

Τ.Ε.Ι. Κρήτης, Παράρτημα Ρεθύμνου
Τμήμα Μουσικής Τεχνολογίας & Ακουστικής

**ΑΥΤΟΜΑΤΗ ΑΚΟΥΣΤΙΚΗ ΤΑΥΤΟΠΟΙΗΣΗ ΤΟΥ
ΚΟΚΚΙΝΟΥ ΡΥΓΧΩΤΟΥ ΚΑΘΑΡΟΥ ΤΩΝ
ΦΟΙΝΙΚΟΕΙΔΩΝ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ



Σπουδαστές
Σταμάτης Ευστράτιος
Φωτίου Αναστάσιος

Επιβλέπων
Δρ. Ποταμίτης Ηλίας

Ρέθυμνο Οκτώβριος 2006

Θα θέλαμε να ευχαριστήσουμε τον επιβλέποντα καθηγητή μας, Δρ. Ηλία Ποταμίτη, για την εμπιστοσύνη που μας έδειξε με την ανάθεση της παρούσας διπλωματικής εργασίας. Η βοήθειά του υπήρξε πολύτιμη για την διεκπεραίωσή της.

*Σταμάτης Ευστράτιος
Φωτίου Αναστάσιος*

ΠΡΟΛΟΓΟΣ

Αυτή η διπλωματική εργασία αποτελεί την προσπάθειά μας όσον αφορά την ανάπτυξη και την, εκ βαθέων, μελέτη ενός ολοκληρωμένου και λειτουργικού συστήματος, για την αυτόματη, ακουστική ανίχνευση ενός συγκεκριμένου εντόμου, της οικογένειας των κολεόπτρων, του ρυγχωτού κανθάρου (*Rhynchophorus Ferrugineus*).

Συγκεκριμένα κάνουμε μία αναφορά στα ήδη υπάρχοντα, αυτόματα συστήματα αναγνώρισης φωνής και μελετάμε τα στάδια που αποτελούν ένα τέτοιο ολοκληρωμένο σύστημα, καθώς και τις μεθόδους επεξεργασίας σήματος που χρησιμοποιούνται σε κάθε υπομονάδα του. Στη συνέχεια ερευνούμε το κατά πόσο, οι προαναφερθείσες μέθοδοι και τεχνικές μπορούν να εφαρμοστούν σε ένα σύστημα ακουστικής ανίχνευσης εντόμων και συγκεκριμένα του εντόμου, με του οποίου την ακουστική συμπεριφορά, ασχολούμαστε. Χρησιμοποιώντας ηχογραφημένα δείγματα ακουστικής εκπομπής του ρυγχωτού κανθάρου, καθώς και ειδικά προγράμματα επεξεργασίας σήματος, αναλύουμε τρόπους μοντελοποίησης των δειγμάτων αυτών με τελικό σκοπό την αυτόματη επιβεβαίωση της ύπαρξης του εντόμου και την αναγνώρισή του. Επίσης προτείνονται μέθοδοι απόκτησης ηχογραφημάτων, της ακουστικής εκπομπής του, καθώς και μηχανήματα που συντελούν στο έργο αυτό.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1. ΕΙΣΑΓΩΓΗ	8
2. ΒΙΟΑΚΟΥΣΤΙΚΗ.....	12
2.1. ΕΙΣΑΓΩΓΗ	12
2.2. ΠΑΡΑΓΩΓΗ ΗΧΩΝ ΤΩΝ ENTOMΩΝ	13
2.2.1.Τριβή	14
2.2.2.Κρούση	14
2.2.3.Δόνηση	14
2.2.4.Μηχανισμός ταλάντωσης.....	15
2.2.5.Εξαγωγή αέρα	15
3. ΚΟΚΚΙΝΟΣ ΡΥΓΧΩΤΟΣ ΚΑΝΘΑΡΟΣ ΤΩΝ ΦΟΙΝΙΚΟΕΙΔΩΝ	16
3.1. ΕΙΣΑΓΩΓΗ	16
3.1.1.Μορφολογία και κύκλος ζωής εντόμου.....	18
3.1.2.Οικολογία	20
3.2. ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ ENTOMΟΥ	21
3.2.1.Καλλιεργητικοί μηχανισμοί - Μηχανικές μέθοδοι.....	21
4. ΑΝΙΧΝΕΥΣΗ ΚΑΙ ΚΑΤΑΓΡΑΦΗ	25
4.1. ΕΙΣΑΓΩΓΗ	25
4.2. ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ	27
4.2.1.Υπερκατευθυντικά μικρόφωνα (Shotgun)	28
4.2.2.Πιεζοηλεκτρικά μικρόφωνα – Αισθητήρες	29
4.2.3.Επιταχυνσιόμετρο (Accelerometer).....	30
4.3. ΔΙΑΧΕΙΡΙΣΗ ΔΕΔΟΜΕΝΩΝ.....	30
4.4. ΑΝΙΧΝΕΥΣΗ ΚΑΙ ΚΑΤΑΓΡΑΦΗ ΤΟΥ ΚΟΚΚΙΝΟΥ ΡΥΓΧΩΤΟΥ ΚΑΝΘΑΡΟΥ ΤΩΝ ΦΟΙΝΙΚΟΕΙΔΩΝ (RED PALM WEEVIL)	31
4.5. ΣΥΣΚΕΥΕΣ ΑΝΙΧΝΕΥΣΗΣ.....	31
Συσκευή ακουστικής ανίχνευσης AED 2000.....	31
4.5.1.Εισαγωγή	31
4.5.2.Χαρακτηριστικά συσκευής.....	32
4.5.3.Χαρακτηριστικά αισθητήρα	35
Συσκευή ακουστικής ανίχνευσης Laar WD 60 Pro CSC.....	36
4.5.4.Εισαγωγή	36
4.5.5.Χαρακτηριστικά συσκευής.....	37
Σύστημα ακουστικής καταγραφής Avisoft Bioacoustics UltraSoundGate	38
4.5.6.Εισαγωγή	38
4.5.7.Χαρακτηριστικά συσκευής.....	38

5. ΗΧΟΓΡΑΦΗΣΗ ΚΑΙ ΑΝΑΛΥΣΗ ΦΑΣΜΑΤΟΣ ΤΟΥ ΚΟΚΚΙΝΟΥ ΡΥΓΧΩΤΟΥ ΚΑΝΘΑΡΟΥ ΤΩΝ ΦΟΙΝΙΚΟΕΙΔΩΝ	41
5.1. ΗΧΟΓΡΑΦΗΣΗ	41
5.2. ΑΝΑΛΥΣΗ ΗΧΗΤΙΚΗΣ ΕΚΠΟΜΠΗΣ	42
5.2.1.Φασματογράφημα (spectrogram) και τρισδιάστατη ανάλυση συχνοτήτων.	42
5.2.2.Ανάλυση φασματογραφημάτων του κόκκινου ρυγχωτού κανθάρου των φοινικοειδών (R.P.W)	44
6. ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΙΑΣ.....	48
6.1. ΒΙΟΜΕΤΡΙΚΗ.....	48
6.2. ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΙΑΣ	49
6.3. ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΗΤΗ ΚΑΘΟΡΙΣΜΕΝΟΥ ΚΕΙΜΕΝΟΥ ΚΑΙ ΑΚΑΘΟΡΙΣΤΟΥ ΚΕΙΜΕΝΟΥ.....	50
6.3.1.Καθορισμένο κείμενο.....	50
6.3.2.Ακαθόριστο κείμενο	50
6.4. INTRASPEAKER ΚΑΙ INTERSPEAKER ΠΟΙΚΙΛΟΜΟΡΦΙΑ	51
6.5. ΠΗΓΕΣ ΣΦΑΛΜΑΤΩΝ ΕΠΙΒΕΒΑΙΩΣΗΣ	51
6.6. ΣΧΕΔΙΑΣΜΟΣ ΣΥΣΤΗΜΑΤΟΣ ΑΥΤΟΜΑΤΗΣ ΑΝΑΓΝΩΡΙΣΗΣ ΟΜΙΛΙΑΣ	52
6.7. ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	53
6.7.1.Η διαδικασία κβαντισμού των παραμέτρων	54
7. ΜΕΘΟΔΟΙ ΕΞΑΓΩΓΗΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΚΑΙ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ	55
7.1. ΣΤΑΔΙΑ ΕΞΑΓΩΓΗΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ	55
7.1.1.Προέμφαση.....	55
7.1.2.Εφαρμογή παραθύρου (Windowing)	56
7.2. ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ.....	57
7.2.1.Συντελεστές γραμμικής πρόβλεψης - Linear Prediction Coefficients (LPC)	57
7.2.2.Συντελεστές CEPSTRUM	60
7.2.3.Mel – Frequency Cepstral Coefficients (MFCC)	62
7.3. ΜΕΘΟΔΟΙ ΑΝΤΙΣΤΟΙΧΙΣΗΣ ΠΡΟΤΥΠΩΝ – ΜΟΝΤΕΛΟΠΟΙΗΣΗ	63
7.3.1.Δυναμική αναδίπλωση στο χρόνο. Dynamic Time Warping (DTW)	63
7.3.2.Διανυσματικός κβαντισμός - Vector Quantization (VQ) ..	63
7.3.3.Μοντέλα γκαουσιανών μειγμάτων - Gaussian Mixture Models (GMM)	69
7.3.4.Κρυφά μαρκοβιανά μοντέλα - Hidden Markov Models (HMM)	77
8. ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ.....	84

8.1. ΕΙΣΑΓΩΓΗ	84
8.2. ΠΕΙΡΑΜΑΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ	85
8.2.1.Αναγνώριση εντόμου.....	85
8.2.2.Ταυτοποίηση εντόμου.....	87
8.3. ΠΑΡΟΥΣΙΑΣΗ ΚΑΙ ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	87
8.3.1.Πείραμα 1.1 Αναγνώριση του R.P.W. από πλήθος ακμαίων εντόμων	87
8.3.2.Πείραμα 1.2 Αναγνώριση του R.P.W. από πλήθος προνυμφών.....	89
8.3.3.Πειράματα αναγνώρισης του R.P.W. με μεταβλητή διάρκεια δειγμάτων 2.1, 2.2	90
8.3.4.Πείραμα 3.1 Ταυτοποίηση του R.P.W. από πλήθος ακμαίων εντόμων	91
8.3.5.Πείραμα 3.2 Ταυτοποίηση του R.P.W. από πλήθος προνυμφών.....	91
9. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ	93
ΠΑΡΑΡΤΗΜΑ Α.....	95
ΠΑΡΑΡΤΗΜΑ Β.....	98
ΠΑΡΑΡΤΗΜΑ Γ	109
ΕΙΚΟΝΟΓΡΑΦΙΑ	111
ΒΙΒΛΙΟΓΡΑΦΙΑ	115

1 ΚΕΦΑΛΑΙΟ

1. ΕΙΣΑΓΩΓΗ

Πάνω στον πλανήτη μας, έχουν εντοπιστεί και καταγραφεί πάνω από 900.000 είδη εντόμων και εικάζεται ότι υπάρχει δεκαπλάσιος αριθμός που δεν έχει, ακόμα, καταγραφεί. Αυτό έχει σαν αποτέλεσμα, τα έντομα, να αποτελούν το μεγαλύτερο και πολυπληθέστερο οικοσύνολο στη γη.

Το ζήτημα της αντιμετώπισης των παρασίτων αποτελεί ένα μεγάλο και σημαντικό κεφάλαιο της επιστήμης που αποκαλούμε Βιολογία. Αφορά δε και επηρεάζει άμεσα ζωτικούς τομείς ενασχόλησης των ανθρώπων ενός τόπου, όπως είναι η γεωργία, η ανθοκομία και η κηπευτική. Το πρόβλημα παρουσιάζεται κυρίως σε φυτά, άνθη, δέντρα αλλά και τους καρπούς αυτών, σε μορφή αποθηκευμένων προϊόντων, όπου ορισμένα έντομα μπορούν να βρουν καταφύγιο, να τραφούν και να αναπαραχθούν, βλάπτοντας, όμως, ανεπανόρθωτα τους οργανισμούς στους οποίους παρασιτούν.

Επιστήμονες, ανά τον κόσμο, που έχουν ασχοληθεί με το συγκεκριμένο πρόβλημα, έχουν καταλήξει στο συμπέρασμα ότι η ανίχνευση και η παρακολούθηση αυτών των επιβλαβών εντόμων είναι σημαντική για ένα ολοκληρωμένο σύστημα διαχείρισης παρασίτων. Η χρήση υποκατάστατων μεθυλοβρομίου για την καταπολέμηση των εντόμων, στηρίζεται στην γνώση του κατάλληλου χρόνου και του κατάλληλου τόπου της εφαρμογής τους μέσα και γύρω από αποθηκευμένα προϊόντα ή φυτά, όπως επίσης και στην ύπαρξη ενός τρόπου υπολογισμού της αποτελεσματικότητά τους. Έτσι, λοιπόν, η χωροταξική κατανομή των πληθυσμών μπορεί να ελαχιστοποιήσει την χρήση των μεθόδων ελέγχου εντόμων που βασίζονται σε πολυδάπανες αναλύσεις.¹

Τα έντομα αναγνωρίζονται κυρίως από το σχήμα τους και από τον ήχο που εκπέμπουν, τα οποία αποτελούν βασικά γνωρίσματα για κάθε είδος. Η ανίχνευση και η αναγνώριση του είδους του κάθε εντόμου, συνήθως, γίνεται χειρωνακτικά, με τη χρήση μεθόδων παγίδευσης και παρακολούθησης. Η τοποθέτηση παγίδων μπορεί να παρέχει πληροφορίες σχετικά με την χωροταξική κατανομή πληθυσμών σε μια συγκεκριμένη περιοχή. Παρόλα αυτά η χρήση των προ-

¹ Hagstrum, D. W., P. W. Flinn, and D. Shuman. (1996). Automated Monitoring Using Acoustical Sensors for Insects in Farm-Stored Wheat. *J. Econ. Entomol.* 89(1):211-217.

αναφερθέντων μεθόδων έχει περιοριστεί εξ αιτίας του χρόνου και της προσπάθειας που απαιτείται για τη προσέγγιση των παγίδων, την χειρωνακτική επιθεώρηση των περιεχομένων τους και την δημιουργία μιας ακριβής βάσης δεδομένων χωροταξικής ανάλυσης.

Η αξιοσημείωτη πρόοδος που έχει σημειωθεί τα τελευταία χρόνια στην τεχνολογία των ηλεκτρονικών υπολογιστών, καθώς και στις τεχνικές επεξεργασίας σήματος και αναγνώρισης προτύπων, έχει δημιουργήσει τις κατάλληλες συνθήκες για τη δημιουργία ενός αυτόματου, ηλεκτρονικού συστήματος ανίχνευσης και αναγνώρισης ειδών. Η βασική ιδέα της συγκεκριμένης εφαρμογής, βασίζεται στο γεγονός ότι η ακουστική εκπομπή ενός εντόμου συνίσταται από επαναλαμβανόμενα πρότυπα (μοτίβα), τα οποία είναι μοναδικά για κάθε είδος.

Εκμεταλλευόμενοι τις αρχές και τις μεθόδους των συστημάτων αναγνώρισης ομιλίας και επιβεβαίωσης ομιλητή, στην παρούσα εργασία, επιχειρούμε να αναπτύξουμε ένα σύστημα ανίχνευσης και επιβεβαίωσης βιοακουστικού σήματος. Ο στόχος ενός τέτοιου συστήματος, στην περίπτωση μας, είναι διπλός. Αφ' ενός ζητούμε να ανιχνεύεται η ύπαρξη ή μη του ρυγχωτού κανθάρου, μέσα στο φοίνικα, αφ' εταίρου, να αναγνωρίζουμε το έντομο ανάμεσα σε άλλα, λιγότερο επιβλαβή, που πιθανόν να παρασιτούν στην ίδια περιοχή.

Εν κατακλείδι, θα μπορούσαμε να αναφέρουμε τον βραχυπρόθεσμο στόχο της εργασίας μας, ο οποίος σχετίζεται με την συνδρομή μας όσον αφορά την ανάπτυξη και διεκπεραίωση ενός ολοκληρωμένου, αυτόματου συστήματος ανίχνευσης παρασίτων και συγκεκριμένα του κόκκινου ρυγχωτού κανθάρου των φοινικοειδών. Ο σκοπός αυτής της προσπάθειας είναι η εξάλειψη του ανθρώπινου παράγοντα από την διαδικασία παρακολούθησης και εντοπισμού των εντόμων. Με αυτόν τον τρόπο ελαχιστοποιείται το ανθρώπινο σφάλμα που προκύπτει από την εφαρμογή εμπειρικών μεθόδων στην προαναφερθείσα διαδικασία.

Η οργάνωση της εργασίας

ΚΕΦΑΛΑΙΟ 2:

Το κεφάλαιο αυτό έχει σαν επικεφαλίδα «Βιοακουστική» και χωρίζεται σε δύο μέρη. Το πρώτο μέρος είναι εισαγωγικό και παρατίθεται ο ορισμός της βιοακουστικής καθώς και το αντικείμενο που πραγματεύεται και μελετά. Επίσης συμπεριλαμβάνεται και ο σκοπός αυτού του κλάδου της επιστήμης της βιολογίας. Στο δεύτερο μέρος αναλύεται ο τρόπος με τον οποίο συμπεριφέρονται τα έντομα στις βιολογικές τους ανάγκες και πώς αυτές συνδέονται με τους ήχους που παράγουν.

ΚΕΦΑΛΑΙΟ 3:

Ο τίτλος του συγκεκριμένου κεφαλαίου είναι «κόκκινος ρυγχωτός κάνθαρος των φοινικοειδών R.P.W.»

Εδώ καταγράφονται επιγραμματικά, τα έντομα που παρασιτούν στους φοίνικες και είναι οι σημαντικότεροι εχθροί αυτών. Όσον αφορά τον RED PALM WEEVIL γίνεται μια εκτενής αναφορά στις περιοχές που έχει βρεθεί – προσβάλει, στα είδη των φοινικοειδών που μπορεί να προσβάλει όπως και σε αυτά που έχει παρατηρηθεί ότι είναι ανθεκτικά απέναντι σε αυτόν. Περιγράφεται, αναλυτικά, ο κύκλος ζωής του, οι συνήθειές του και η συμπεριφορά του κατά την διάρκεια της ζωής του και, τέλος, παραθέτονται τρόποι ανίχνευσης, αντιμετώπισής και καταπολέμησης του εντόμου.

ΚΕΦΑΛΑΙΟ 4:

Το περιεχόμενο του 4^{ου} κεφαλαίου σχετίζεται με την ανίχνευση και την καταγραφή των εντόμων από τις ακουστικές τους εκπομπές.

Χωρίζεται σε πέντε μέρη, εκ των οποίων, στο πρώτο γίνεται αναφορά στην φυσιολογία των εντόμων, τους χώρους στους οποίους μπορούν να ανιχνευθούν και με τι είδους εξοπλισμό (αισθητήρες, μικρόφωνα) μπορεί να γίνει η ανίχνευση και η καταγραφή τους.

Στο δεύτερο μέρος καταγράφονται οι σημαντικότερες και πιο γνωστές μέθοδοι ανίχνευσης και ηχογράφησης.

Στο τρίτο και τέταρτο μέρος, περιγράφονται τα μέσα καταγραφής και διαχείρισης των δεδομένων που έχουν καταγραφεί και πώς μπορεί να γίνει, εν τέλει, η ανίχνευση και η καταγραφή του Red Palm Weevil με βάση τα δεδομένα του πρώτου μέρους του κεφαλαίου και της φυσιολογίας του εντόμου που εξετάζουμε.

Τέλος, στο πέμπτο μέρος παραθέτονται οι λειτουργίες και τα τεχνικά χαρακτηριστικά τριών συσκευών, ανίχνευσης και καταγραφής, που, κατά την γνώμη μας, καλύπτουν περισσότερο τις απαιτήσεις όσον αφορά την ανίχνευση του R.P.W.

ΚΕΦΑΛΑΙΟ 5:

Το κεφάλαιο αυτό, το οποίο χωρίζεται σε δύο μέρη, σχετίζεται με την ηχογράφηση του κόκκινου ρυγχωτού κανθάρου και την ανάλυση του φάσματος εκπομπής του.

Στο πρώτο μέρος καταγράφεται, αναλυτικά, ο τρόπος με τον οποίο γίνεται η ανίχνευση και, εν τέλει, η ηχογράφηση με την χρήση πιεζοηλεκτρικών μικροφώνων, αισθητήρων ή επιταχυνσιομέτρου.

Στο επόμενο μέρος παρουσιάζουμε την πολύ σημαντική ανάλυση της ηχητικής του εκπομπής, τον τρόπο με τον οποίο παρήχθη, ποια είναι τα αποτελέσματά της και ποιοι είναι, τελικά, οι ήχοι που παράγει το συγκεκριμένο είδος, οι οποίοι το χαρακτηρίζουν, το κάνουν αναγνωρίσιμο και του δίνουν την ιδιότητα να ξεχωρίζει από τα άλλα είδη της οικογενείας του.

ΚΕΦΑΛΑΙΟ 6:

Στο κεφάλαιο 6 γίνεται μία εισαγωγή στην αναγνώριση ομιλίας. Εξηγείται ο ορισμός της βιομετρικής, τι είναι η αναγνώριση ομιλίας, πού χρησιμοποιείται και σε ποιες κατηγορίες χωρίζεται. Επίσης, γίνεται μια εκτεταμένη αναφορά στην αναγνώριση ομιλητή καθορισμένου και ακαθόριστου κειμένου όπως και στα πλεονεκτήματα και μειονεκτήματα αυτών. Αναλύονται οι ορισμοί της *intraspeaker* και της *interspeaker* ποικιλομορφίας καθώς και ποιες μπορεί να είναι οι αιτίες που μπορούν να φέρουν εσφαλμένα αποτελέσματα στην αναγνώριση. Τέλος, περιγράφονται οι φάσεις για τον σχεδιασμό ενός συστήματος αναγνώρισης ομιλίας και αναλύεται ο *front end* μηχανισμός, τα στάδια του και οι λειτουργία του μέσα στο σύστημα.

ΚΕΦΑΛΑΙΟ 7:

Εδώ παρουσιάζονται, αναλυτικά, οι μέθοδοι εξαγωγής χαρακτηριστικών και μοντελοποίησης.

Αναλύονται, επίσης, τα στάδια της προέμφασης, εφαρμογής παραθύρου και οι τεχνικές εξαγωγής χαρακτηριστικών (*LPC*, συντελεστές *cepstrum*, *MEL cepstrum*). Εν κατακλείδι, παρουσιάζονται οι μέθοδοι αντιστοίχισης προτύπων και μέθοδοι μοντελοποίησης όπως *DTW*, *VQ*, *HMM* και *GMM*.

