bodten M, De Boo T, Van der Werf T, Daniels O "*The Vibratory Innocent Heart Murmur in Schoolchildren: Difference in Auscultatory Findings Between School Medical Officers and Pediatric Cardiologist*", Pediatr Cardiol 15 (1994) 282-287.
28. Shaver J A, Salerni R, Reddy P S "*Normal and Abnormal Heart Sounds in Cardiac Diagnosis*", 10 (1985) 1-68.
29. http://www.nhlbi.nih.gov/health/health-topics/topics/heartmurmur/
30. http://www.babyspace.gr/Article/child-heart-problems/138-1327.html
31. http://www.sciencedirect.com/science/article/pii/S0957417406003927#secx3
32. http://www.sciencedirect.com/science/article/pii/S1746809411000292
33. http://www.bsignetics.com

8 Συχνές Συντομογραφίες

PCG	**P**hono**C**ardio**G**ram
ECG	**E**lectro**C**ardio**G**ram
GMM	**G**aussian **M**ixture **M**odel
MFCCs	**M**el-**F**requency **C**epstral **C**oefficient**s**
PLPs	**P**erceptual **L**inear **P**rediction coefficient**s**
HTK	**H**idden Markov model **T**ool **K**it
CMN	**C**epstral **M**ean **N**ormalization
NNGE	**N**earest **N**eighbor **G**eneralized **E**xemplars
CART	**C**lassification **A**nd **R**egression **T**rees