2 ΚΕΦΑΛΑΙΟ

2. ΒΙΟΑΚΟΥΣΤΙΚΗ

2.1. ΕΙΣΑΓΩΓΗ

Η Βιοακουστική είναι ένας από τους πιο σύγχρονους κλάδους της Βιολογίας. Δανειζόμενη μεθόδους της φυσικής μελετά τους ήχους που παράγουν οι ζωικοί οργανισμοί. Η κατανόηση του τρόπου παραγωγής τους, των χαρακτηριστικών και κυρίως της βιολογικής σημασίας τους είναι, επίσης, τα βασικά ερωτήματα που απασχολούν την Βιοακουστική.²

Τα ζώα, σε ολόκληρο τον πλανήτη, επικοινωνούν μεταξύ τους εκπέμποντας ήχους – φωνήσεις. Χερσαία ζώα, αμφίβια, υδρόβια, πουλιά, κητοειδή και έντομα παράγουν ήχους που χαρακτηρίζουν το είδος τους. Αυτή η πληθώρα των ήχων παρέχει από μόνη της μία πηγή πληροφόρησης, η οποία μπορεί να χρησιμοποιηθεί για να εξερευνηθεί η πολυμορφία των φωνήσεων και της επικοινωνίας μεταξύ των εκάστοτε ειδών.

Αυτή την δραστηριότητα των ειδών εκμεταλλεύεται η Βιοακουστική, με σκοπό τη μελέτη της συμπεριφοράς των ζώων και γενικότερα την παρακολούθηση των ειδών που απειλούνται με εξαφάνιση, καθώς και την παρακολούθηση παρασιτικών εντόμων που προκαλούν σοβαρές καταστροφές σε καλλιέργειες και προϊόντα και άλλα. Επίσης η βιοακουστική μπορεί να είναι ένα πολύ χρήσιμο όργανο όσον αφορά την περιβαλλοντική εκπαίδευση.

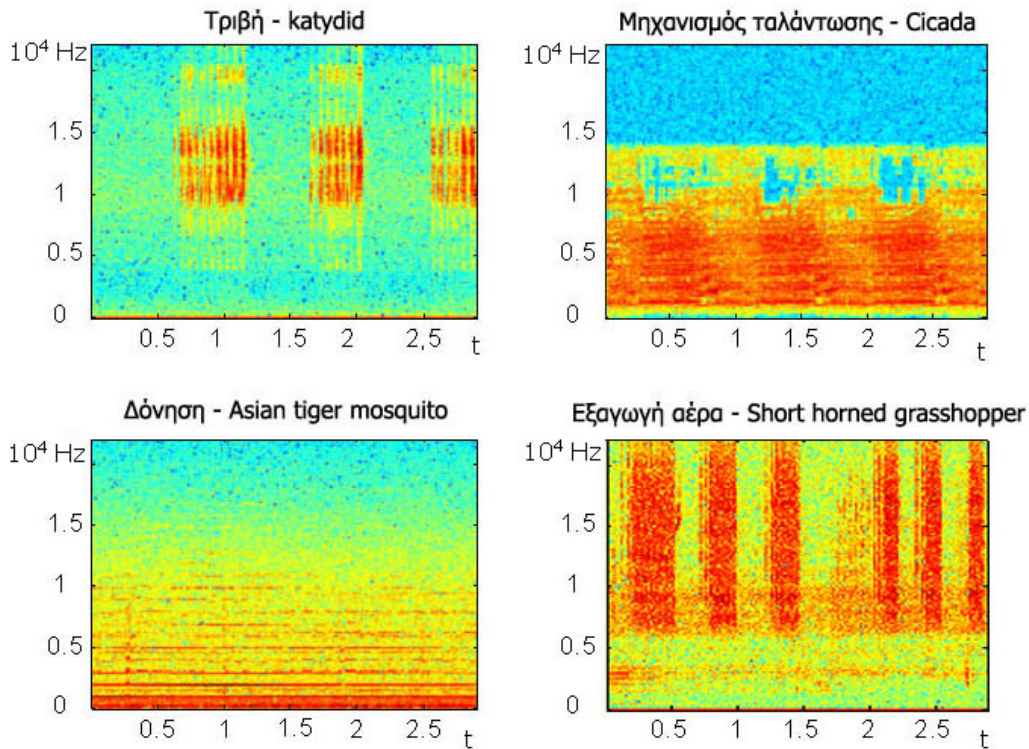
Η αποτελεσματική εφαρμογή της βιοακουστικής είναι άρρηκτα συνδεδεμένη με την τεχνολογική ανάπτυξη διότι αυτή η επιστήμη απαιτεί την χρήση πολύπλοκων καταγραφέων ήχου (sound recorders), μικροφώνων και υπολογιστικών προγραμμάτων καλής ποιότητας.

Οι βιοακουστικές μελέτες έχουν εφαρμοσθεί σε πολλά είδη ζώων όπως τα πουλιά, τα αμφίβια, τα θηλαστικά, τα έντομα και τα ερπετά. Ενδεικτικά αναφέρουμε ότι από την ηχητική εκπομπή των νυχτερί-

² <http://www.pelagosinstitute.gr/gr/research/dinamiko.html>

δων ή των αρουραίων, γίνεται πιθανός ο εντοπισμός τους, δεδομένου ότι τα είδη αυτά ζουν από τη φύση τους στο σκοτάδι. Η μέθοδος αυτή ονομάζεται ηχοεντοπισμός (echo localization). Αυτός ο τρόπος εντοπισμού έχει επίσης εφαρμογή στα κητοειδή όσον αφορά την αναγνώριση και κατανόηση του πολυσύνθετου τρόπου επικοινωνίας τους. Σχετικά με τα πουλιά, εξαιτίας του ποικίλου τρόπου φώνησης που χρησιμοποιούν, η βιοακουστική θέτει τα θεμέλια για την εκ βάθρων γνώση των δραστηριοτήτων, της συμπεριφοράς και της επικοινωνίας τους.

Η γραφική αναπαράσταση του ήχου που παράγει το κάθε είδος, με ειδικά προγράμματα, αποτελεί ένα ισχυρό εργαλείο για την ταυτοποίηση των ειδών εξαιτίας του γεγονότος ότι η διαφορετικότητα των ακουστικών εκπομπών των διαφόρων ειδών, η οποία δε γίνεται αντιληπτή από το ανθρώπινο αυτί, γίνεται αντιληπτή, με αυτόν τον τρόπο.



Σχ 2.1 Φασματογραφήματα φωνήσεων εντόμων

2.2. ΠΑΡΑΓΩΓΗ ΗΧΩΝ ΤΩΝ ΕΝΤΟΜΩΝ

Έχει παρατηρηθεί ένας συγκεκριμένος αριθμός τρόπων συμπεριφοράς οι οποίοι συνδέονται με την παραγωγή ήχων στα έντομα.

Ο πρώτος τρόπος περιλαμβάνει εκείνες τις περιπτώσεις όπου τα έντομα παράγουν ήχους με σκοπό να προσεγγίσουν τα θηλυκά κοινά τους ή να ζητήσουν από τα θηλυκά να δώσουν το στίγμα τους

προς εντοπισμό από την ακουστική τους εκπομπή ή να καλέσουν σε σύναξη τα υπόλοιπα έντομα.

Ο δεύτερος τρόπος συμπεριφοράς αποτελείται από εκείνες τις περιπτώσεις όπου

α) Ο ήχος παράγεται ως αντίδραση στην παρουσία ή στην δραστηριότητα άλλων οργανισμών. Συγκεκριμένα, αρσενικά, θηλυκά ή μη ώριμα έντομα, παράγουν ακουστικές εκπομπές προκειμένου να δηλώσουν την ενόχλησή τους ή να προειδοποιήσουν άλλα έντομα για κίνδυνο

β) Ένα αρσενικό έντομο μπορεί να χρησιμοποιήσει το τραγούδι του προκειμένου να ενημερώσει τα υπόλοιπα αρσενικά ότι έχει την περιοχή υπό την επίβλεψη του

γ) Ένα θηλυκό παράγει ακουστική εκπομπή κατά την παρουσία ενός αρσενικού, του ίδιου είδους.

Έκτος από την παραγωγή ήχου ως μέσο επικοινωνίας, ο ήχος μπορεί να παράγεται μη επιτηδευμένα ως αποτέλεσμα της σίτισης του εντόμου, της πτήσης του ή της μετακίνησής του. Ο μηχανισμός παραγωγής του ήχου στα έντομα μπορεί να συμπυχθεί ως εξής: Η συστολή των μυών οδηγεί σε δόνηση του μηχανισμού παραγωγής ήχου και τελικά στην εξαγωγή του ήχου και στην εκπομπή του. Οι ήχοι παράγονται από τα έντομα με πέντε διαφορετικούς τρόπους (σχήμα 2.1):

2.2.1. Τριβή

Η τριβή ανάμεσα σε δύο μέρη του σώματος, συνήθως ακούγεται σαν τερέτισμα π.χ. (γρύλος, ακρίδα, κατσαρίδα, σκαθάρι, σκώρος, πεταλούδα, μυρμήγκι, προνύμφη σκαθαριού κ.α.)³

2.2.2. Κρούση

Χτυπώντας κάποιο σημείο του σώματος όπως είναι τα πόδια (ακρίδα band – winged), η άκρη της κοιλιάς (κατσαρίδα) ή το κεφάλι (σκαθάρι death – watch), ακούγεται συνήθως σαν ένας συνεχής χτύπος.³

2.2.3. Δόνηση

Η ταλάντωση μερών του σώματος όπως τα φτερά που συνήθως ακούγεται σαν βόμβος ή υπόκωφο βουητό δονώντας κάποια μέρη του σώματος στον αέρα. (μέλισσες, σφήκες, κουνούπια, μύγες, κ.λ.π.).³

³ Todor Ganchev, Ilyas Potamitis, Nikos Fakotakis “*Automatic Acoustic Identification of Insects using Fusion of Classifiers: The case of Crickets and Cicadas*” §2 Acoustic Behavior of Insects

2.2.4.Μηχανισμός ταλάντωσης

Η γρήγορη συστολή και διαστολή των δονούμενων μυών (δονούμενες μεμβράνες που μοιάζουν με τύμπανα). Συνήθως ακούγεται σαν ένας συνεχής ήχος που αποτελείται από "clicks" π.χ. (τζιτζίκια, leafhopper, treehopper, spittlebug)³

2.2.5.Εξαγωγή αέρα

Η διαδικασία κατά την οποία το έντομο εξάγει αέρα ή υγρό από το σώμα του. Συνήθως ακούγεται σαν σφύριγμα π.χ.(ακρίδες short - horned).³

3 ΚΕΦΑΛΑΙΟ

3. ΚΟΚΚΙΝΟΣ ΡΥΓΧΩΤΟΣ ΚΑΝΘΑΡΟΣ ΤΩΝ ΦΟΙΝΙΚΟΕΙΔΩΝ

RED PALM WEEVIL - *Rhynchophorus ferrugineus*

3.1. ΕΙΣΑΓΩΓΗ

Υπάρχουν πολλοί σοβαροί εχθροί των φοινικοειδών, οι περισσότεροι των οποίων δεν έχουν παρατηρηθεί στην Ελλάδα π.χ. *Paysandisia archon* (Lepidoptera: Castniidae), *Trips palmi* (Thysanoptera: Thripidae). Αρκετά είδη κολεόπτρων της οικογενείας Curculionidae αποτελούν, επίσης, σοβαρούς εχθρούς των φοινικοειδών. Ως πιο επικίνδυνα αναφέρονται τα *Rhynchophorus ferrugineus*, *R. palmarum*, ακολούθως τα *R. Pheonicis*, *R. Vunleratus* και *R. Bilineatus*, ενώ, μικρότερης οικονομικής σημασίας, αναφέρεται ότι είναι τα *Dynamis borassii*, *Rhynchophorus quadrangulus* και *Metamasius cinnamominus*. Ειδικά το *Rhynchophorus ferrugineus* (κόκκινος ρυγχωτός κάνθαρος των φοινικοειδών, red palm weevil, coconut weevil, Asiatic palm weevil, Indian red palm weevil) θεωρείται σήμερα ως ο σημαντικότερος εχθρός των φοινικοειδών.⁴



Εικ. 3.1 Red Palm Weevil

Ο κόκκινος ρυγχωτός κάνθαρος καταγράφηκε για πρώτη φορά το 1891 στην Ινδία. Επίσης καταγράφηκε ως κυριότερος εχθρός της ινδικής καρύδας (κοκκοφοίνικα) το 1906, ενώ το 1917 περιγράφηκε ως το πιο επικίνδυνο έντομο τις χουρμαδιάς στο Punjab της Ινδίας. Το 1918 αναφέρεται ότι προκάλεσε σοβαρές καταστροφές στις χουρμαδιές της Μεσοποταμίας (Ιράκ) αλλά δεν συλλέχτηκε κανένα αντιπροσωπευτικό δείγμα που να το επιβεβαιώνει. Στα μέσα του 1980 βρέθηκε στις χώρες του Αραβικού κόλπου. Θεωρείται το πιο επιβλα-

⁴ Κοντοδήμας, Δ. - Οικονόμου, Δ. - Θυμάκης, Ν. - Χ.Μεντή και Ανάγνου – Βερονίκη, Μ. «*Νέος σοβαρός εχθρός των φοινικοειδών, το κολεόπτερο Rhynchophorus Ferrugineus (Olivier) (Curculionidae: Dryophorinae)*», Γεωργία και κτηνοτροφία, τεύχος 1, 54-58, 2006

βές έντομο των χουρμάδων της Μέσης Ανατολής. Αναλυτικότερα έχει καταγραφεί σε χώρες της Ωκεανίας (Αυστραλία, Νέα Παπούα - Γουινέα, νήσοι του Σολομώντος) και της Ασίας (Μπαγκλαντές, Μπαχρέιν, Καμπότζη, Κίνα, Ινδία, Ινδονησία, Ιράκ, Ιράν, Ιαπωνία, Κουβέιτ, Λάος Μαλαισία, Μυανμαρ, Ομάν, Πακιστάν, Φιλιππίνες, Κατάρ, Σαουδική Αραβία, Ιορδανία, Σρι Λάνκα, Ταϊβάν, Ταϊλάνδη, Ηνωμένα Αραβικά Εμιράτα, Βιετνάμ) και πρόσφατα έχει βρεθεί και σε Μεσογειακές χώρες (Ισπανία, Παλαιστινιακή αρχή, Αίγυπτος, Τουρκία, Ελλάδα).

Η είσοδος του *Rhynchophorus ferrugineus* στον ελλαδικό χώρο αποτελεί σοβαρό πρόβλημα. Στην Ελλάδα χρησιμοποιούνται περίπου 30 είδη φοινικοειδών για φυτεύσεις, σε εξωτερικούς χώρους, τα περισσότερα εκ των οποίων αποτελούν γνωστούς ξενιστές του *Rhynchophorus ferrugineus*. Για ορισμένα είδη όπως π.χ. το *Syagrus romanzoffiana* το οποίο χρησιμοποιείται τα τελευταία χρόνια εκτεταμένα στην κηποτεχνία, δεν υπάρχουν στοιχεία.



Εικ. 3.2 Η εξάπλωση του R.P.W.

Επίσης ιδιαίτερως ανησυχητική είναι η ανεύρεση του συγκεκριμένου εντόμου στην Κρήτη όπου βρίσκεται το φοινικόδασος στο Βάι, που είναι το τελευταίο δάσος στη γη του φοίνικα του Θεοφράστου, *Phoenix theophrasti*.

Ο *Rhynchophorus ferrugineus* βρέθηκε και στο Ηράκλειο της Κρήτης



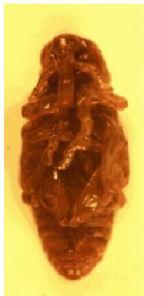
Εικ. 3.3 *Phoenix theophrasti*

το Νοέμβριο του 2005 από τον Γεωπόνο κ. Δημήτρη Οικονόμου και με την επιμέλεια του κ. Θυμάκη Νικολάου προσκομίστηκαν δείγματα στο Μπενάκειο Φυτοπαθολογικό Ινστιτούτο. Από εκεί στάλθηκαν στο Μουσείο Φυσικής Ιστορίας του Λονδίνου όπου επιβεβαιώθηκε η ταυτότητα του εντόμου. Αυτή τη στιγμή δεν υπάρχουν περισσότερα στοιχεία για την κατάσταση των προσβολών.



Εικ. 3.4 Ο κύκλος ζωής του R.P.W

3.1.1. Μορφολογία και κύκλος ζωής εντόμου⁵



Εικ. 3.5
Νύμφη

Όλα τα στάδια του εντόμου (ώο, προνύμφη, νύμφη, ενήλικο) εξελίσσονται στο εσωτερικό του φοίνικα και ο βιολογικός κύκλος δεν μπορεί να ολοκληρωθεί πουθενά αλλού.

Τα θηλυκά γεννούν γύρω στα 300 αυγά σε ξεχωριστές οπές ή πληγές του φοίνικα. Τα αυγά του *Rhynchophorus ferrugineus* έχουν λευκό, κρεμ, χρώμα και ωοειδές σχήμα και το μήκος τους φτάνει το 1.1mm.

Οι προνύμφες (άποδες) εκκολάπτονται μέσα σε διάστημα 2-3-ημερών και ανοίγουν οπές στο εσωτερικό των φοινίκων. Μετακινούνται από περισταλτικές μυϊκές συσπάσεις του σώματος και τρέφονται από μαλακούς χυμώδεις ιστούς, απορρίπτοντας έτσι όλα τα ινώδη υλικά του φυτού. Το μήκος, της πλήρως αυξημένης προνύμφης είναι 50mm και το πλάτος της 20mm.

Η περίοδος που το έντομο βρίσκεται στο στάδιο της προνύμφης ποικίλει από 1- 3 μήνες. Κάθε προνύμφη του *Rhynchophorus ferrugineus* κατασκευάζει μια νυμφική θήκη με τις ίνες του φοίνικα, σχήματος οβάλ, με μήκος 60mm και πλάτος 30mm. Η νύμφωση διαρκεί 14 - 21 ημέρες.



Εικ. 3.6 Τα αυγά του R.P.W.



Εικ. 3.7 Προνύμφες του R.P.W

⁵ <http://www.redpalmweevil.com/life.htm>

Το ενήλικο είναι ένας κοκκινωπός, καφέ κύλινδρος με ένα προεξέχον καμπυλόγραμμο ράμφος-ρύγχος. Ποικίλει αρκετά στο μέγεθος και έχει περίπου 35mm μήκος και 12mm πλάτος. Το κεφάλι και το ρύγχος του αποτελούν το 1/3 περίπου του συνολικού του μήκους.

Το ρύγχος του φέρει ένα μικρό ζεύγος σιαγώνων, στο τέλος του, και ένα ζεύγος κεραιών κοντά στην βάση του. Το χρώμα του, από πάνω (κυρτή επιφάνεια), είναι κόκκινο καφέ, ενώ από την κάτω πλευρά (κοίλη επιφάνεια) είναι καφέ σκούρο. Το ρύγχος του αρσενικού στη άκρη του καλύπτεται από κοντά, κιτρινοκαφέ τριχίδια, σε αντίθεση με το θηλυκο, του οποίου το ρύγχος είναι γυμνό, πιο λεπτό, περισσότερο καμπυλωτό και ελαφρώς μακρύτερο.

Τα μάτια του *R. Ferrugineus* είναι μικρά και χωρίζονται στην βάση του ρύγχους.

Το μεσαίο μέρος του εντόμου έχει κοκκινωπό χρώμα και φέρει μερικές μαύρες βούλες οι οποίες διαφέρουν σε σχήμα, μέγεθος και αριθμό.

Το έλυτρο, το λεπτό κεράτινο κάλυμμα των φτερών, έχει κόκκινο καφέ χρώμα και φέρει επιμήκειες μαύρες ραβδώσεις.



Εικ. 3.8 Το κουκούλι του R.P.W



Εικ. 3.9



Εικ. 3.10 Το κεφάλι



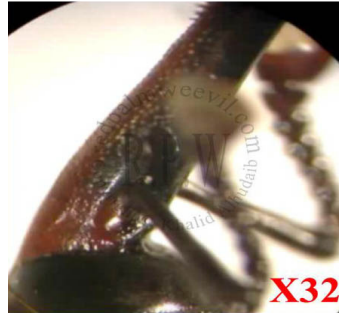
Εικ. 3.11 Ο θώρακας



Εικ. 3.12 Το ρύγχος του αρσενικού



Εικ. 3.13 Το ρύγχος του αρσενικού



Εικ. 3.14 Οι κεραίες



Εικ. 3.15 Οι κεραίες

3.1.2. Οικολογία⁶

Το *Rhynchophorus ferrugineus* προσβάλλει πολλά είδη φοινικοειδών όπως *Areca catechu*, *Arenga engleri*, *Caryota urens*, *Cocos nucifera*, *Corypha elata*, *Elaeis guineensis*, *Livistona desipiens*, *L. Chinensis*, *L. Australis*, *Phoenix canariensis*, *L. Sylvestris*, *P. Roebelenii*, *Sabal palmetto*, *S. Causiarum*, *Trachycarpus fortunei*, *Washingtonia filifera* κ.α. καθώς και τα *Agave Americana* και *Saccharum officinarum*. Τα είδη *Chamaerops humilis* και *Washingtonia robusta* αναφέρονται ως ανθεκτικά στο έντομο αυτό.

Η προσβολή ξεκινά από την κορυφή (στεφάνη) όπου τα θηλυκά ωοτοκούν. Οι προνύμφες ανοίγουν στοές έως και ένα μέτρο κατά μήκος του κορμού αλλά και κατά μήκος των φύλλων. Η προσβολή οδηγεί στο θάνατο του φοινικοειδούς.



Εικ. 3.16 Sabal Causiarum



Εικ. 3.17 Phoenix Roebelenii



Εικ. 3.18. Washingtonia filifera

⁶ http://www.eppo.org/quarantine/Albert_List/insects/rhyefe.htm



Εικ. 3.19 Livistona Australis

3.2. ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ ΕΝΤΟΜΟΥ

Φυτοϋγειονομικός έλεγχος

Σύμφωνα με την αναφορά του EPPO (European and Mediterranean plant protection organization) το έντομο διαδίδεται με την μεταφορά των φυτών (ή τμήματα των φυτών) από περιοχές που υπάρχουν προσβεβλημένα φυτά. Ως εκ τούτου, πρέπει να εφαρμόζονται αυστηρά μέτρα καραντίνας σε διεθνή και εθνικά επίπεδα.

Επίσης πρέπει να γίνονται εκτεταμένες και λεπτομερείς παρατηρήσεις σε όλες τις περιοχές που καλλιεργούνται, πωλούνται ή φυτεύονται φοινικοειδή, ιδιαίτερα μετά την διαπίστωση του εχθρού σε κάποια περιοχή.⁷

3.2.1. Καλλιεργητικοί μηχανισμοί - Μηχανικές μέθοδοι

Η καλή υγιεινή του αγρού και η καλλιεργητικές τεχνικές αποτελούν τα σημαντικότερα μέτρα για την αποτροπή της προσβολής. Συγκεκριμένα απαιτείται⁸:

⁷ http://www.eppo.org/quarantine/Albert_List/insects/rhyefe.htm

⁸ Θυμάκης Ν., Δ. Κοντοδήμας, Οικονόμου Δ., Δ. Ραπτόπουλος. «Ο εχθρός των φοινικοειδών *Rhynchophorus Ferrugineus* (Coleoptera: Curculionidae). Υφιστάμενη κατάσταση και προτάσεις για την αντιμετώπισή του.» Πνευματικό κέντρο Δήμου Χερσονήσου 31 Μάιου 2006



Εικ. 3.20 Φοινικές που έχουν προσβληθεί από τον *Rhynchophorus ferrugineus*



- Καθαρισμός των κεφαλών των φοινικοειδών περιοδικά ώστε να εμποδίζεται η αποσάθρωση των οργανικών υλικών στις μασχάλες των φύλλων του φυτού.
- Αποφυγή κλαδεμάτων-πληγών.
- Όταν κόβονται τα πράσινα φύλλα πρέπει η τομή να γίνεται 120cm μακριά από την βάση.
- Τα φοινικοειδή που είναι προσβεβλημένα από ασθένειες είναι περισσότερο επιρρεπή από τον *Rhynchophorus ferrugineus*. Τα φοινικοειδή αυτά πρέπει να θεραπεύονται με την χρήση κατάλληλων μυκητοκτόνων και μετά να εφαρμόζονται εντομοκτόνα ώστε να παρεμποδιστεί η εναπόθεση ωών.
- Η καταστροφή όλων των νεκρών ή βαριά προσβεβλημένων φοινικοειδών είναι το σημαντικότερο μέτρο για τον περιορισμό του εντόμου.

Βιολογική αντιμετώπιση.

Έχει παρατηρηθεί παρασιτισμός από ακάρεα. Ακόμα υπάρχει δυνατότητα για βιολογική αντιμετώπιση με εντομοπαθογόνους μικροοργανισμούς (νηματώδεις, μύκητες ή ιούς), με πιο αποτελεσματική τη χρήση νηματωδών. Επίσης έχει δοκιμαστεί πειραματικά και η εξαπόλυση στείρων αρρένων.

Χημική καταπολέμηση.

Η χημική καταπολέμηση συνίσταται σε προληπτικές επεμβάσεις με εντομοκτόνα (ιδίως μετά το κλάδεμα) ώστε να αποτραπεί η ωοτοκία και σε θεραπευτικές επεμβάσεις στα προσβεβλημένα δέντρα με έγχυση εντομοκτόνων κ.α.

Τα σημαντικότερα προβλήματα στην αντιμετώπιση του *Rhynchophorus ferrugineus* είναι η δυσκολία εντοπισμού των προσβεβλημένων φυτών. Η διαπίστωση της προσβολής γίνεται κατά κανόνα όταν αυτή έχει φθάσει σε μη αναστρέψιμο στάδιο. Προς το σκοπό της εξακρίβωσης της προσβολής σε πρώιμο στάδιο έχουν αναπτυχθεί τεχνικές και μηχανήματα για τον εντοπισμό του ήχου που παράγεται από τις προνύμφες μέσα στον κορμό ή την καταγραφή της μεταβολής της αναπνοής και διαπνοής του προσβεβλημένου φοινικοειδούς. Στο Ισραήλ χρησιμοποιήθηκαν επίσης και ειδικά εκπαιδευμένοι σκύλοι που εντοπίζουν την προσβολή.⁹

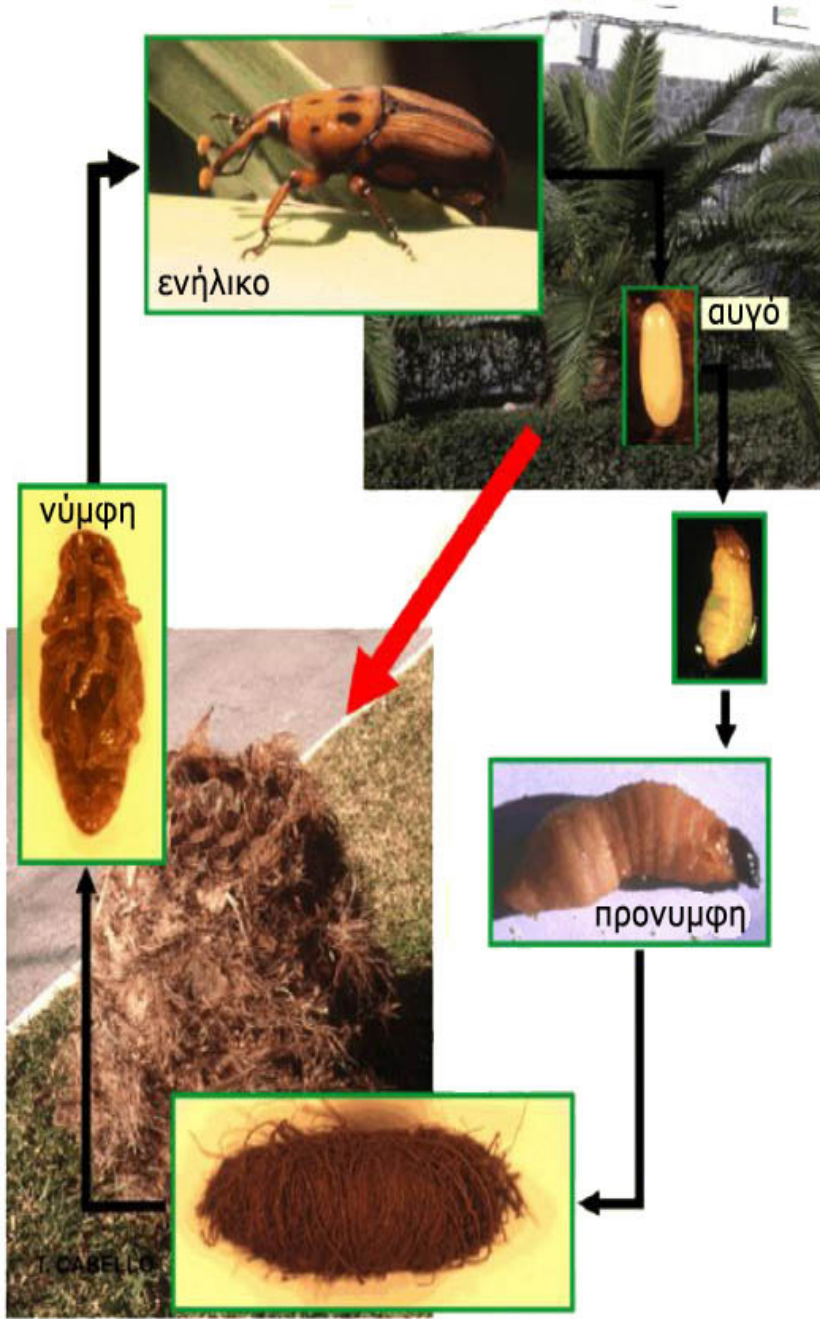


Εικ. 3.21



Εικ. 3.22

⁹ http://www.seea.es/divulgac/galima/Plagas/Rhynchophorus_ferrugineus/



Εικ. 3.23



Εικ. 3.24



Εικ. 3.25



Εικ. 3.26

4 ΚΕΦΑΛΑΙΟ

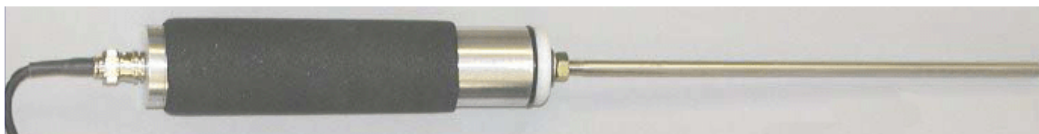
4. ΑΝΙΧΝΕΥΣΗ ΚΑΙ ΚΑΤΑΓΡΑΦΗ

4.1. ΕΙΣΑΓΩΓΗ

Η ανίχνευση, παρακολούθηση και καταγραφή των ειδών είναι μια πολύ σημαντική διαδικασία, με σκοπό την μελέτη της συμπεριφοράς τους, των συνθηκών τους και γενικότερα την καταγραφή και καταμέτρηση του πληθυσμού του κάθε είδους.

Όσον αφορά την ανίχνευση των εντόμων και γενικά παρασίτων πρέπει να αναλογιστούμε τον τόπο που ζουν και δραστηριοποιούνται.

Μεγάλος αριθμός εντόμων φέρει πτερύγια, μηχανισμός που βοηθάει να πετάξουν και να διανύσουν μεγάλες αποστάσεις, όπως για παράδειγμα κουνούπια, ακρίδες και άλλα, σε αντίθεση, για παράδειγμα, με τα μυρμήγκια ή τους τερμίτες που δεν φέρουν πτερύγια, επομένως δραστηριοποιούνται στις περιοχές που μπορούν να φτάσουν περπατώντας. Φυσικά, υπάρχει περίπτωση, ακόμα και αυτά που δεν έχουν πτητικό μηχανισμό, να διανύσουν μεγάλες αποστάσεις παρասυρόμενα από τον αέρα ή να ταξιδεύουν μέσα σε φυτά, αγροτικά προϊόντα κ.α.



Εικ. 4.1 Αισθητήρας ανίχνευσης SP-1 Probe της Acoustic Emission Consulting.

Τα περισσότερα έντομα λειτουργούν σαν πρωταρχικοί αποικοδομητές, στην λειτουργία αποσύνθεσης της οργανικής ύλης, με σκοπό την ανακύκλωση των οργανικών και ανόργανων στοιχείων της φύσης. Αυτό σημαίνει ότι προσβάλουν οτιδήποτε υπάρχει γύρω μας.

Από φυτά μέχρι ζώα, νεκρά ή ζωντανά. Υπάρχουν έντομα και παράσιτα, όπως είναι ο δάκος και η μελίγκρα, στις ελιές, που καταστρέφουν τον καρπό και το φύλλωμα του υγιούς δέντρου. Ένα άλλο παράδειγμα είναι το τσιμπούρι και ο ψύλλος, που προσβάλουν ζώα πίνοντας το αίμα τους και μπορούν, σε μερικές περιπτώσεις, ακόμα και να σκοτώσουν το ζώο. Ακόμα και στο σώμα του ανθρώπου, υπάρχουν παράσιτα και συγκεκριμένα υπάρχει ένα είδος παρασίτου που

ζει πάνω στο δέρμα, κάτω από το μάτι, και τρέφεται με τα νεκρά κύτταρα του δέρματος.

Αναδιατυπώνοντας, έντομα και παράσιτα μπορούν να ανιχνευθούν πάνω στο χώμα, μέσα σε αυτό, στους κορμούς των δέντρων, στα φυτά, μέσα σε καρπούς – προϊόντα, και ιπτάμενα στον αέρα. Έτσι, γνωρίζοντας κάποιες βασικές δραστηριότητες, συμπεριφορές και χαρακτηριστικά του κάθε εντόμου, μπορούμε να το εντοπίσουμε. Εάν βρίσκεται μέσα στο χώμα, μέσα σε κάποιο προϊόν ή μέσα σε ξύλο, κορμό δέντρου, η ακουστική του ανίχνευση μπορεί να γίνει με αισθητήρες ανίχνευσης (sensors), οι οποίοι μπορούν να εισέρχονται μέσα στο χώμα ή το ξύλο ή να εφάπτονται πάνω σε επιφάνειες, που μέσα ή κάτω από αυτές βρίσκεται το έντομο.

Στην περίπτωση που χρειάζεται να ανιχνευθεί έντομο το οποίο πετάει η χρήση των αισθητήρων που αναφέραμε είναι απαγορευτική. Για μια τέτοια ανίχνευση, το κατάλληλο μέσο, είναι το μικρόφωνο. Όπως είναι λογικό, τα μικρόφωνα αυτά πρέπει να είναι κατευθυντικά ώστε να περιορίζεται η λήψη ήχων, οι οποίοι βρίσκονται εκτός του άξονα λήψης (off axis) του μικροφώνου. Σε περιπτώσεις που θέλουμε να αυξήσουμε την κατευθυντικότητα ενός μη κατευθυντικού μικροφώνου, μπορεί να γίνει χρήση παραβολικού κατόπτρου. Όλα αυτά βεβαίως ισχύουν, όταν πρόκειται για ήχους οι οποίοι είναι μέσα στο ακουστικό φάσμα, δηλαδή από 20Hz ως 20kHz.



Εικ. 4.2 Αισθητήρας ανίχνευσης μέσα σε ξύλο



Εικ. 4.3 Αισθητήρας ανίχνευσης μέσα στο χώμα



Εικ. 4.4 Αισθητήρες ανίχνευσης (επαφής) της Acoustic Emission Consulting

Όταν όμως το φάσμα της ηχητικής εκπομπής του εντόμου, που θέλουμε να ανιχνεύσουμε, μπορεί να είναι και πάνω από το φάσμα των 20-20000Hz, τότε πρέπει να γίνει χρήση ultrasound μικροφώνου. Τέτοια μικρόφωνα χρησιμοποιούνται, συνήθως, για την ανίχνευση νυχτερίδων και ποντικιών.

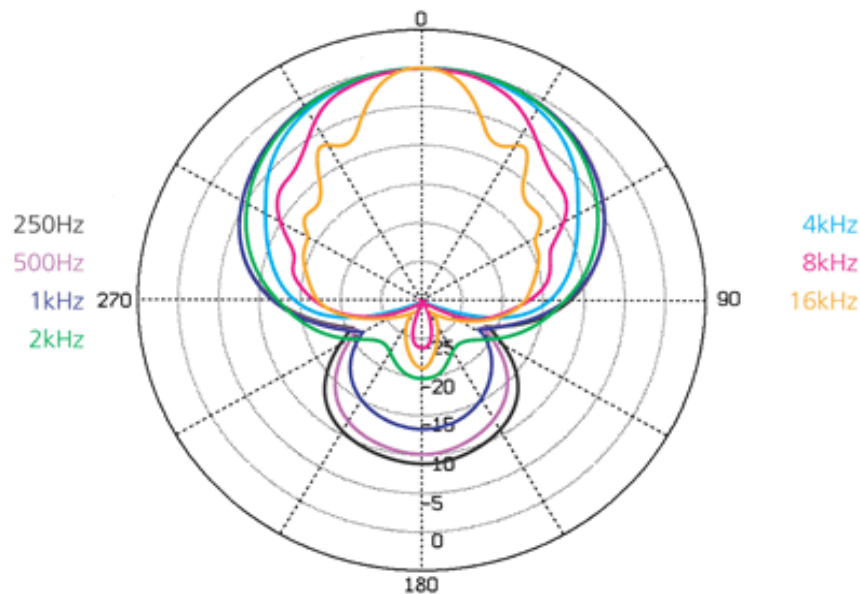


Εικ. 4.5 Ultrasound μικρόφωνο της Avisoft Bioacoustics

4.2. ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ

Υπάρχουν διάφοροι τρόποι συλλογής δεδομένων όσον αφορά την ανίχνευση και την ηχογράφηση διαφόρων ειδών. Στην περίπτωση των εντόμων, θα εξετάσουμε μερικούς από αυτούς και ύστερα θα αποφανθούμε για το ποιος είναι καταλληλότερος όσον αφορά τον εντοπισμό και την ηχογράφηση του Red Palm Weevil.

Οι πιο γνωστές μέθοδοι ανίχνευσης και ηχογράφησης, διεκπεραιώνονται με:

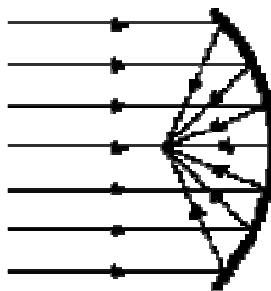


Εικ. 4.6 Πολικό διάγραμμα κατευθυντικού μικροφώνου

4.2.1. Υπερκατευθυντικά μικρόφωνα (Shotgun)

Το υπερκατευθυντικό μικρόφωνο είναι ένα μικρόφωνο του οποίου το διάφραγμα έχει κατασκευαστεί με τέτοιο τρόπο ώστε το πολικό του διάγραμμα να είναι πολύ στενό. Δηλαδή λαμβάνει περισσότερο τους ήχους που βρίσκονται on axis, σε σχέση με ήχους που έρχονται από άλλες γωνίες (Εικ.4.6). Αντί κατευθυντικού μικροφώνου, μπορούμε να χρησιμοποιήσουμε ένα κοινό μικρόφωνο σε συνδυασμό με παραβολικό κάτοπτρο.

Η μέθοδος αυτή στηρίζεται στη χρήση μικροφώνων προκειμένου να ηχογραφηθούν οι διάφοροι ήχοι που παράγει ένα έντομο (π.χ. «ερωτικό» κάλεσμα, ήχος των φτερών του, ήχος την ώρα που τρώει κ.λ.π.). Στη συγκεκριμένη περίπτωση το μικρόφωνο τοποθετείται με κατεύθυνση προς το έντομο, ενώ είναι συνδεδεμένο με ένα ψηφιακό recorder όπου καταλήγουν τα ακουστικά (audio) σήματα. Πριν το recorder, όμως, προηγείται μία βαθμίδα ενίσχυσης και η πιθανή εφαρμογή ενός φίλτρου, που εξαρτάται από το «αντικείμενο» της ηχογράφησης. Τα δεδομένα, στη συνέχεια, ψηφιοποιούνται και αναλύονται από ειδικά λογισμικά, επεξεργασίας ήχου και αποθηκεύονται σε ηλεκτρονικό υπολογιστή.



Εικ. 4.7 Το παραβολικό κάτοπτρο.



Εικ. 4.7 Κατευθυντικό μικρόφωνο σε συνδυασμό με παραβολικό κάτοπτρο

4.2.2. Πιεζοηλεκτρικά μικρόφωνα – Αισθητήρες

Το πιεζοηλεκτρικό μικρόφωνο περιέχει ένα πιεζοηλεκτρικό κρύσταλλο, από χαλαζία ή άλλο υλικό, το οποίο έχει την ιδιότητα όταν του ασκείται πίεση να εμφανίζονται στην επιφάνειά του ηλεκτρικά φορτία. Ο κρύσταλλος αυτός εφάπτεται πάνω στο διάφραγμα του μικροφώνου. Με την ταλάντωση του διαφράγματος πιέζεται ο κρύσταλλος και δημιουργεί στην επιφάνειά του μεταβαλλόμενα ηλεκτρικά φορτία.^{10,11}

Σε αυτή την περίπτωση, αυτοί οι ειδικοί αισθητήρες τοποθετούνται, απ' ευθείας, πάνω στην επιφάνεια, όπου εικάζεται ότι υπάρχουν έντομα. Η επιφάνεια αυτή μπορεί να είναι το έδαφος, κάποια ξύλινη επιφάνεια, ένα δέντρο, αποθηκευμένα γεωργικά προϊόντα κ.λ.π. Ο αισθητήρας λαμβάνει δονήσεις που μπορεί να προέρχονται από την κίνηση του εντόμου μέσα σε μια επιφάνεια ή από το θόρυβο που κάνει όταν τρώει ή όταν εκτελεί άλλες λειτουργίες. Οι δονήσεις μετατρέπονται σε ηλεκτρικό σήμα που οδηγούνται σε σύστημα καταγραφής και επεξεργασίας δεδομένων.



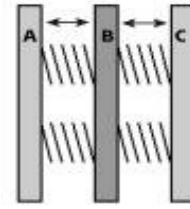
Εικ. 4.8 Ο ειδικός αισθητήρας εισάγεται στο χώμα για την ανίχνευση παρασίτων

¹⁰<http://www.kybernografoi.gr/modules.php?name=News&file=article&sid=171&mode=thread&order=0&thold=0>

¹¹<http://www.focusmag.gr/articles/view-main-category.rx?cid=68>

4.2.3. Επιταχυνσιόμετρο (Accelerometer)

Το επιταχυνσιόμετρο υπολογίζει την επιτάχυνση χρησιμοποιώντας ένα ολοκληρωμένο κύκλωμα, που αρχικά είχε σχεδιαστεί για να ελέγχει την ενεργοποίηση των αερόσακων στα αυτοκίνητα. Αυτό το ολοκληρωμένο κύκλωμα αποτελείται από μικροσκοπικές πλάκες, με περίβλημα από σιλικόνη, οι οποίες είναι τοποθετημένες και συνδεδεμένες όπως οι επαφές ενός πυκνωτή. Οι πλάκες αυτές δονούνται όταν επιταχύνονται και κατά την δόνηση τους αλλάζει η χωρητικότητα. Ένα άλλο κύκλωμα, που περιλαμβάνεται στη διάταξη που αναφέραμε, υπολογίζει την χωρητικότητα και την μετατρέπει σε τάση. Τέλος, ένας ενισχυτής ενισχύει και φιλτράρει τα σήματα στις εξόδους του ολοκληρωμένου κυκλώματος.¹²



Εικ. 4.9 Οι πλάκες του επιταχυνσιόμετρου

Όσον αφορά τη συγκεκριμένη μέθοδο, χρησιμοποιούνται μεταλλικά καρφιά, τα οποία εισάγονται μέσα στην επιφάνεια την οποία ερευνούμε (έδαφος, κορμός δέντρου, γεωργικά προϊόντα κ.λ.π.). Το επιταχυνσιόμετρο συνδέεται, μαγνητικά, πάνω στο καρφί με σκοπό να λάβει τις ακουστικές δονήσεις που, πιθανόν, να λαμβάνουν χώρα, 40 cm γύρω από το καρφί, καθώς και τη μετέπειτα μετατροπή τους σε ηλεκτρικό σήμα. Από αυτό το στάδιο και μετά, ακολουθεί η γνωστή διαδικασία επεξεργασίας και αποθήκευσης της audio πληροφορίας.

4.3. ΔΙΑΧΕΙΡΙΣΗ ΔΕΔΟΜΕΝΩΝ

Εδώ πρέπει να σημειωθεί ότι τα δεδομένα μπορούν να καταγραφούν στον σκληρό δίσκο ενός φορητού υπολογιστή (lap top), σε μία κινητή μονάδα δίσκου (hard disk recording) ή σε κάποιο μέσο μαγνητικής καταγραφής, που στηρίζεται στην χρήση κασέτας, όπως ένα DAT. Επίσης, τα δεδομένα μπορούν να αποστέλλονται μέσω δικτύου (internet audio streaming) ή να εκπέμπονται ασύρματα, μέσω δορυφόρου (ή άλλης τεχνολογίας ασύρματης επικοινωνίας) σε ένα κεντρικό σύστημα. Εκεί μπορεί να γίνεται η επεξεργασία τους, που εστιάζεται κυρίως στην ταξινόμηση και κατηγοριοποίηση των ήχων. Αυτή η τεχνική εξυπηρετεί μακροπρόθεσμους στόχους, που επικεντρώνονται στην ιδέα της αυτόματης καταγραφής πληθυσμού μιας ζωικής ομάδας και γενικά της πανίδας μιας περιοχής, όπου υπάρχει πλήρης απουσία του ανθρώπινου παράγοντα και όλες οι καταγραφές μπορούν να γίνουν από μακριά.¹³

¹² <http://en.wikipedia.org/wiki/Accelerometer>

¹³ Vlad M. Trifa, “A framework for bird songs detection, recognition and localization using acoustic sensor networks”, Master thesis, UCLA, 17 February 2006, §3.1.1

4.4. ΑΝΙΧΝΕΥΣΗ ΚΑΙ ΚΑΤΑΓΡΑΦΗ ΤΟΥ ΚΟΚΚΙΝΟΥ ΡΥΓΧΩΤΟΥ ΚΑΝΘΑΡΟΥ ΤΩΝ ΦΟΙΝΙΚΟΕΙΔΩΝ (RED PALM WEEVIL)

Το έντομο που εξετάζουμε στην εργασία μας, ο *Rhynchophorus Ferrugineus*, παρασιτεί μέσα σε κορμούς δένδρων, και συγκεκριμένα, ορισμένων φοινικοειδών. Εξ' αιτίας αυτού του γεγονότος, η μέθοδος που θα ενδεικνυόταν, περισσότερο, όσον αφορά την ανίχνευση και καταγραφή του, είναι εκείνη της χρήσης ειδικού αισθητήρα, ο οποίος θα λαμβάνει δονήσεις μέσα από το δέντρο. Η χρήση ενός κατευθυντικού μικροφώνου είναι απαγορευτική καθώς ο λόγος σήματος προς το θόρυβο βάθους θα ήταν πολύ μικρός, οπότε και η διαδικασία της αναγνώρισης θα γινόταν πιο δύσκολη.

4.5. ΣΥΣΚΕΥΕΣ ΑΝΙΧΝΕΥΣΗΣ

Συσκευή ακουστικής ανίχνευσης AED 2000¹⁴

4.5.1. Εισαγωγή

Το όργανο "AED 2000" σε συνδυασμό με ,διαφόρων τύπων, ειδικούς αισθητήρες (sensors), όπως ο SP-1L αποτελεί ένα αποτελεσματικό εργαλείο, όσον αφορά την ανίχνευση εντόμων που παρασιτούν σε μία περιοχή. Οι εφαρμογές του είναι πολυάριθμες και εντοπίζονται κυρίως στην ανίχνευση:

- Τερμιτών που ζουν μέσα σε ξύλα, σε δέντρα και στο έδαφος
- Κανθάρων, σκαθαριών και διαφόρων τύπων κάμπιας που τρέφονται με ρίζες
- Του μυρμηγκιού, γένους *campanotus*,
- Μέλισσας, γένους *xylocopa*
- Red Palm Weevil (ρυγχωτός κάνθαρος), που παρασιτεί σε ορισμένα είδη φοίνικα
- Ασιατικών σκαθαριών, γένους *Longhorn*
- Διαφόρων εντόμων που παρασιτούν σε αποθηκευμένα κοκκώδη προϊόντα.

¹⁴ <http://www.aeconsulting.com>



Εικ. 4.10 Το AED – 2000 της Acoustic Emission Consulting

4.5.2. Χαρακτηριστικά συσκευής

- Απλή λειτουργία πληκτρολογίου 12 θέσεων για πλοήγηση στο μενού
- Ένδειξη LCD, τεσσάρων γραμμών x 16 χαρακτήρες για επιλογή και ρύθμιση παραμέτρων
- Ρυθμιζόμενο Gain και Threshold, με LED που αναβοσβήνει όταν ξεπεραστεί η τιμή Threshold
- Αποδιαμορφωμένη έξοδος audio σήματος, που βοηθάει στον χαρακτηρισμό του σήματος. Σημαντικό χαρακτηριστικό σε εφαρμογές ανίχνευσης διαρροής χαμηλής συχνότητας, παρακολούθησης κατάστασης μηχανημάτων και εντοπισμού παρασίτων.
- Σειριακή θύρα, συνεχούς σύνδεσης, σε υπολογιστή.
- Λειτουργία με μπαταρίες ή ρεύμα (AC). Χρησιμοποιεί 2 κανονικές ή επαναφορτιζόμενες μπαταρίες των 9 Volt.



Εικ. 4.11 Αισθητήρες ανίχνευσης (επαφής) της Acoustic Emission Consulting

- Έξοδος AC σήματος για σύνδεση με παλμογράφο ή ψηφιακή ηχογράφιση.
- Ευρύ Bandwidth 1KHz – 2.5KHz
- Λειτουργία φίλτρου ζώνης με μεταβαλλόμενο bandwidth
- Αύξηση Gain μέχρι και 20 dB. 40 dB εσωτερικού Gain
- Μεταβαλλόμενο Threshold από 0 έως 5 Volt με διαβαθμίσεις των 20 mV
- Έξοδος DC ρεύματος 24 Volt, για σύνδεση με προενισχυτές ισχύος, και αισθητήρες εσωτερικής προενίσχυσης.

Αισθητήρας SP1 Probe¹⁵

Το μοντέλο αισθητήρα SP1 Probe είναι σχεδιασμένο για μετρήσεις επαφής σε συνδυασμό με τη συσκευή AED 2000. Διαθέτει ένα set βιδωτών εξαρτημάτων, πολλαπλών εφαρμογών, που το μετατρέπουν σε πολυχρηστικό εργαλείο.

Ο αισθητήρας SP1 Probe διατίθεται σε 2 εκδόσεις :

Το SP1L είναι ένα μοντέλο αισθητήρα χαμηλής συχνότητας με έναν συντονισμένο κρύσταλλο στα 40 KHz και έναν ολοκληρωμένο προενισχυτή 40 dB. Με τη χρήση, ειδικά σχεδιασμένων, φίλτρων, ο SP1L μπορεί να λειτουργήσει στο 1KHz, για εφαρμογές ανίχνευσης διαρροής νερού και ανίχνευσης παρασιτικών εντόμων μέσα σε ξύλο, σε προϊόντα και στο χώμα.

Το υψηλόσυχο μοντέλο SP1H είναι σχεδιασμένος για εφαρμογές των 100KHz και πάνω όπως είναι η ανίχνευση διαρροής σε σωληνώσεις υψηλής πίεσης, ανίχνευση διαρροής σε βαλβίδες, παρακολούθηση λειτουργίας μηχανημάτων και ανίχνευση μερικής αποφόρτισης μετασχηματιστών.

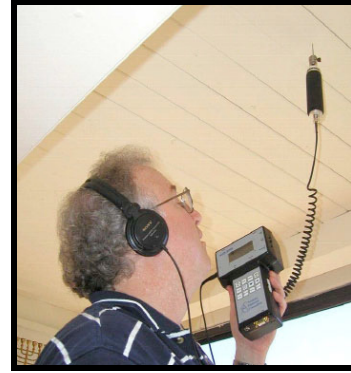


Εικ. 4.12 Ακροσωλήνια επαφής για ανίχνευση διαρροής και για παρακολούθηση λειτουργίας μηχανημάτων

Μαζί με τους αισθητήρες, διατίθεται, επίσης, ένα σύνολο βοηθητικών προεκτάσεων όπως:

¹⁵ <http://www.aeconsulting.com>

- Ένας ισχυρός μαγνήτης NeBFe για επαφή με φερομαγνητικά υλικά.
- Ένα ανοξειδωτο, ατσάλινο ακροσωλήνιο, κυρίως για εισαγωγή σε ξύλινες επιφάνειες και στο έδαφος.
- Ένα ανοξειδωτο, ατσάλινο ακροσωλήνιο, με ελαστικό περίβλημα, για εισαγωγή σε μαλακές επιφάνειες, ικανό να ελαχιστοποιεί τον θόρυβο επαφής σε χαμηλόσυχνες εφαρμογές.



Εικ. 4.13 Το AED – 2000. Ανίχνευση τερμιτών σε ξύλινη οροφή.

Αισθητήρας ABP – 1 Airborne Probe¹⁶

Ο εναέριος αισθητήρας ABP – 1 Airborne Probe είναι σχεδιασμένος για να χρησιμοποιείται μαζί με το AED – 2000, για ένα σύνολο ειδικών εφαρμογών όπως:

- Ανίχνευση διαρροής σε συστήματα που χρησιμοποιούν πεπιεσμένο αέριο.
- Εκφόρτιση κορόνας από ηλεκτρικούς απομονωτές
- Ανίχνευση παρασίτων σε αποθηκευμένους κόκκους προϊόντων
- Ανίχνευση εντόμων που ανοίγουν τρύπες σε δέντρα ή φυτά.

Ο, υψηλής ευαισθησίας, αισθητήρας ABP-1 Airborne Probe είναι συντονισμένος στα 40KHz (χαμηλή περιοχή των υπερήχων) και διαθέτει σύνθετη αντίσταση που να ταιριάζει σε εναέρια ηλεκτρική σύζευξη. Σε συνδυασμό με τις μεθόδους επεξεργασίας του AED – 2000 (μέτρηση threshold και συνεχής επεξεργασία RMS σήματος) είναι ικανός να διακρίνει παλμικές αλλά και συνεχείς πηγές θορύβου. Ένας κώνος 11° ,στο εσωτερικό του αισθητήρα, παρέχει έλεγχο της κατευθυντικότητας.



Εικ. 4.14 Ο αισθητήρας ABP – 1 Airborne Probe

¹⁶ <http://www.aeconsulting.com>

4.5.3. Χαρακτηριστικά αισθητήρα

- Συμπαγής χειρολαβή - πιστόλι από αλουμίνιο
- Εναέριος αισθητήρας των 40 KHz
- Κώνος 11° για κατευθυντικό έλεγχο λήψης
- Ολοκληρωμένος προενισχυτής 40 dB
 - + 24 V DC Τάση
 - Σύνθετη αντίσταση εξόδου 50 Ohm
 - BNC βυσμάτωση
- Διαστάσεις και βάρος
 - 6.5" x 2.4" x 1" (16.5 cm x 6.1cm x 2.5 cm)
 - 10 oz (284 gms)



Εικ. 4.15 Ο ABP – 1 Airborne Probe στην ανίχνευση διαρροής αερίου



Εικ. 4.16 Ο ABP – 1 Airborne Probe στην ανίχνευση εντόμων μέσα σε κόκκους σταριού



Εικ. 4.17 Ο ABP – 1 Airborne Probe στον έλεγχο εκφόρτισης κορόνας

Συσκευή ακουστικής ανίχνευσης Laar WD 60 Pro CSC¹⁷

4.5.4. Εισαγωγή

Το Laar WD 60 Pro CSC είναι, επίσης, ένα μηχάνημα ανίχνευσης που έχει χρησιμοποιηθεί κυρίως για την ανίχνευση καταστροφικών εντόμων μέσα σε φοίνικες, όπως ο Red Palm Weevil. Κατασκευαστής του είναι ο Benedikt Von Laar, ο οποίος και έχει χρησιμοποιήσει το Laar WD 60 Pro CSC για την ανίχνευση του R.P.W. σε φοινικιές. Το μηχάνημα αυτό είναι μια συσκευή χειρός, που είναι κατασκευασμένη να λαμβάνει μικρές ακουστικές δονήσεις και η αποτελεσματικότητά του έχει, επίσης, δοκιμαστεί σε Γερμανικά είδη σκαθαριών που ζουν μέσα στα ξύλα.



Εικ. 4.18 Το Laar WD 60 Pro CSC.

Το standard πακέτο της συσκευής περιέχει:

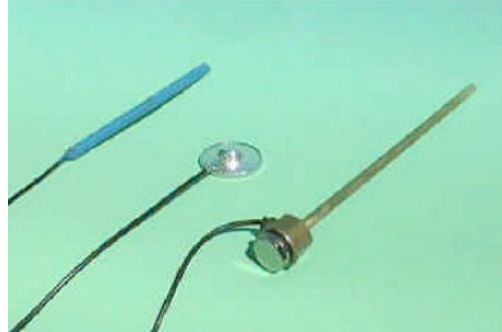
- Το Laar WD 60 Pro CSC
- Ένα ζευγάρι ακουστικά
- Το εγχειρίδιο
- Audio CD με Samples
- CD – Rom με πληροφορίες σχετικά με το προϊόν
- Ένα τρυπάνι των 8mm για εφαρμογή σε ξύλο

¹⁷ <http://www.laartech.biz>

Διατίθενται, προαιρετικά, μπαταρίες για το τρυπάνι, εξοπλισμός για επούλωση πληγής του δέντρου, βίδες μαρκαρίσματος τρύπας.

4.5.5. Χαρακτηριστικά συσκευής

- Δυνατότητα ανίχνευσης ακουστικών δονήσεων από 50 Hz έως 250 KHz
- Line έξοδος σήματος. για σύνδεση με διάφορα μέσα καταγραφής (DAT, Hard Disk)
- Bandwidth: 50Hz – 30KHz
- Έξοδος για ακουστικά
- Είσοδος για μικρόφωνο – αισθητήρα
- Δυνατότητα αλλαγής του αισθητήρα
- Ένδειξη χαμηλής μπαταρίας
- Τάση λειτουργίας : 9V



Εικ. 4.19 Διάφοροι τύποι αισθητήρων (από αριστερά: υδρόφιλος, επαφής και ο απλός)

Εκτός από την ανίχνευση εντόμων το Laar WD 60 Pro CSC προσφέρεται και για χρήση σε εφαρμογές, όπως, ανίχνευση ακουστικής εκπομπής μέσα στο νερό ή και υπερήχων.



Εικ. 4.20 Εφαρμογή μέσα στο νερό με χρήση υδρόφωνου



Εικ. 4.21 Το Laar WD 60 Pro CSC συνδεδεμένο με σύστημα παρακολούθησης (monitoring)

Σύστημα ακουστικής καταγραφής Avisoft Bioacoustics UltraSoundGate¹⁸

4.5.6. Εισαγωγή

Η συσκευή UltraSoundGate της Avisoft Bioacoustics είναι ένα ολοκληρωμένο σύστημα καταγραφής, που χρησιμοποιείται για την ανίχνευση εντόμων, ποντικιών, νυχτερίδων, πουλιών και κητοειδών. Η φιλοσοφία αυτού του οργάνου διαφέρει από τα άλλα δύο που φαίνονται παραπάνω εξ' αιτίας του γεγονότος ότι έχει τέσσερα ανεξάρτητα κανάλια. Πράγμα που σημαίνει ότι μπορεί να λειτουργεί σαν τέσσερις αυτόνομες συσκευές.



Εικ. 4.22 Avisoft Bioacoustics UltraSoundGate 416 (μπροστά)



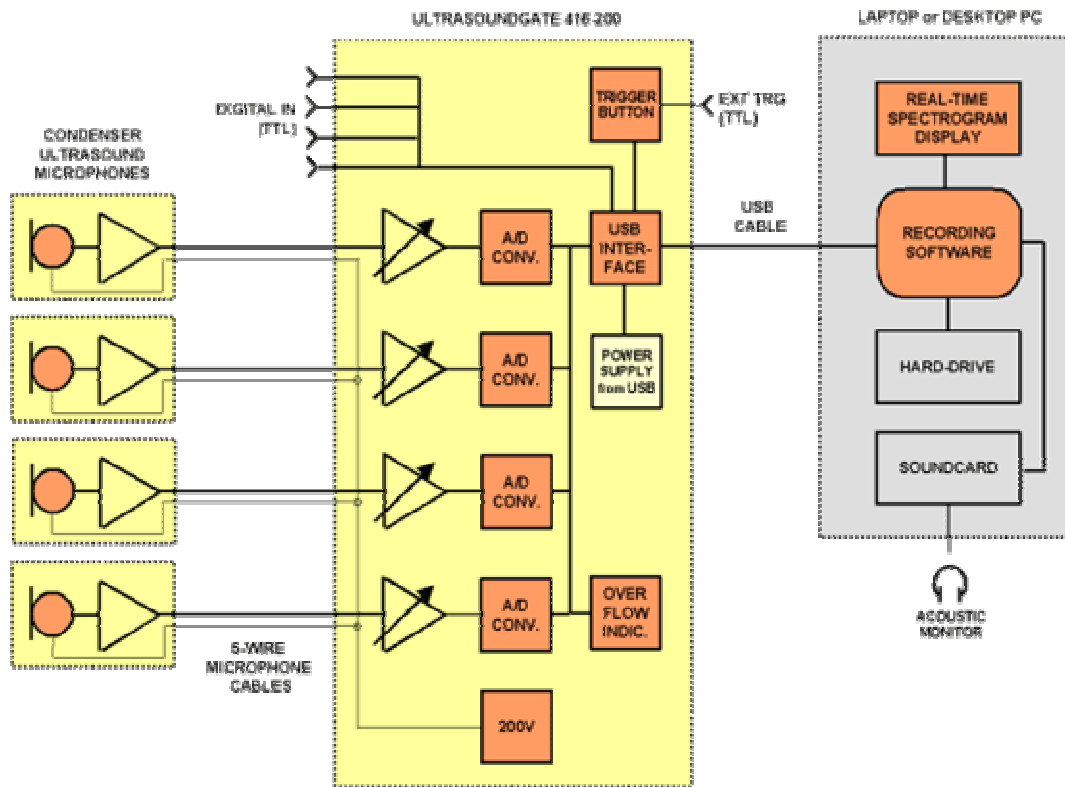
Εικ. 4.23 Avisoft Bioacoustics UltraSoundGate 416 (πίσω)

4.5.7. Χαρακτηριστικά συσκευής

- 4 ξεχωριστοί AD μετατροπείς με 16bit ανάλυση / 1MHz δειγματοληψία.
- Αριθμός καναλιών (XLR βύσματα 5 pins, θηλυκά.): 4 διαφορετικά (4 διαφορετικοί AD μετατροπείς)
- Τροφοδοσία μικροφώνου : 5volt, 20mA , 200volt.
- Μέγιστη ευαισθησία της XLR εισόδου (peak to peak) +-20mvolt.
- Εξωτερικές εισοδοι, 2.5mm mini jack βύσματα: 4 (μία για κάθε αναλογικό κανάλι)
- Εναλλακτικά sample rates : 50, 62.5, 75, 100, 125, 150, 166, 187, 214, 250, 300, 375, 500, 750kHz (8 bit μόνο).

¹⁸ <http://www.avisoft.com>

- Επιλογή ανάλυσης 8 ή 16 bit
- Maximum συναθροιστικά sample rate: 500 kHz στα 16 bit και 1MHz στα 8 bit.



Εικ. 4.24 Το μπλοκ διάγραμμα του UltraSoundGate 416

- Maximum bandwidth αναλογικής εισόδου: 20Hz....370kHz
- Trigger διακόπτης, για απομακρυσμένο έλεγχο της ηχογράφησης στο σκληρό δίσκο, σε συγχρονισμό με το εκάστοτε χρησιμοποιούμενο software.
- USB 1.1 πρωτόκολλο
- Λειτουργία τροφοδότησης μέσω διαύλου (BUS)
- Plug&play λειτουργία για όλες τις συσκευές που συνδέονται.
- Ένδειξη υπερχειλίσης
- Συμπαγές περίβλημα από αλουμίνιο (103x56x145mm συμπεριλαμβανομένου των διακοπών, βάρος: 600gr)
- Υδατοστεγές άθραυστο κουτί μεταφοράς anti-static

Επιπρόσθετα, η συσκευή διατίθεται και με δικό της λογισμικό για την καταγραφή και επεξεργασία των ήχων. Τα μικρόφωνα που χρη-

σιμοποιούνται από την συσκευή δεν είναι όπως τα συμβατικά πυκνωτικά μικρόφωνα, αλλά είναι με 5pin XLR και χρειάζονται διαφορετική τροφοδοσία.



Εικ. 4.25 Υδρόφωνο της Avisoft με μετατροπέα από BNC σε 5 pin XLR ρυθμιζόμενο φίλτρο διέλευσης υψηλών και προενισχυτή



Εικ. 4.26 Μικρόφωνο της Avisoft με επίπεδη απόκριση συχνότητας από 15kHz ως 180kHz

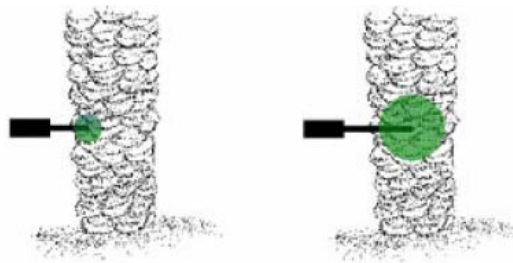
5 ΚΕΦΑΛΑΙΟ

5. ΗΧΟΓΡΑΦΗΣΗ ΚΑΙ ΑΝΑΛΥΣΗ ΦΑΣΜΑΤΟΣ ΤΟΥ ΚΟΚΚΙΝΟΥ ΡΥΓΧΩΤΟΥ ΚΑΝΘΑΡΟΥ ΤΩΝ ΦΟΙΝΙΚΟΕΙΔΩΝ

5.1. ΗΧΟΓΡΑΦΗΣΗ

Όταν ο red palm weevil προσβάλει κάποιο φοίνικα, το φυτό στην αρχή δεν παρουσιάζει κάποια συμπτώματα της προσβολής του. Όταν αρχίσει να φαίνεται ότι ο φοίνικας έχει προσβληθεί από το έντομο, είναι πια αργά για την σωτηρία του και είναι καταδικασμένος. Για τον λόγο αυτόν η ανίχνευση του, μέσα στο φοινικοειδές, είναι μια μακροχρόνια και επίπονη διαδικασία.

Αφού έχει επιλεγεί ο κατάλληλος εξοπλισμός πρέπει να εξετάζουμε τους φοίνικες έναν έναν ξεχωριστά. Είναι προτιμότερο να χρησιμοποιείται αισθητήρας ανίχνευσης, ο οποίος θα μπαίνει μέσα στο φυτό, ώστε η στάθμη θορύβου βάθους να βρίσκεται σε πολύ χαμηλότερα επίπεδα απ' ότι στην περίπτωση που θα γινόταν χρήση αισθητήρα επαφής ή μικροφώνου (βλ. κεφ 4). Αναλυτικότερα η διαδικασία είναι η εξής. Αρχικά πρέπει να ανοιχθεί μια οπή στον φοίνικα από την οποία θα εισάγουμε τον αισθητήρα στο εσωτερικό του φυτού. Η οπή αυτή είναι προτιμότερο να γίνει όσο το δυνατό πλησιέστερα στην κεφαλή του φοινικοειδούς, γιατί σε αυτό το σημείο συνηθίζει να γεννάει τα αυγά του, το έντομο. Δεδομένου ότι ο αισθητήρας μπορεί να ανιχνεύσει δονήσεις που γίνονται γύρω του, μέχρι και 40cm, συνιστάται, η ίδια διαδικασία, να γίνει σε περισσότερα από ένα σημεία, κατά μήκος του κορμού του φοινικοειδούς, ανάλογα με το ύψος του.



Εικ. 5.1 Η ευαισθησία του ανιχνευτή αυξάνεται ανάλογα με το βάθος εισαγωγής

Είναι σημαντικό να χρησιμοποιείται κάποιο ειδικό προϊόν το οποίο βοηθάει το φυτό να επουλώσει τις οπές - πληγές, που ανοίχτηκαν

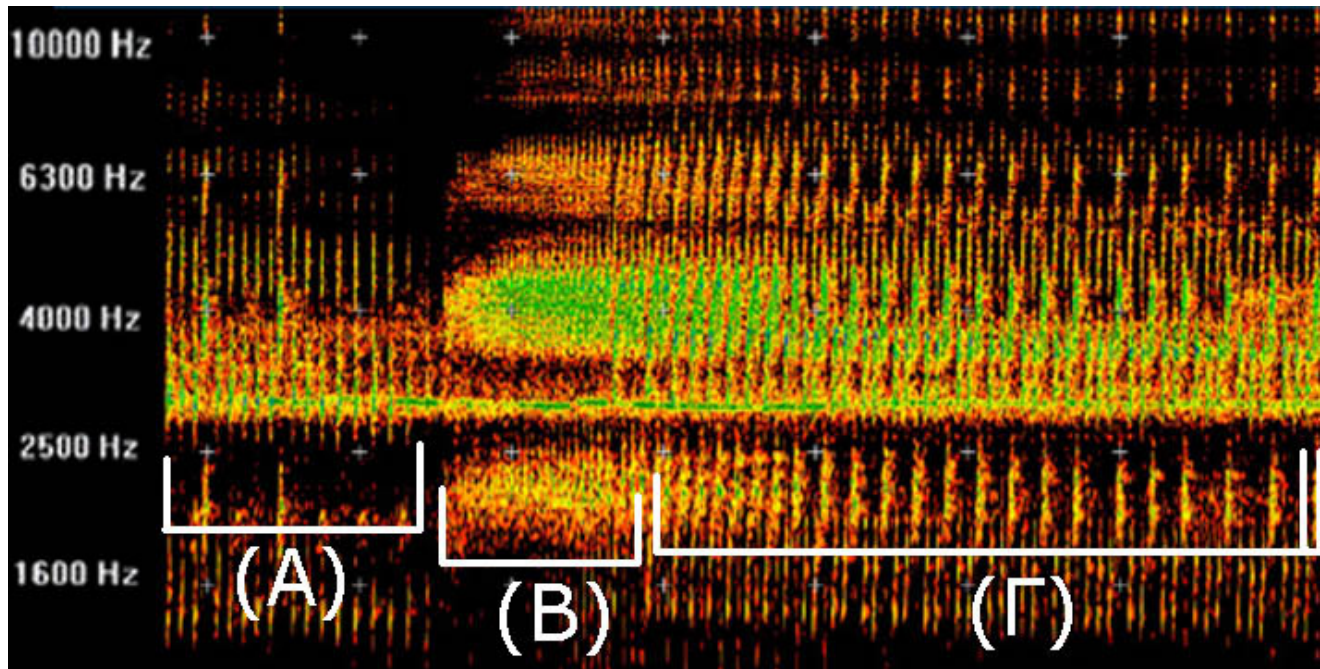
νωρίτερα στο σώμα του. Τα δεδομένα που λαμβάνει ο αισθητήρας περνούν από ενισχυτικές μονάδες και φίλτρα και στη συνέχεια καταγράφονται σε ψηφιακά μέσα αποθήκευσης.

5.2. ΑΝΑΛΥΣΗ ΗΧΗΤΙΚΗΣ ΕΚΠΟΜΠΗΣ

Όπως έχουμε αναφέρει και σε προηγούμενο κεφάλαιο, τα διάφορα είδη παράγουν ηχητικές εκπομπές, των οποίων η διαφορετικότητα δε γίνεται αντιληπτή από το ανθρώπινο αυτί. Η γραφική αναπαράσταση του ήχου, οπτικοποιεί τις διαφορές αυτές, βοηθάει στην ηχητική ανάλυση και αποτελεί ένα ισχυρό εργαλείο για την ταυτοποίηση των ειδών.

5.2.1. Φασματογράφημα (spectrogram) και τρισδιάστατη ανάλυση συχνοτήτων.¹⁹

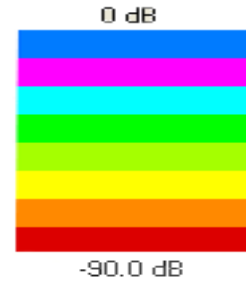
Οι δύο πιο δημοφιλείς γραφικές αναπαραστάσεις είναι το φασματογράφημα (spectrogram) ή τρισδιάστατη ανάλυση συχνοτήτων.



Σχ. 5.1 Φασματογράφημα φώνησης τζίτζικιού.

¹⁹ Διγαλάκης Βασίλης. Σημειώσεις μαθήματος: «Εισαγωγή στην επεξεργασία Φωνής» Τμήμα Ηλεκτρολόγων Μηχανικών Και Μηχανικών Υπολογιστών. Πανεπιστήμιο Κρήτης

Το φασματογράφημα είναι ένα γράφημα του οποίου οι άξονες αναπαριστούν, ο μεν του x τον χρόνο και ο δε y την συχνότητα. Η τρίτη παράμετρος, που είναι το χρώμα, για έγχρωμο γράφημα, ή η πυκνότητα του χρώματος, για γράφημα που χρησιμοποιεί την κλίμακα του γκρι, απεικονίζει το πλάτος έντασης σε dB.



Σχ. 5.2 Αντιστοιχία χρωμάτων του φασματογραφήματος

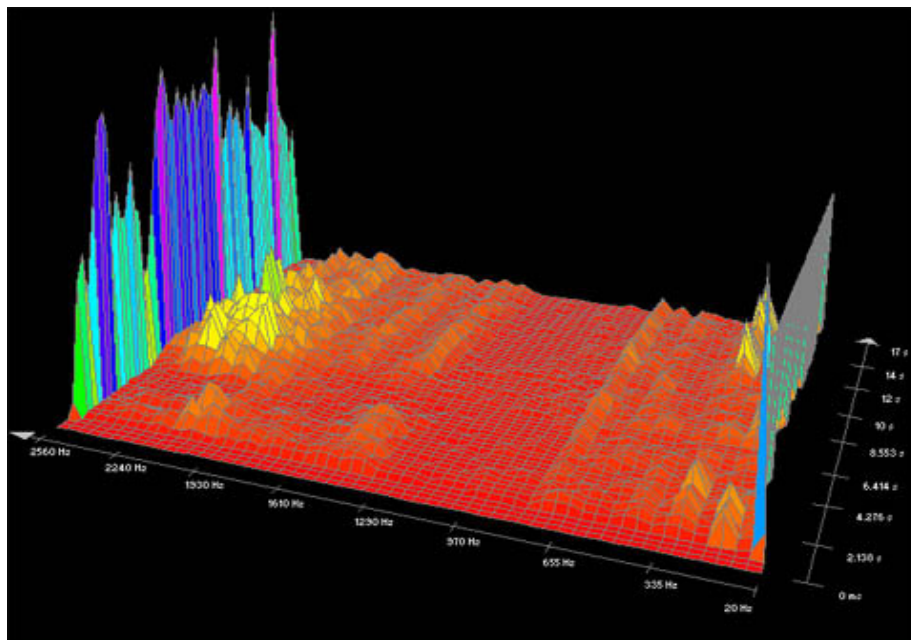
Το γράφημα σχηματίζεται υπολογίζοντας τον Short-time Fourier Transformation (STFT) ενός παραθύρου της φωνής, ανά μικρά χρονικά διαστήματα και σχεδιάζοντας το μέτρο. Σύμφωνα με την αρχή της Αβεβαιότητας (Uncertainty Principle), δεν μπορούμε να έχουμε ακριβή ανάλυση και στο χρόνο και στη συχνότητα.

Τα φασματογραφήματα χωρίζονται σε δύο κατηγορίες. Στα Ευρείας ζώνης και στα Στενής ζώνης.

Το Ευρείας Ζώνης φασματογράφημα (Wideband Spectrogram) δίνει έμφαση στην ανάλυση, σε σχέση με το χρόνο, έτσι ώστε να παρακολουθεί ταχείες μεταβολές του φάσματος και υπολογίζεται χρησιμοποιώντας μικρών εύρων παράθυρα (περίπου 5-20 sec).

Το Στενής Ζώνης φασματογράφημα (Narrowband Spectrogram) δίνει έμφαση στην ανάλυση, σε σχέση με τη συχνότητα, έτσι ώστε να αναπαριστά το φάσμα με αρκετή λεπτομέρεια. Υπολογίζεται χρησιμοποιώντας παράθυρα, με μεγαλύτερο εύρος (της τάξης των 50 msec).

Στην τρισδιάστατη ανάλυση συχνοτήτων, οι δύο κάθετοι άξονες αντιστοιχούν στο χρόνο και στην συχνότητα, όπως και στο φασματογράφημα και το διαφορετικό χρώμα σε διαφορετικό πλάτος έντασης.



Σχ. 5.3 Τρισδιάστατη ανάλυση συχνοτήτων φώνησης τζίτζικιού . Οι χαμηλές συχνότητες που παρουσιάζονται, προέρχονται από τον μηχανικό θόρυβο.

5.2.2. Ανάλυση φασματογραφήματων του κόκκινου ρυγχωτού κανθάρου των φοινικοειδών (R.P.W)

Τα φασματογραφήματα χωρίζονται σε δύο βασικές κατηγορίες, βάση του περιεχομένου τους. Τα χρονικά και τα φασματικά. Τα χρονικά είναι αυτά στα οποία η πληροφορία παρουσιάζεται να επαναλαμβάνεται σε ίσα χρονικά διαστήματα, έχοντας, πάντα, το ίδιο πλάτος έντασης και το ίδιο συχνοτικό περιεχόμενο. Τα φασματικά, είναι τα γραφήματα των οποίων η πληροφορία είναι συγκεχυμένη. Μπορεί να υπάρχουν μόνο κάποιες φράσεις που να επαναλαμβάνονται μεν αλλά άτακτα και χωρίς τα ίδια πλάτη έντασης δε, έχοντας πάντα το ίδιο συχνοτικό περιεχόμενο.

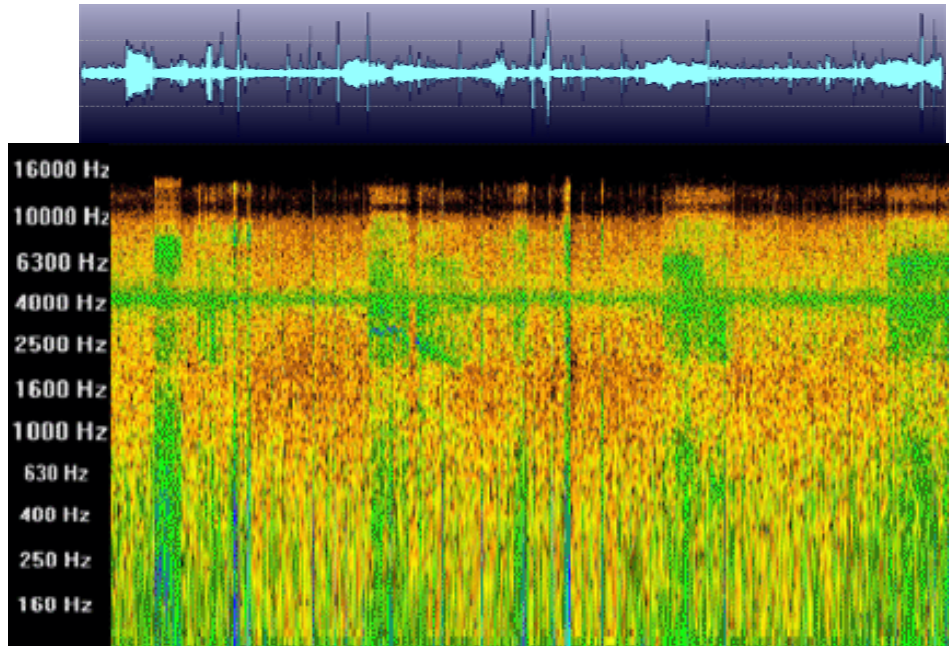
Τα έντομα παράγουν ήχους, οι οποίοι τα βοηθούν στην προσέλκυση του θηλυκού εντόμου από το αρσενικό ή για να δηλώσουν την ενόχληση τους από άλλα έντομα. Μπορούν, ακόμα, να παράγουν ήχους, μη επιτηδευμένα, κατά την διάρκεια κάποιων λειτουργιών τους, όπως, για παράδειγμα, κατά την σίτισή τους, την πέψη τους και άλλα.

Στο σχήμα 5.1 φαίνεται το φασματογράφημα της φώνησης ενός τζιτζικιού. Είναι χρονικό, αφού μπορεί εύκολα να παρατηρηθεί ότι αποτελείται από μια συγκεκριμένη φόρμα, που επαναλαμβάνεται και αποτελείται από τρία διαφορετικά μοτίβα. Το (Α) το (Β) και το (Γ). Η συγκεκριμένη φώνηση έχει σκοπό να προκαλέσει ερωτικά το θηλυκό του είδους.

Στην περίπτωση του red palm weevil οι ηχητικές εκπομπές που παράγονται και εξετάζουμε προέρχονται από την προνύμφη που βρίσκεται μέσα στον φοίνικα. Η μόνη δραστηριότητά της, είναι η σίτιση με την όποια θα συλλέξει θρεπτικά συστατικά που θα οδηγήσουν στην επιτυχή νύμφωσή της όπως και η διαδικασία δημιουργίας της νυμφικής θήκης. Επομένως δεν καλείται θέμα ερωτικού καλέσματος, ούτε επίδειξη κυριαρχίας σε κάποια περιοχή. Οι ηχητικές εκπομπές που μπορεί να προκύψουν από την διαδικασία σίτισης είναι, πρώτον, κατά την ίδια την διαδικασία σίτισης, η ηχητική εκπομπή, δηλαδή, την στιγμή που οι σιαγόνες της προνύμφης θρυμματίζουν το ξύλο, δεύτερον, η κίνησή της την ώρα που τρέφεται ή η κίνηση μέσα στις στοές, που η ίδια έχει ανοίξει και τρίτον, ήχοι που παράγονται κατά την διαδικασία της πέψης ή την εξαγωγή αέρα από το σώμα της.²⁰

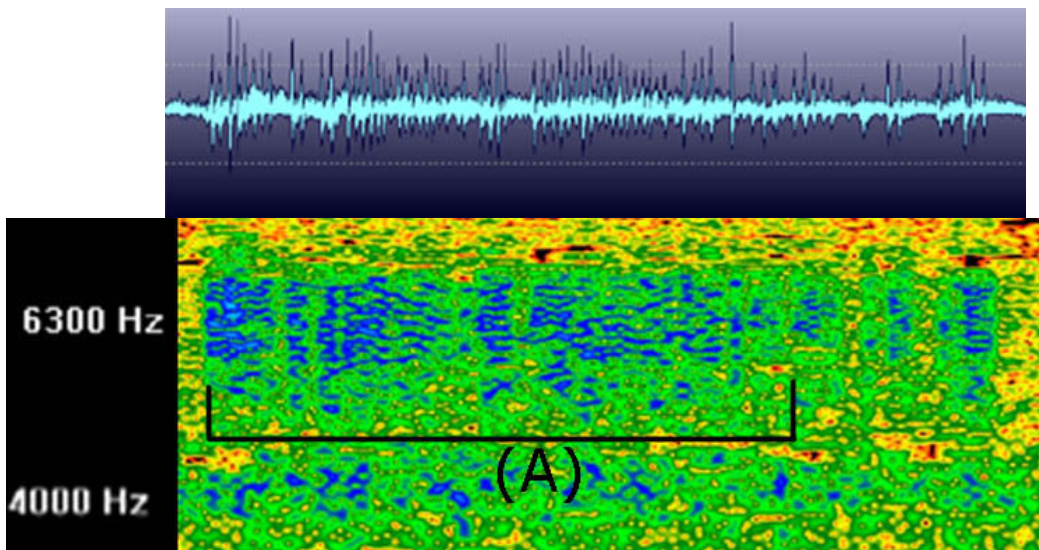
Παρακάτω παραθέτονται φασματογραφήματα από τις χαρακτηριστικότερες εκπομπές του *Rhynchophorus ferrugineus*.²⁰

²⁰ Benedikt von Laar, Inc. “*The bioacoustic detection of the Red Palm Weevil*” April 2002 § “Results Of Measurement And Analysis”

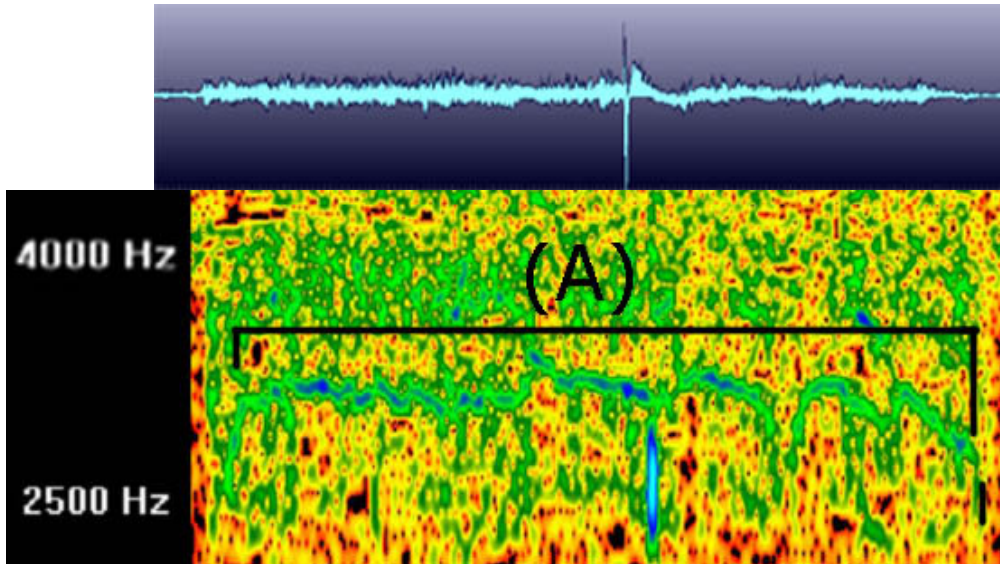


Σχ. 5.4

Στο σχήμα 5.4 παρατίθεται ένα ολοκληρωμένο φασματογράφημα στο οποίο καταγράφονται όλες οι λειτουργίες της προνύμφης. Παρατηρείται ότι το γράφημα είναι ακαθόριστο, χωρίς εμφανή επαναλαμβανόμενα μοτίβα όπως για παράδειγμα, στην περίπτωση του τζίτζικιού.



Σχ. 5.5



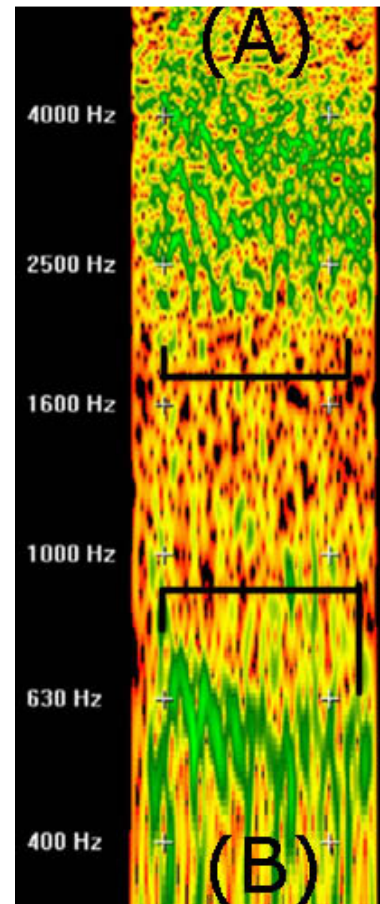
Σχ. 5.6

Η εκπομπή της προνύμφης βρίσκεται στο φάσμα των 400Hz έως 16000Hz με την κύρια δραστηριότητά του να εμφανίζεται από τα 2kHz ως τα 7kHz. Συχνότητες που βρίσκονται κάτω από τα 1000Hz σχετίζονται με το ξύλο, πού είτε πρόκειται περί ανακλάσεων είτε περί τριβής της προνύμφης πάνω σε αυτό.

Στα φασματογραφήματα 5.5, 5.6 και 5.7 στα σημεία (A), (A) και (B) αντιστοίχως, φαίνεται ο ήχος που παράγεται κατά την διαδικασία της πέψης και της αποβολής υγρών από το σώμα της προνύμφης.

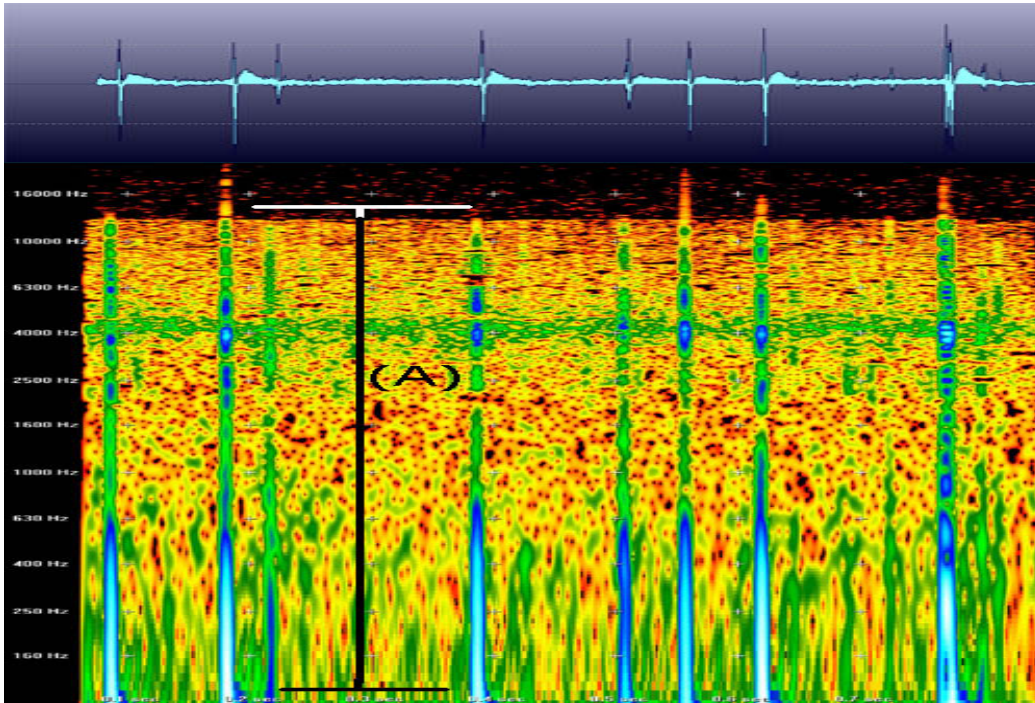
Στο (A) μέρος του γραφήματος 5.7 αναπαριστάται ο ήχος πού παράγεται κατά την κίνηση.

Η ακουστική εκπομπή, κατά τη διαδικασία της πέψης και της αποβολής υγρών, μπορεί να μην έχει πάντα την ίδια ένταση αλλά η συχνοτική της απόκριση και η εξέλιξη της στον χρόνο την χαρακτηρίζει.



Σχ. 5.7

Στο 5.8 οι κάθετες, στον άξονα του χρόνου, γραμμές, που φαίνονται, είναι ο ήχος που παράγεται από την διαδικασία της σίτισης. Είναι ο ήχος του ξύλου από την τριβή των δοντιών της προνύμφης πάνω σε αυτό.



Σχ. 5.8

6 ΚΕΦΑΛΑΙΟ

6. ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΙΑΣ

6.1. ΒΙΟΜΕΤΡΙΚΗ

Η Βιομετρική αναφέρεται στην αυτόματη αναγνώριση ενός ατόμου, βασισμένη σε ψυχολογικά και χαρακτηριστικά συμπεριφοράς. Υπάρχουν πολύ τύποι βιομετρικών χαρακτηριστικών: αναγνώριση προσώπου, αναγνώριση αποτυπωμάτων, γεωμετρία δαχτύλων, γεωμετρία χεριού, αναγνώριση ίριδας, αναγνώριση φλέβας και αναγνώριση φωνής και υπογραφής.

Η μέθοδος της βιομετρικής αναγνώρισης προτιμάται, έναντι των παραδοσιακών μεθόδων που περιλαμβάνουν κώδικες και προσωπικούς αριθμούς ταυτοποίησης (PINs) για διάφορους λόγους:

Το άτομο, προς αναγνώριση, πρέπει να παρίσταται στο σημείο της αναγνώρισης. Η ταυτοποίηση που βασίζεται σε βιομετρικές τεχνικές απαλλάσσει τον χρήστη από το να θυμάται κωδικούς ή να συναλλάσσεται χρησιμοποιώντας κάρτες ή ταυτότητες.

Σαν αποτέλεσμα της επανάστασης στην τεχνολογία της πληροφορίας, είναι αναγκαίο να περιοριστεί η πρόσβαση σε προσωπικά και ευαίσθητα δεδομένα, αντικαθιστώντας τα PINs και τους κωδικούς. Οι βιομετρικές τεχνικές προσφέρουν περισσότερες ανέσεις στον χρήστη και μπορούν να αποτρέψουν πιθανή, μη εξουσιοδοτημένη πρόσβαση σε αυτόματες ταμιακές μηχανές (ATM), κινητά τηλέφωνα, πιστωτικές κάρτες, Η/Υ, workstations και δίκτυα υπολογιστών.

Οι PINs και οι κωδικοί μπορούν να ξεχαστούν και οι μέθοδοι αναγνώρισης που σχετίζονται με κάρτα, όπως είναι το διαβατήριο, το δίπλωμα οδήγησης, η κάρτα ασφάλειας μπορούν να κλαπούν ή να χαθούν. Διάφοροι τύποι βιομετρικών συστημάτων χρησιμοποιούνται για αναγνώριση σε πραγματικό χρόνο. Υπάρχουν βιομετρικά συστήματα τα οποία χρησιμοποιούν σάρωση ίριδας και αμφιβληστροειδούς και γεωμετρία προσώπου και χεριού. Τα περισσότερα, όμως, σχετίζονται

με αναγνώριση προσώπου, φωνής και αντιστοίχιση δακτυλικών αποτυπωμάτων.²¹

6.2. ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΙΑΣ

Στις μέρες μας, η αναγνώριση ομιλίας αποτελεί μία σημαντική και αξιόπιστη τεχνολογία, συνεχώς εξελισσόμενη, που έχει χρησιμοποιηθεί σε πολυάριθμες βιομηχανικές και μη εφαρμογές. Μερικά παραδείγματα είναι:

- Συστήματα ελέγχου πρόσβασης.
- Τραπεζικές συναλλαγές μέσω τηλεφώνου.
- Βιομετρική σύνδεση σε εμπορικά συστήματα που σχετίζονται με το τηλέφωνο.
- Υπηρεσίες πληροφοριών και κρατήσεων.
- Έλεγχος ασφαλείας για εμπιστευτικές πληροφορίες.

Οι εφαρμογές αυτές χωρίζονται σε 3 μεγάλες κατηγορίες, την «αναγνώριση ομιλητή» την «επιβεβαίωση ομιλητή» και την «αναγνώριση ομιλίας».

Αναγνώριση ομιλητή χρησιμοποιείται για τον προσδιορισμό μια άγνωστης φωνής και την αντιστοίχσή της σε μία φωνή από ένα σύνολο γνωστών φωνών. Δηλαδή απαντάμε στο ερώτημα «Ποιος είναι;».

Όσον αφορά τη δεύτερη κατηγορία εφαρμογών, γίνεται πιστοποίηση και εξακρίβωση της ταυτότητας του ομιλητή σε σχέση με την ταυτότητα του ομιλητή που έχουμε θέσει ως μοντέλο σύγκρισης, εκ των προτέρων. Δηλαδή στην περίπτωση αυτή απαντάμε στο ερώτημα «Είναι αυτός ο X ομιλητής ή όχι;».

Στην τελευταία κατηγορία, που είναι αρκετά πιο πολύπλοκη από τις άλλες δύο, προσπαθούμε να αναγνωρίσουμε το περιεχόμενο του λόγου του ομιλητή.

Σε αυτό το κεφάλαιο παρουσιάζονται και αναλύονται κάποιες μέθοδοι, της συγκεκριμένης τεχνολογίας και θα επικεντρωθούμε κυρίως στις δύο πρώτες κατηγορίες εφαρμογών, την «αναγνώριση ομιλητή» και την «επιβεβαίωση ομιλητή».

Στόχος μας είναι η κατασκευή μοντέλων πιθανότητας από «φωνητικά» δείγματα του ρυγχωτού κάνθαρου, πρώτον, για να μπορούμε να διαπιστώσουμε την ύπαρξή του μέσα στο δέντρο και, δεύτερον, για να μπορούμε να τον ξεχωρίσουμε ανάμεσα σε άλλα έντομα, που πιθανόν να παρασιτούν στην ίδια περιοχή.

²¹ Gerik Alexander von Graevenitz “*About Speaker Recognition Technology*”, Bergdata Biometrics GmbH, Germany, §. “*Overview About Biometrics*”

6.3. ΑΝΑΓΝΩΡΙΣΗ ΟΜΙΛΗΤΗ ΚΑΘΟΡΙΣΜΕΝΟΥ ΚΕΙΜΕΝΟΥ ΚΑΙ ΑΚΑΘΟΡΙΣΤΟΥ ΚΕΙΜΕΝΟΥ

Τα συστήματα αναγνώρισης ομιλητή ταξινομούνται σε καθορισμένου κειμένου και ακαθόριστου. Στην πρώτη περίπτωση το σύστημα «γνωρίζει», εκ των προτέρων, την φράση «κλειδί» ή τον κωδικό που θα προφέρει ο χρήστης.

Αυτά τα συστήματα απαιτούν από τον χρήστη να επαναλάβει κάποιες συγκεκριμένες εκφράσεις, που συνήθως περιέχουν το ίδιο κείμενο με τα δεδομένα της «εκπαίδευσής» του. Σε αυτή την περίπτωση η γνώση λέξεων ή ακολουθιών λέξεων μπορεί να εκμεταλλευτεί προκειμένου να βελτιωθεί η απόδοση.

Όσον αφορά την αναγνώριση ομιλητή ακαθόριστου κειμένου, το σύστημα δεν γνωρίζει, εξ' αρχής, το κείμενο που προφέρεται από τον χρήστη, είναι πιο ευέλικτο και τα μοντέλα πιθανότητας γίνονται πιο πολύπλοκα.

Ανακεφαλαιώνοντας, τα πλεονεκτήματα και τα μειονεκτήματα των δύο αυτών περιπτώσεων είναι τα εξής:

6.3.1. Καθορισμένο κείμενο

- Πλεονεκτήματα
 - Γενικά, καλύτερη απόδοση όσο περιορίζεται η διαδικασία.
 - Το σύστημα βελτιώνεται όσο περιορίζεται το μέγεθος του κειμένου προς αναγνώριση.
 - Υπάρχει αυξημένη ασφάλεια, καθώς το σύστημα είναι δύσκολο να εξαπατηθεί.
- Μειονεκτήματα
 - Απαιτεί συγκεκριμένη και ακριβή άρθρωση του λόγου προκειμένου να γίνει η ταυτοποίηση.
 - Απαιτείται δημιουργία μοντέλων και για τη δυναμική φύση του λόγου.

6.3.2. Ακαθόριστο κείμενο

- Πλεονεκτήματα
 - Το σύστημα είναι πιο ευέλικτο και μπορεί να χρησιμοποιηθεί και για μη συγκεκριμένο λόγο.
 - Δεν χρειάζεται η δημιουργία μοντέλων για την δυναμική του λόγου, οπότε υπάρχει υπολογιστικό κέρδος.
- Μειονεκτήματα

- Απαιτεί την χρήση μεγάλης ποσότητας, φωνητικά, ισορροπημένου λόγου.
- Εξ' αιτίας της ποικιλομορφίας του κειμένου, τα μοντέλα είναι πολύπλοκα.

6.4. INTRASPEAKER ΚΑΙ INTERSPEAKER ΠΟΙΚΙΛΟΜΟΡΦΙΑ

Η διαφοράς των χαρακτηριστικών μεταξύ διαφορετικών ομιλητών, ονομάζεται interspeaker ποικιλομορφία (interspeaker variance).

Η interspeaker ποικιλομορφία προκαλείται από τη διαφορετικότητα των φωνητικών χαρακτηριστικών, που παρατηρούνται από άτομο σε άτομο και παρέχει χρήσιμη πληροφορία για τον διαχωρισμό διαφόρων ομιλητών. Ένα άλλο είδος ποικιλομορφίας είναι η intraspeaker ποικιλομορφία (intraspeaker variance), που προκαλείται όταν ένας ομιλητής προφέρει την ίδια λέξη ή πρόταση αλλά δεν μπορεί να επαναλάβει την άρθρωση με, ακριβώς, τον ίδιο τρόπο από την μία εφαρμογή στην άλλη.

Η intraspeaker variance περιλαμβάνει τον διαφορετικό ρυθμό ομιλίας, την συναισθηματική κατάσταση ομιλητή και το περιβάλλον ομιλίας. Είναι ο κύριος παράγοντας που προκαλεί τη μείωση της απόδοσης του συστήματος αναγνώρισης ομιλητή. Γι' αυτό είναι επιθυμητό να επιλέγονται οι παράμετροι που παρουσιάζουν μικρότερη intraspeaker και περισσότερη interspeaker ποικιλομορφία. Σε περισσότερες εφαρμογές αναγνώρισης ομιλητή, είναι πιθανό να μειώνεται η intraspeaker ποικιλομορφία, απαιτώντας από τον χρήστη να προφέρει την δοκιμαστική πρόταση, που περιέχει το ίδιο κείμενο ή λεξιλόγιο όπως οι προτάσεις εκπαίδευσης. Αυτή είναι περίπτωση των μεθόδων αναγνώρισης ομιλητή καθορισμένου κειμένου.²²

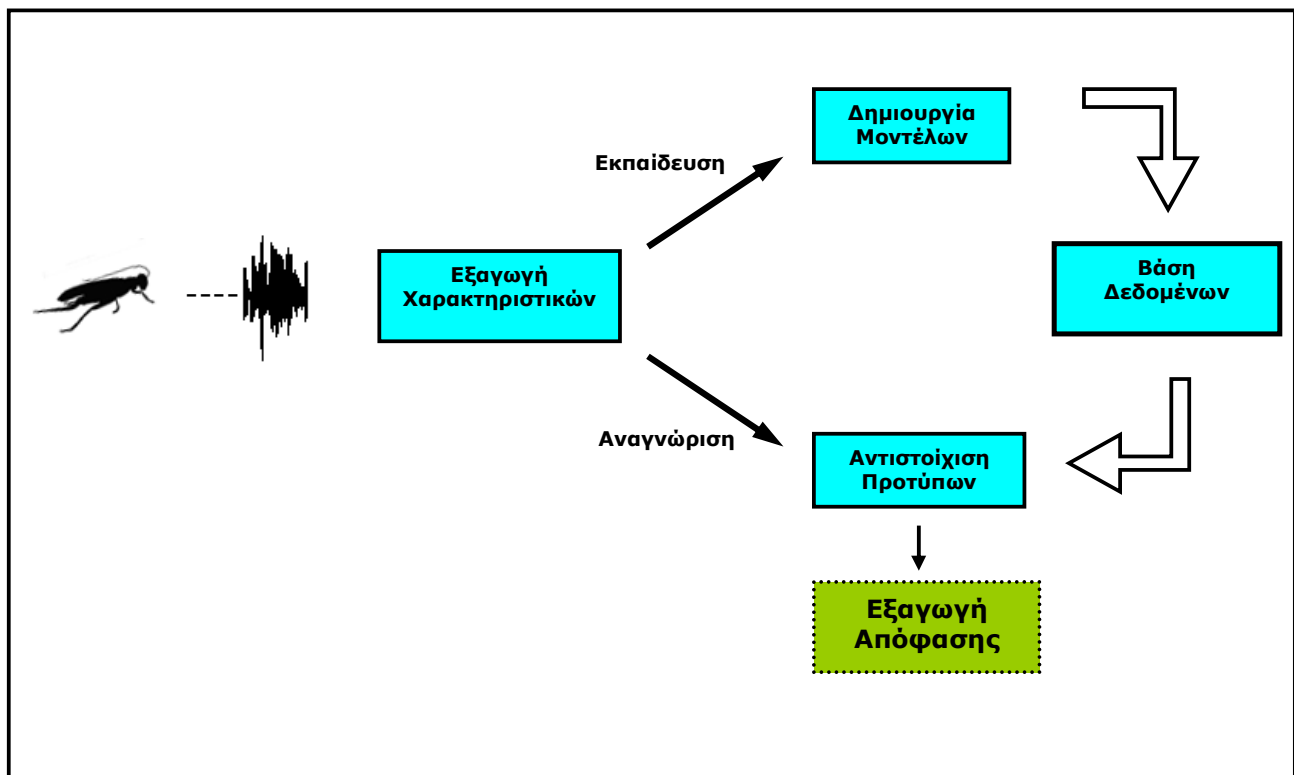
6.5. ΠΗΓΕΣ ΣΦΑΛΜΑΤΩΝ ΕΠΙΒΕΒΑΙΩΣΗΣ

Οι πηγές σφαλμάτων επιβεβαίωσης που μπορούν να προκύψουν είναι οι εξής:²³

²² Gerik Alexander von Graevenitz "About Speaker Recognition Technology", Bergdata Biometrics GmbH, Germany, §"Intraspeaker and Interspeaker Variance"

²³ Gerik Alexander von Graevenitz "About Speaker Recognition Technology", Bergdata Biometrics GmbH, Germany, §"Sources of Verification Error"

- Φράσεις που έχουν διαβαστεί ή εκφραστεί λανθασμένα
- Ακραίες συναισθηματικές συνθήκες.
- Ο τρόπος έκφρασης της φράσης είναι διαφορετικός από τον τρόπο που χρησιμοποιήθηκε στη καταγραφή.
- Αλλαγή θέσης λήψης του μικροφώνου.
- Ασυμβατότητα καναλιού (π.χ. χρήση διαφορετικών μικροφώνων για καταγραφή και επιβεβαίωση).
- Διαφορετική ταχύτητα προφοράς κατά την διάρκεια της επιβεβαίωσης σε σύγκριση με τα δεδομένα της εκπαίδευσης.
- Ασθένεια (π.χ. το κρύωμα μπορεί να αλλοιώσει τη φωνή).
- Η ηλικία.



Σχ. 6.1 Τα στάδια λειτουργίας ενός ολοκληρωμένου συστήματος αναγνώρισης ομιλίας

6.6. ΣΧΕΔΙΑΣΜΟΣ ΣΥΣΤΗΜΑΤΟΣ ΑΥΤΟΜΑΤΗΣ ΑΝΑΓΝΩΡΙΣΗΣ ΟΜΙΛΙΑΣ

Ένα ολοκληρωμένο σύστημα αναγνώρισης ομιλίας λειτουργεί σε 2 φάσεις. Αρχικά, προηγείται η φάση της **εκπαίδευσης (training)** του συστήματος και επακολουθεί η φάση της **αναγνώρισης**, που είναι και ο τελικός στόχος στον οποίο αποσκοπούμε. Αναλύοντας τους όρους :

- **Εκπαίδευση συστήματος:** Εισάγουμε στο σύστημα φωνητικά δείγματα γνωστής προέλευσης, από τα οποία, με ειδική επεξεργασία, θα δημιουργηθούν τα μοντέλα πιθανότητας.
- **Αναγνώριση:** Κάθε νέο, άγνωστο, φωνητικό δείγμα που εισάγεται στο σύστημα συγκρίνεται με τα ήδη υπάρχοντα, γνωστά δείγματα προκειμένου να επιτευχθεί η αναγνώριση, επιβεβαίωση ή ταυτοποίηση του ομιλητή.

Πρέπει να σημειώσουμε, όμως, ότι κάθε μοντέλο πιθανότητας που δημιουργείται αποθηκεύεται σε μία βάση δεδομένων, η οποία χρησιμεύει για την ταξινόμηση του κάθε νέου είδους που ανιχνεύεται. Οι βαθμίδες λειτουργίας του συστήματος φαίνονται στο σχήμα 6.1 ενώ στη συνέχεια αναλύονται τα επί μέρους στάδια λειτουργίας.

6.7. ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ (FRONT-END ΜΗΧΑΝΙΣΜΟΣ)

Μια βασική υπόθεση στην οποία βασίζονται οι αναγνωριστές, είναι ότι το σήμα της φωνής μπορεί να θεωρηθεί σαν στάσιμο (stationary-δηλαδή τα χαρακτηριστικά του φάσματος μπορούν να θεωρηθούν σταθερά) σε ένα διάστημα λίγων msec. Το σήμα φωνής χωρίζεται σε ένα σύνολο από διαστήματα (frames) και για κάθε διάστημα υπολογίζεται ένα ομαλοποιημένο φάσμα. Τα διαστήματα έχουν συνήθως μήκος 10 msec και αλληλεπικαλύπτονται δίνοντας έτσι ένα παράθυρο μεγαλύτερου μήκους. Χρησιμοποιώντας για παράδειγμα παράθυρο Hamming και εφαρμόζοντας ανάλυση Fourier ή γραμμικής πρόβλεψης (LPC) παίρνουμε τις βασικές φασματικές παραμέτρους τις οποίες, με διάφορους μετασχηματισμούς, μετατρέπουμε σε κάποια κατάλληλη μορφή.

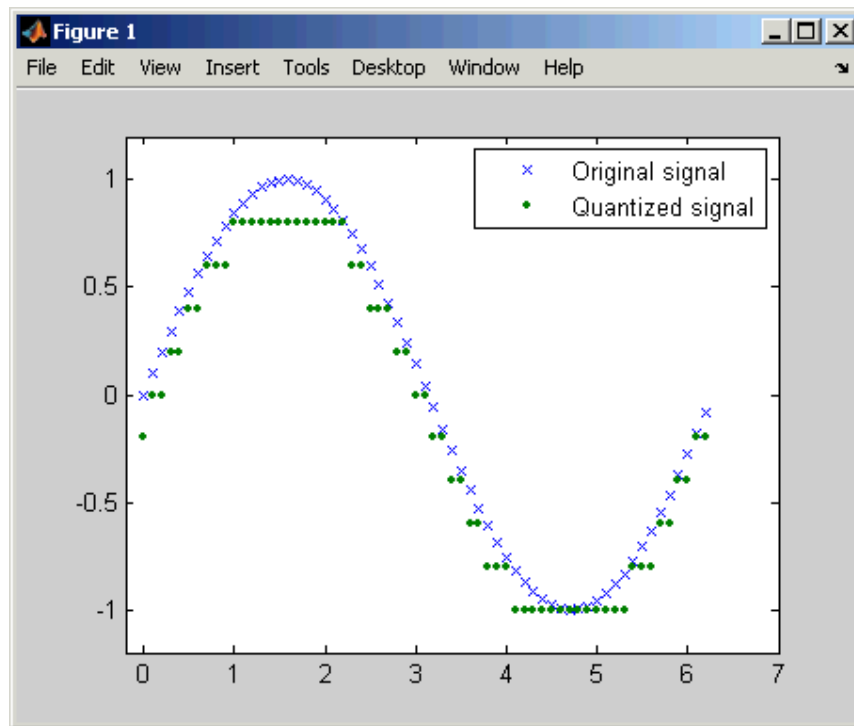
Μια τέτοια μορφή, για τα ακουστικά διανύσματα, είναι οι *mel-frequency cepstral coefficients (MFCCs)*. Αφού πάρουμε το φάσμα του σήματος, εφαρμόζουμε μια κλίμακα (*mel-frequency*) η οποία σχεδιάζεται, ουσιαστικά, ώστε να προσεγγίζει την φασματική ανάλυση της ανθρώπινης ακοής, που είναι γραμμική έως τα 1000Hz και λογαριθμική από εκεί και πέρα. Η κλίμακα αυτή έχει αποδειχθεί, πειραματικά, ότι βελτιώνει την ακρίβεια αναγνώρισης. Στη συνέχεια, για να κάνουμε την ισχύ του φάσματος Γκαουσιανή εφαρμόζουμε λογαριθμική συμπίεση. Τέλος, αφού εφαρμόσουμε Discrete Cosine Transform, παίρνουμε ένα ακουστικό διάνυσμα, που αποτελείται από 12 cepstral συντελεστές και την ενέργεια του σήματος.

6.7.1. Η διαδικασία κβαντισμού των παραμέτρων

Το τελικό αποτέλεσμα, επομένως, είναι η παραγωγή ενός *MFCC* διανύσματος για κάθε *frame* (δηλαδή για κάθε 10 msec) φωνής. Ο βασικός σκοπός αυτής της εργασίας, είναι ο βέλτιστος διανυσματικός κβαντισμός των παραπάνω διανυσμάτων. Η διαδικασία αυτή περιγράφεται εκτενώς σε επόμενο κεφάλαιο. Εδώ πολύ σύντομα θα αναφέρουμε απλά την κεντρική ιδέα, που είναι η εξής :

Το *MFCC* διάνυσμα Y χωρίζεται κατάλληλα σε μικρότερα υποδιανύσματα y_1, y_2, \dots, y_L και κάθε τέτοιο υποδιάνυσμα κβαντίζεται στο πιο κοντινό σε αυτό υποδιάνυσμα, από ένα σύνολο από τέτοια κατάλληλα επιλεγμένα υποδιανύσματα (το σύνολο αυτό ονομάζεται *codebook*)

Το αποτέλεσμα του κβαντισμού είναι η συμπίεση του σήματος (αφού κάθε υποδιάνυσμα αναπαρίσταται από το δείκτη του υποδιανύσματος στο οποίο κβαντίστηκε) και η επιτάχυνση της αναγνώρισης .



Σχ. 6.2 Κβαντοποίηση σήματος

7 ΚΕΦΑΛΑΙΟ

7. ΜΕΘΟΔΟΙ ΕΞΑΓΩΓΗΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΚΑΙ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ

Η βασική ιδέα της διαδικασίας της εξαγωγής χαρακτηριστικών είναι η ανάλυση του σήματος με σκοπό την εξαγωγή της τιμής ενός συνόλου χαρακτηριστικών, σε ένα περιοδικό και διακριτό βήμα χρόνου, και η χρήση αυτής της πληροφορίας για την σύγκριση μεταξύ διαφορετικών φωνητικών δειγμάτων.

Το ζητούμενο είναι, ουσιαστικά, η αναπαράσταση του σήματος, μέσω μιας παραμετρικής φόρμας, και η συμπίεση της πληροφορίας που απαιτείται για να αποθηκευτεί, εξάγοντας μόνο την χρήσιμη πληροφορία από αυτό.

7.1. ΣΤΑΔΙΑ ΕΞΑΓΩΓΗΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

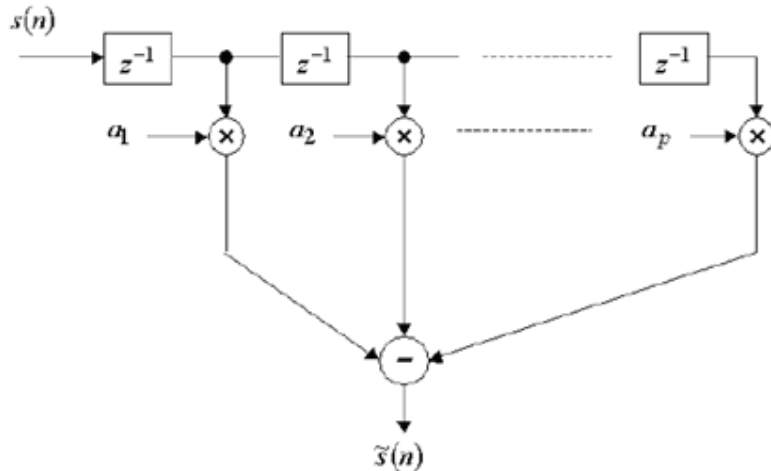
Πριν την εξαγωγή των χαρακτηριστικών από ένα φωνητικό δείγμα, προηγείται η φάση της προεπεξεργασίας. Εδώ πρέπει να σημειώσουμε ότι η εξαγωγή χαρακτηριστικών ορίζεται ως «η παραμετροποίηση του αρχικού σήματος ή φωνητικού δείγματος και η συμπιεσμένη αναπαράστασή του σε μορφή διανυσματικών μεγεθών, με σκοπό την ταξινόμηση και την σύγκρισή του με άγνωστα φωνητικά δείγματα». Τα επί μέρους στάδια της προεπεξεργασίας, που είναι η προέμφαση και η εφαρμογή παραθύρου, (Windowing) αναλύονται παρακάτω.

7.1.1. Προέμφαση

Πριν την το στάδιο της κυρίως επεξεργασίας του σήματος, εφαρμόζεται ένα φίλτρο με σκοπό να εξομαλύνει τις φασματικές ανομοιομορφίες που υπάρχουν στο εκάστοτε σήμα. Όσον αφορά την περίπτωση της ανθρώπινης φωνής γίνεται μία αύξηση στις υψηλές συχνότητες του φωνητικού δείγματος, έτσι ώστε να αντισταθμιστεί η

φυσική εξασθένηση της ενέργειας, που παρατηρείται στο φάσμα του, κατά την αύξηση της συχνότητας. Συνήθως χρησιμοποιείται ένα 1^{ης} τάξεως *FIR* φίλτρο. Η χαρακτηριστική εξίσωση ενός *FIR* φίλτρου n τάξεως είναι η εξής:

$$\bar{s}(n) = (a_1 \cdot s[n]) - (a_2 \cdot s[n-1]) - \dots (a_p \cdot s[n-p]) \quad 7.1$$



Σχ. 7.2 Το διάγραμμα ροής ενός *FIR* φίλτρου

7.1.2. Εφαρμογή παραθύρου (*Windowing*)

Μετά την εφαρμογή του φίλτρου, συνήθως, χρησιμοποιείται ένα παράθυρο εξομάλυνσης με σκοπό να μειωθούν οι ασυνέχειες που υπάρχουν στο πρωτογενές σήμα. Ουσιαστικά ένα παράθυρο αποτελεί μία αριθμητική ακολουθία η οποία πολλαπλασιάζεται με το σήμα μας, με σκοπό να το ομαλοποιήσει.

Οι σημαντικότεροι τύποι παραθύρων παρατίθενται στο *παράρτημα Α*²⁴

²⁴ <http://www.wikipedia.org>

7.2. ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

Η βασική ιδέα της φάσης της εξαγωγής χαρακτηριστικών είναι να αναπαραστήσουμε το σήμα με την μορφή παραμετρικών μεγεθών και να το συμπιέσουμε εξάγοντας μόνο την σημαντική πληροφορία από αυτό. Τα χαρακτηριστικά που θεωρούνται σημαντικά πρέπει να πληρούν τις παρακάτω προδιαγραφές:

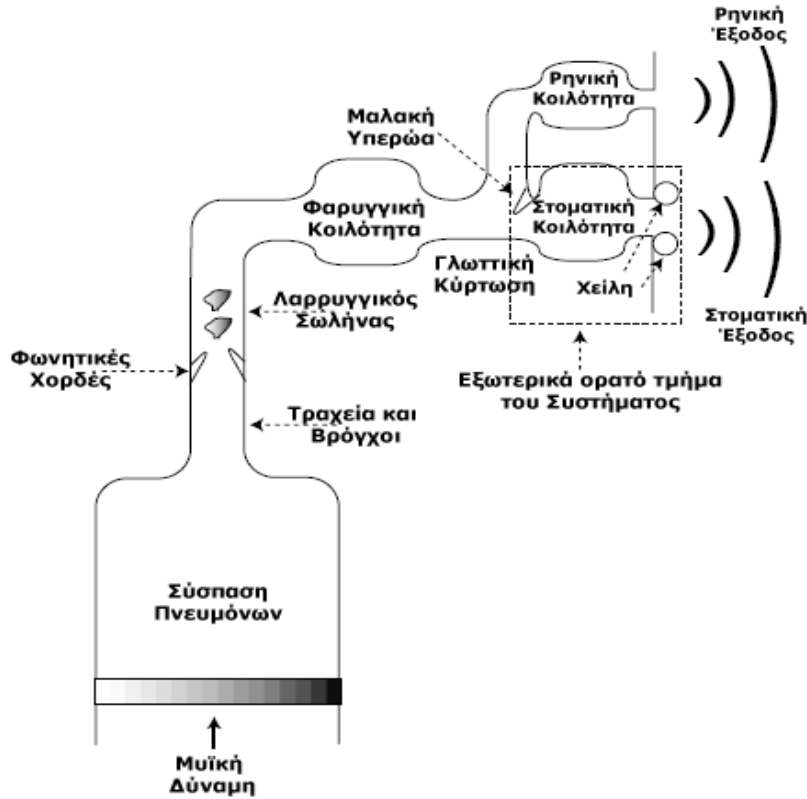
- Μεγάλη inter - speaker ποικιλομορφία
- Χαμηλή intra - speaker ποικιλομορφία
- Εύκολο στη μέτρηση
- Δύσκολο στο να παραποιηθεί και να συγκαλυφθεί
- Ανθεκτικό στην παραμόρφωση και στον θόρυβο
- Όσο το δυνατό ανεξάρτητο από τα άλλα χαρακτηριστικά

Στην πραγματικότητα όμως δεν υπάρχει χαρακτηριστικό που να πληρεί όλες αυτές τις προδιαγραφές, ταυτόχρονα.

7.2.1. Συντελεστές γραμμικής πρόβλεψης - *Linear Prediction Coefficients (LPC)*

Τα τελευταία χρόνια, η *LPC* έχει αποτελέσει μια από τις πιο ισχυρές τεχνικές όσον αφορά την ανάλυση ομιλίας. Στην ουσία αυτή η τεχνική είναι η βάση κάποιων πιο πρόσφατων και πολύπλοκων αλγορίθμων οι οποίοι χρησιμοποιούνται για τον υπολογισμό παραμέτρων του λόγου, όπως, για παράδειγμα, το τονικό ύψος, τα formants, το φάσμα κ.λ.π. Με βάση την θεωρία της γραμμικής πρόβλεψης, οι τιμές των νέων δειγμάτων φωνής, μπορούν να εκτιμηθούν (προβλεφθούν) από τα προηγούμενα χρησιμοποιώντας ένα γραμμικό συνδυασμό, η ισοδύναμα, ή ένα *FIR* (finite impulse response) φίλτρο.

Η μέθοδος γραμμικής πρόβλεψης απαιτεί τον υπολογισμό των συντελεστών a_k από δεδομένα φωνής.



Σχ. 7.1 Ανθρώπινο σύστημα παραγωγής φωνής

Το μοντέλο γραμμικής πρόβλεψης και η ανάλυση γραμμικής πρόβλεψης (linear prediction analysis) βασίζεται στο ψηφιακό μοντέλο παραγωγής της φωνής: $H(z) = G(z) V(z)$

όπου:

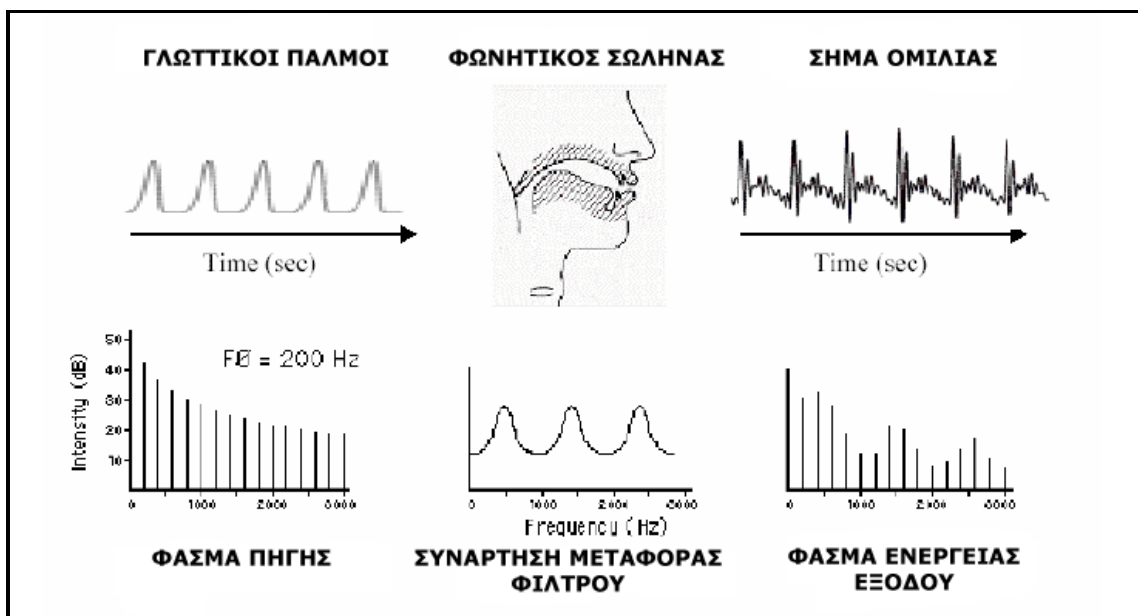
$$G(z) = \text{απόκριση γλωττικού παλμού} = \frac{1}{(1 - cz^{-1})^2}, c \approx \Gamma$$

$$V(z) = \text{συνάρτηση μεταφοράς φωνητικού σωλήνα} = \frac{1}{\prod_{k=1}^{p'} (1 - c_k z^{-1})^2}$$

G = κέρδος

Συνολικά, η συνάρτηση μεταφοράς του φωνητικού συστήματος μπορεί να προσεγγισθεί με

$$H(z) = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (7.2)$$



Σχ. 7.3 Μοντέλο παραγωγής φωνής

Με βάση το μοντέλο αυτό, το σήμα της φωνής στο πεδίο του χρόνου δίδεται από την σχέση

$$s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n) \quad (7.3)$$

Το σφάλμα πρόβλεψης (prediction error) δίδεται από την σχέση

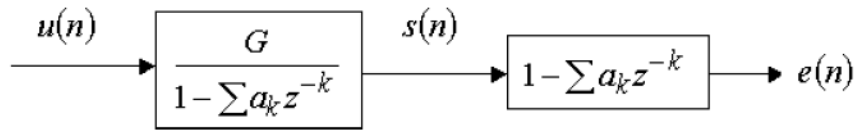
$$e(n) = s(n) - \bar{s}(n) = s(n) - \sum_{k=1}^p a_k s(n-k) \quad (7.4)$$

Μπορεί, δηλαδή, να υπολογιστεί από το σήμα της φωνής με την χρήση του φίλτρου σφάλματος πρόβλεψης (prediction error filter):

$$A(z) = 1 - \sum_{k=1}^p a_k z^{-k} \quad (7.5)$$

Το φίλτρο $A(z)$ ονομάζεται και αντίστροφο φίλτρο (inverse filter) του μοντέλου παραγωγής και έτσι έχουμε το σύστημα που φαίνεται στο σχήμα 7.4. Οι συντελεστές a_k είναι, στην γενική περίπτωση, δια-

φορετικοί από τους πραγματικούς συντελεστές του μοντέλου παραγωγής.



Σχ. 7.4 Υπολογισμός σφάλματος γραμμικής πρόβλεψης με αντίστροφο φίλτρο, όπου $u(n)$ η πηγή διέγερσης.

Ελαχιστοποίηση Τετραγωνικού Σφάλματος²⁵

Οι συντελεστές a_k μπορούν να υπολογιστούν έτσι ώστε να ελαχιστοποιηθεί το συνολικό τετραγωνικό σφάλμα πρόβλεψης:

$$\min_{a_1, \dots, a_p} \sum_n e^2(n) \quad (7.6)$$

όπου το συνολικό τετραγωνικό σφάλμα ελαχιστοποιείται σε κάποιο παράθυρό του σήματός της φωνής.

Επίσης, άλλη μια παραμετροποίηση που βασίζεται στην *LPC*, μπορεί να διεξαχθεί χρησιμοποιώντας το *cepstrum*, το οποίο μπορεί εύκολα να υπολογιστεί μια απλή αναδρομή όπως:

$$c_n = -a_n + \frac{1}{n} \sum_{i=1}^{n-1} (n-i) a_i c_{n-i} \quad (7.7)$$

7.2.2. Συντελεστές *CEPSTRUM*

Το *cepstrum* είναι το αποτέλεσμα της εφαρμογής μετασχηματισμού Fourier στο συχνοτικό φάσμα ενός σήματος.

$$C(q) = 2T \{ \ln |G(z)| + \ln |V(z)| \} \quad (7.8)$$

²⁵ Μπλέτσας Α.Α., «Ελληνικό Σύστημα Αυτόματης Εκφώνησης Κειμένου (*Text To -Speech System*)», §2.4 “LPC: Γενική Αρχή – μέθοδος ανάλυσης της φωνής”, Διπλωματική εργασία, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Πολυτεχνική Σχολή, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Νοέμβριος 1998

Όπου

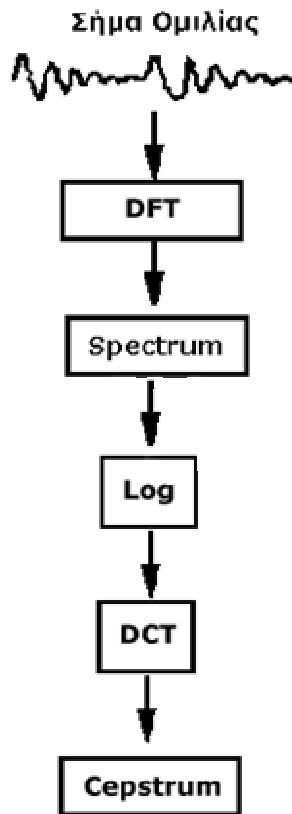
$C(q)$ = το cepstrum

$G(z)$ = απόκριση γλωπτικού παλμού

$V(z)$ = συνάρτηση μεταφοράς φωνητικού σωλήνα

$T\{\}$ = Μετασχηματισμός - Συνήθως DCT (αντίστροφος FFT)

Όπως προαναφέρθηκε, το σήμα ομιλίας παράγεται μέσω της γλωπτικής διέγερσης σε συνδυασμό με το «φιλτράρισμα» από το σύστημα του φωνητικού σωλήνα. Οπότε, κατά την παραγωγή του cepstrum προκύπτουν δύο ειδών συντελεστές, οι αργοί και οι γρήγοροι (slow και fast αντίστοιχα). Οι μεν αργοί αντιπροσωπεύουν τις αργά μεταβαλλόμενες παραμέτρους του φωνητικού σωλήνα και οι δε γρήγοροι το γρήγορα μεταβαλλόμενο σήμα διέγερσης. Για την διαδικασία της αντιστοίχισης προτύπων, επιλέγονται οι αργοί ως συντελεστές που θα τεθούν υπό σύγκριση, διότι είναι αυτοί που καθορίζουν την διαφορετικότητα ανάμεσα σε ξεχωριστούς ομιλητές.



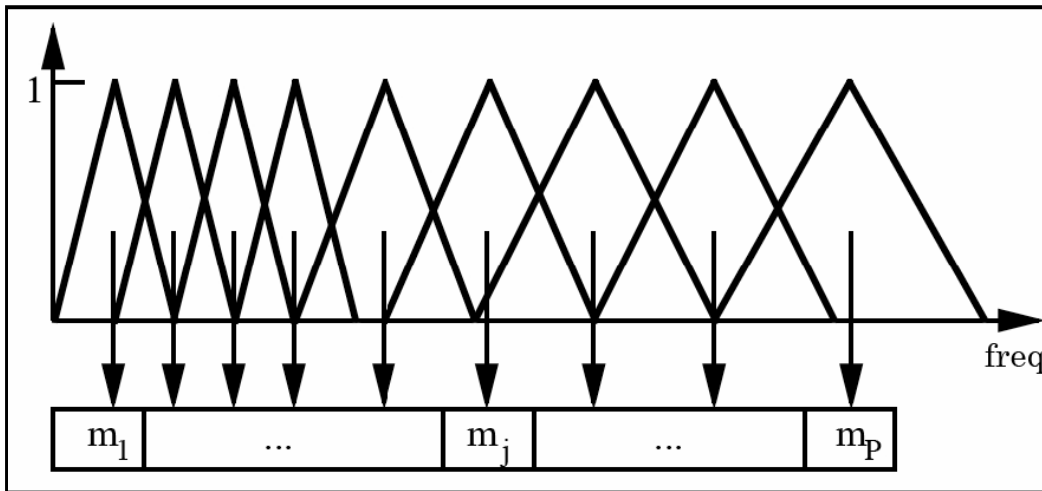
Σχ.7.5 Η διαδικασία παραγωγής του Cepstrum

7.2.3. Mel – Frequency Cepstral Coefficients (MFCC)

Mel cepstrum

Επιστημονικές έρευνες έχουν αποδείξει ότι το ανθρώπινο αυτί δεν αντιλαμβάνεται γραμμικά τις συχνότητες και είναι πειραματικά αποδεδειγμένο ότι η χρήση ενός όμοιου, με την ανθρώπινη ακοή, μοντέλου για την κωδικοποίηση ενός σήματος, μπορεί να επιφέρει μεγάλη βελτίωση στην απόδοση του συστήματος αναγνώρισης. Για την εξαγωγή των MFCCs χρησιμοποιούμε την Filter bank analysis. Αυτή η διαδικασία είναι όμοια με το αποτέλεσμα του μετασχηματισμού Fourier αλλά η κύρια διαφορά είναι ότι οι ζώνες συχνότητας είναι λογαριθμικά τοποθετημένες αντί για γραμμικά. Η λογαριθμική κλίμακα ονομάζεται Mel κλίμακα και προσδιορίζεται ως εξής²⁶:

$$Mel(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (7.9)$$



Σχ. 7.6 MEL Filter Bank

Η τράπεζα φίλτρων (Filter bank), που προαναφέρθηκε, υλοποιείται παίρνοντας τον μετασχηματισμό Fourier για ένα παράθυρο, ενός ηχητικού δείγματος και, έπειτα, πολλαπλασιάζοντας τους συντελεστές έντασης, που προκύπτουν, χρησιμοποιώντας τριγωνικά φίλτρα. Συνήθως χρησιμοποιούμε τις παραμέτρους του cepstrum (MFCCs) οι

²⁶ Διγαλάκης Βασίλης. Σημειώσεις μαθήματος: «Εισαγωγή στην επεξεργασία Φωνής» Τμήμα Ηλεκτρολόγων Μηχανικών Και Μηχανικών Υπολογιστών. Πανεπιστήμιο Κρήτης

οποίοι μπορούν να υπολογιστούν λογαριθμώντας τα πλάτη – εντάσεις, m_j , της τράπεζας φίλτρων και εφαρμόζοντας, έπειτα, μετασχηματισμό διακριτού συνημίτονου (DCT).

$$c_i = \sqrt{\frac{2}{N} \sum_{j=1}^N m_j \cos\left(\frac{\pi i}{N}(j-0.5)\right)} \quad (7.10)$$

7.3. ΜΕΘΟΔΟΙ ΑΝΤΙΣΤΟΙΧΙΣΗΣ ΠΡΟΤΥΠΩΝ – ΜΟΝΤΕΛΟΠΟΙΗΣΗ

7.3.1. Δυναμική αναδίπλωση στο χρόνο. *Dynamic Time Warping (DTW)*

Η τεχνική *DTW* αποτελεί έναν αλγόριθμο που υπολογίζει την ομοιότητα, μεταξύ δύο ακολουθιών, οι οποίες μπορεί να ποικίλουν, όσον αφορά την εξέλιξή τους στον χρόνο και την ταχύτητά τους. Η ιδέα της δυναμικής ευθυγράμμισης του χρόνου έχει χρησιμοποιηθεί για να λύσει προβλήματα που σχετίζονται με την σύγκριση της φασματικής ακολουθίας του λόγου. Όσον αφορά τα σήματα ομιλίας, τα διάφορα χαρακτηριστικά μοτίβα του ίδιου φωνητικού δείγματος, σπάνια πραγματοποιούνται με τον ίδιο ρυθμό κατά την διάρκεια της ομιλίας. Αυτό όμως δεν πρέπει να χαρακτηρίζεται σαν ανομοιότητα γιατί με αυτόν τον τρόπο οδηγούμαστε σε εσφαλμένο συμπέρασμα. Στόχος είναι η κανονικοποίηση (normalizing) των διακυμάνσεων του ρυθμού ομιλίας έτσι ώστε να μπορούμε να συγκρίνουμε τα διάφορα φωνητικά δείγματα με συναφή τρόπο.

Τα συστήματα που λειτουργούν με την μέθοδο *DTW* προσδιορίζουν μία ακουστική είσοδο, χρησιμοποιώντας ένα προκαθορισμένο σύνολο διανυσμάτων, τα πρότυπα, που προκύπτουν από τις παραμέτρους του σήματος. Η αντιστοίχιση πραγματοποιείται υπολογίζοντας την Ευκλείδεια απόσταση μεταξύ των προκαθορισμένων προτύπων και των προτύπων που τίθενται προς αναγνώριση.²⁷

7.3.2. Διανυσματικός κβαντισμός - *Vector Quantization (VQ)*

Ο διανυσματικός κβαντισμός είναι μία διαδικασία κατά την οποία ένα συνεχές σήμα προσεγγίζεται ως μια ψηφιακή αναπαράσταση,

²⁷ Joseph A. Kogan and Daniel Margoliash “Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study” Journal of the Acoustical Society of America 103 (4), 2185-2196, January 1998

που περιλαμβάνει ένα σύνολο παραμέτρων, με σκοπό την μοντελοποίηση ενός ολοκληρωμένου προτύπου (διάνυσμα).

Ένας διανυσματικός κβαντιστής καθορίζεται πλήρως από ένα codebook, το οποίο αποτελείται από ένα σύνολο σταθερών, πρότυπων διανυσμάτων. Η VQ μειώνει τον πλεονασμό δεδομένων και αυτό προκαλεί, αναπόφευκτα, την παραμόρφωση μεταξύ των αρχικών σημάτων. Ένα βασικό πρόβλημα στην τεχνική VQ είναι η ελαχιστοποίηση της παραμόρφωσης που εισάγεται. Η διαδικασία VQ περιλαμβάνει²⁸:

- την εκτίμηση της παραμόρφωσης
- την παραγωγή συγκεκριμένου αριθμού πρότυπων διανυσμάτων.

Δύο τυπικές τεχνικές VQ που ελαχιστοποιούν την παραμόρφωση που δημιουργείται, αναφέρονται παρακάτω.

Έστω:

$$x = (x_1, x_2, \dots, x_d) \in R^d \quad (7.11)$$

ένα d-διάστατο διάνυσμα.

Στη διαδικασία VQ, το διάνυσμα x αντιστοιχίζεται σε ένα άλλο d-διάστατο διάνυσμα z με πραγματικές τιμές

$$z = q(x) \quad (7.12)$$

όπου q είναι ο παράγοντας κβαντοποίησης.

Γενικά, το z είναι ένα στοιχείο ενός πεπερασμένου συνόλου τιμών

$$Z = \{z_i, 1 \leq i \leq L\} \quad (7.13)$$

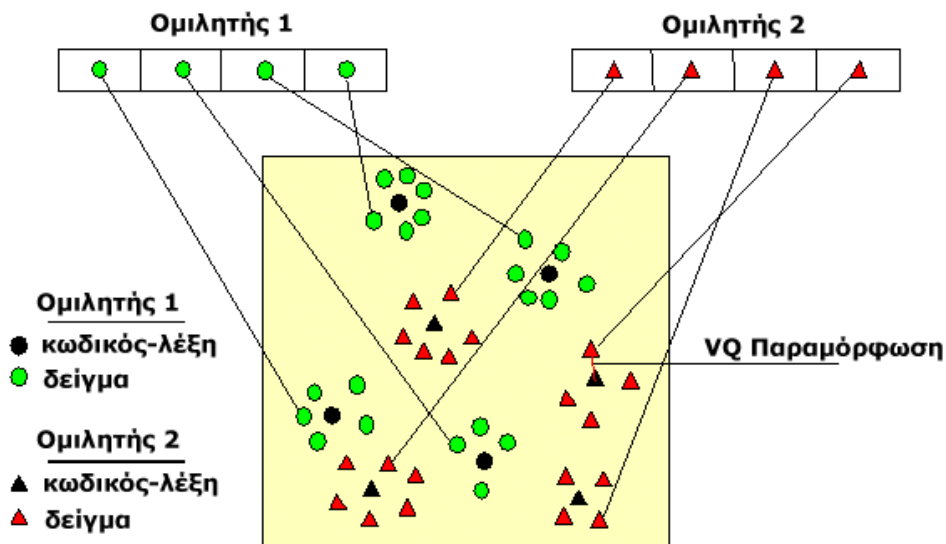
όπου

²⁸ Περσικάκης Μανόλης, «Κωδικοποίηση φωνής για αναγνώριση σε ενσύρματα και ασύρματα δίκτυα δεδομένων (Διαδίκτυο, GPRS) μέσα από διάλογο 2 Kbps», Πολυτεχνείο Κρήτης, τμήμα ηλεκτρολόγων μηχανικών και μηχανικών υπολογιστών, εργαστήριο τηλεπικοινωνιών πληροφορίας και δικτύων, Διπλωματική εργασία, Δεκέμβριος 2000, §2.5, §2.6

$$z_i = (z_{i1}, z_{i2}, \dots, z_{id}).$$

Το σύνολο Z αναφέρεται ως "codebook", L είναι το μέγεθός του, και $\{z_i\}$ είναι το σύνολο των κωδικών-λέξεων.

Η διαδικασία δημιουργίας του κβαντιστή είναι επίσης γνωστή ως διαδικασία εκπαίδευσης.



Εικ. 7.7 Η διαδικασία αντιστοίχισης προτύπων με τη μέθοδο VQ . Οι κύκλοι αντιστοιχούν στα ακουστικά διανύσματα του ομιλητή 1 ενώ τα τρίγωνα στον ομιλητή 2. Στην φάση της εκπαίδευσης αναπαράγεται ένα VQ codebook, για κάθε γνωστό ομιλητή, ομαδοποιώντας τα ακουστικά του/της διανύσματα εκπαίδευσης. Οι κωδικοί - λέξεις (codewords ή centroids) αναπαρίστανται στο σχήμα με μαύρους κύκλους και μαύρα τρίγωνα για τον πρώτο και τον δεύτερο ομιλητή αντίστοιχα. Η απόσταση από ένα διάνυσμα στον κοντινότερο κωδικό - λέξη, ενός codebook, ονομάζεται VQ παραμόρφωση. Στην φάση της αναγνώρισης, ένα φωνητικό δείγμα εισόδου, ενός άγνωστου ομιλητή, «κβαντίζεται» διανυσματικά και υπολογίζεται η ολική VQ παραμόρφωση. Με αυτόν τον τρόπο αναγνωρίζεται ο ομιλητής που αντιστοιχεί στο VQ codebook με την μικρότερη ολική παραμόρφωση.

Εκπαίδευση Κβαντιστή Διανύσματος

Ο d -διάστατος χώρος του αρχικού τυχαίου διανύσματος x χωρίζεται σε L περιοχές ή blocks

$$\{C_i, 1 \leq i \leq L\} \tag{7.14}$$

και το κάθε block αντιστοιχίζεται με ένα διάνυσμα z_i . Ο κβαντιστής τότε αναθέτει την κωδική - λέξη z_i , ώστε,

$$q(x) = z_i \quad (7.15)$$

αν το x ανήκει στο C_i .

Οποιοδήποτε δάνυσμα εισόδου, X , που ανήκει στο block C κβαντοποιείται ως z_i . Η μορφή κάθε block μπορεί να είναι διαφορετική και οι θέσεις των κωδικών - λέξεων που αντιστοιχούν στα blocks, καθορίζεται από την ελαχιστοποίηση της μέσης παραμόρφωσης που εισάγεται από τα αντίστοιχα blocks.

Για να αξιολογηθεί το σφάλμα κβαντοποίησης μεταξύ του x και του z , είναι απαραίτητο να καθορισθεί ένα μέτρο παραμόρφωσης

$$D(x, z) = \frac{1}{x} \sum_{i=1}^N \min_j \|x_i - z_j\|^2 \quad (7.16)$$

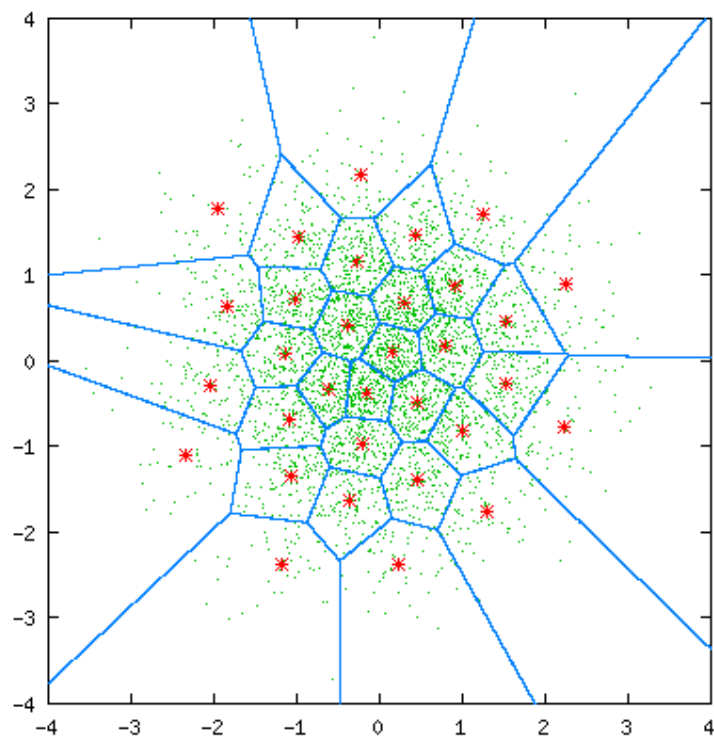
ώστε να μετρηθεί η ποιότητα κβαντοποίησης. Όσο μικρότερο είναι το μέτρο παραμόρφωσης, τόσο περισσότερο ταιριάζει το x με το z . Το μέτρο παραμόρφωσης μεταξύ του x και του z είναι, επίσης γνωστό, ως μέτρο απόστασης. Το μέτρο πρέπει να είναι εντοπίσιμο και υπολογίσιμο ώστε να είναι δυνατή η ανάλυσή του, και οι διαφορές στις τιμές της παραμόρφωσης να συσχετίζονται με την ποιότητα της κβαντοποίησης.

Επίσης, το σταθμισμένο μέσο τετραγωνικό σφάλμα της παραμόρφωσης μπορεί να χρησιμοποιηθεί έτσι, ώστε οι παραμορφώσεις που εισάγονται με την κβαντοποίηση διαφορετικών παραμέτρων να είναι ίσες. Τα άνισα βάρη μπορούν να χρησιμοποιηθούν για να επιτρέψουν σε ορισμένα στοιχεία να είναι σημαντικότερα από άλλα. Μια δημοφιλής επιλογή για άνισα βάρη είναι να χρησιμοποιηθεί ο αντίστροφος του πίνακα διακύμανσης του z

$$d(x, z) = (x - z)' \Sigma^{-1} (x - z) \quad (7.17)$$

όπου Σ ο πίνακας διακύμανσης του z . Η μέτρηση της παραπάνω παραμόρφωσης μπορεί να απλοποιηθεί σε τετραγωνικό σφάλμα παραμόρφωσης,

$$d(x, \hat{x}) = \|x, \hat{x}\| = \sum_{i=0}^{d-1} (x_i - z_i)^2 \quad (7.18)$$



Εικ. 7.8 Το μοντέλο της Διανυσματικής Κβαντοποίησης (codebook). Τα ομαδοποιημένα διανύσματα των χαρακτηριστικών (feature vectors) αναπαρίστανται με πράσινες κουκίδες και κάθε ομάδα αποτελεί ένα κύτταρο (cell). Τα διανύσματα που βρίσκονται μέσα σε ένα κύτταρο, παίρνουν την τιμή του αντίστοιχου κωδικού λέξη (centroid ή code word) που αναπαρίσται με κόκκινη κουκίδα.

Το Πρόβλημα Της Αναγνώρισης

Ορίζοντας το πρόβλημα της αναγνώρισης (αποκωδικοποίησης), λέμε ότι ζητείται να καθοριστεί, με βάση κάποιο κριτήριο, ότι «εστάλη» (προφέρθηκε) η ακολουθία λέξεων W , δεδομένης της ακολουθίας των διανυσμάτων X στην είσοδο του αποκωδικοποιητή. Οι στατιστικές μέθοδοι αναγνώρισης προϋποθέτουν την ύπαρξη κάποιου αντίστοιχου στατιστικού μοντέλου για τον υπολογισμό της πιθανότητας (ή συνάρτηση πιθανοφάνειας). Πρόκειται για το μέγεθος $P(X/W)$. Επιπλέον, ως κριτήριο αποκωδικοποίησης, όπως και σε ένα τυπικό ψηφιακό τηλεπικοινωνιακό σύστημα, είναι η ελαχιστοποίηση πιθανότητας σφάλματος. Με βάση το μοντέλο $P(X|W)$, η πιθανότητα σφάλματος ελαχιστοποιείται αν αποκωδικοποιήσουμε την ακολουθία, εκείνη, W , για την οποία μεγιστοποιείται η a - posteriori πιθανότητα, δεδομένου ότι ο αποκωδικοποιητής «έλαβε» την ακολουθία $\underline{X} = [x_1, x_2, \dots, x_T]$ ²⁹

²⁹ Διγαλάκης Βασίλης. Σημειώσεις μαθήματος: «Εισαγωγή στην επεξεργασία Φωνής» Τμήμα Ηλεκτρολόγων Μηχανικών Και Μηχανικών Υπολογιστών. Πανεπιστήμιο Κρήτης

Χρησιμοποιώντας τον κανόνα του Bayes έχουμε:

$$\hat{W} = \arg \max_{\underline{w}} P(\underline{W} | \underline{X}) = \arg \max_{\underline{w}} \frac{P(\underline{W})P(\underline{X} | \underline{W})}{P(\underline{X})} \quad (7.19)$$

Όπου το $\arg\max$ συμβολίζει το όρισμα που μεγιστοποιεί την αντίστοιχη ποσότητα. Αυτή η εξίσωση δείχνει ότι για να βρεθεί η πιο πιθανή ακολουθία λέξεων \underline{W} , πρέπει να βρεθεί η ακολουθία που μεγιστοποιεί το γινόμενο του $P(\underline{W})$ και του $P(\underline{X} | \underline{W})$. Ο πρώτος από αυτούς τους όρους, $P(\underline{W})$, υπολογίζει την a-priori πιθανότητα της παρατήρησης \underline{W} , ανεξάρτητα από το σήμα που παρατηρήθηκε, με βάση κάποιο στατιστικό μοντέλο, και η πιθανότητα αυτή είναι γνωστή ως γλωσσικό μοντέλο (language model). Ο δεύτερος όρος $P(\underline{X} | \underline{W})$ αναπαριστά την πιθανότητα εμφάνισης μιας ακολουθίας διανυσμάτων \underline{X} δεδομένων μερικών ακολουθιών λέξεων \underline{W} , και αυτή η πιθανότητα είναι γνωστή ως ακουστικό μοντέλο (acoustic model).

- **Εκ των προτέρων πιθανότητα P(A) (Prior probability):**
Η πιθανότητα ενός γεγονότος πριν εξετάσουμε κάποια άλλη πληροφορία.
- **Εκ των υστέρων πιθανότητα P(B|A) - Posterior probability:** η πιθανότητα ενός γεγονότος εξετάζοντας περαιτέρω πληροφορίες που σχετίζονται με αυτό.

Ας θεωρήσουμε την ακολουθία λέξεων $W = \ll \text{τμήμα μουσικής τεχνολογίας και ακουστικής} \gg$. Τότε, το γλωσσικό μοντέλο θα υπολόγιζε την $P(\underline{W})$ ως εξής :

$$\begin{aligned} P(\underline{W}) &= P(\ll \text{τμήμα μουσικής τεχνολογίας και } | \text{ ακουστικής} \gg) = \\ &= P(\ll \text{τμήμα} \gg) \cdot P(\ll \text{μουσικής} \gg | \ll \text{τμήμα} \gg) \cdot P(\ll \text{τεχνολογίας} \gg | \ll \text{τμήμα μουσικής} \gg) \cdot P(\ll \text{και} \gg | \ll \text{τμήμα μουσικής τεχνολογίας} \gg) \cdot P(\ll \text{ακουστικής} \gg | \ll \text{τμήμα μουσικής τεχνολογίας και} \gg) \end{aligned}$$

Η γλωσσική μονάδα που μοντελοποιείται είναι, συνήθως, η λέξη. Για να υπάρχει η δυνατότητα γενίκευσης και να μοντελοποιούνται οι λέξεις που δεν παρατηρήθηκαν στα δεδομένα εκπαίδευσης, χρησιμοποιούνται μικρότερες γλωσσικές μονάδες, όπως είναι το φώνημα (phoneme) ή η συλλαβή. Η όλη διαδικασία ονομάζεται σχεδίαση του ακουστικού μοντέλου. Η κάθε λέξη μετατρέπεται σε μία ακολουθία βασικών ήχων, τα φωνήματα, χρησιμοποιώντας ένα λεξικό προφορών (dictionary). Το λεξικό προφορών είναι ένα αρχείο που περιέχει τις αποδόσεις όλων των λέξεων, που περιέχονται στην γραμματική και πρέπει να συνταχθεί ώστε να περιγράφει ακριβώς τις προφορές

των λέξεων ακόμη και με περισσότερους, του ενός, τρόπους. Η δομή ενός τέτοιου λεξικού φαίνεται στο παρακάτω παράδειγμα:

άγνωστες	a j n o s t e s
άνθρωπος	a+ n t h r o p o s
δυναμικών	d h I n a m I k o+ n
κωνσταντίνος	k o n s t a d i+ n o s
κωσταντίνος	k o s t a d i+ n o s
νοικιασμένο	n I k 2 y a s m e+ n o

7.3.3. Μοντέλα γκαουσιανών μειγμάτων - *Gaussian Mixture Models (GMM)*

Πυκνότητα Πιθανότητας

Η συνάρτηση πυκνότητας πιθανότητας (probability density function - pdf) περιγράφει την πιθανότητα, μια παρατήρηση να ισχύει μέσα σε ένα συγκεκριμένο πεδίο.

$$\Pr(a < x < b) = \int_a^b f(x)dx \quad (7.20)$$

Για ένα D – διάστατο διάνυσμα χαρακτηριστικών, x , η πυκνότητα μείγματος που χρησιμοποιείται για την συνάρτηση της πιθανότητας ορίζεται ως:

$$p(x | \lambda) = \sum_{i=1}^M b_i p_i(x) \quad (7.21)$$

όπου M = αριθμός των επιμέρους Γκαουσιανών
 b_i = στάθμη των συνιστωσών του μείγματος
 p_i = πυκνότητα πιθανότητας των συνιστωσών του μείγματος

Η ολοκληρωμένη πυκνότητα του Γκαουσιανού μείγματος, παραμετροποιείται από:

- Από τους μέσους των διανυσμάτων μ_i
- Από τους πίνακες συνδιακυμάνσεων Σ_i
- Από τις στάθμες του μείγματος b_i

Οι παράμετροι αυτοί αναπαριστώνται από το μέγεθος

$$\lambda = \{\mu_i, \Sigma_i, b_i\} \quad i=1, \dots, M \quad (7.22)$$

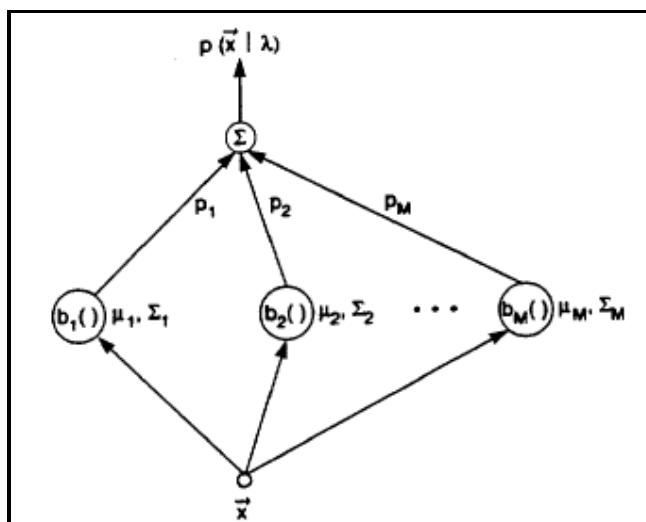
Στην αναγνώριση ομιλητή, κάθε ομιλητής, εκπροσωπείται από ένα *GMM* και αναφέρεται σε αυτό με το μοντέλο λ .

Λόγοι Επιλογής Γκαουσιανών Μοντέλων

Υπάρχουν δύο κύριοι λόγοι που χρησιμοποιούμε πυκνότητες Γκαουσιανών μειγμάτων ως αναπαράσταση της ταυτότητας του ομιλητή. Ο πρώτος λόγος είναι ότι οι επιμέρους πυκνότητες των συστατικών μιας Γκαουσιανής πυκνότητας, με πολλαπλές κορυφές, όπως το *GMM*, μπορούν να μοντελοποιήσουν μερικά σημαντικά σύνολα ακουστικών χαρακτηριστικών.

Είναι λογικό να συμπεράνουμε ότι το ακουστικό φάσμα που αντιπροσωπεύει στην φωνή ενός ομιλητή, μπορεί να αναλυθεί από ένα σύνολο ακουστικών, ομάδων χαρακτηριστικών, οι οποίες αντιπροσωπεύουν κάποια γενικά, φωνητικά γεγονότα, όπως είναι τα φωνήεντα, τα ένρινα και τα σύμφωνα. Οι ομάδες αυτές, προκύπτουν από την διαμόρφωση του φωνητικού σωλήνα, που ποικίλει από ομιλητή σε ομιλητή και είναι χρήσιμες για τον χαρακτηρισμό της ταυτότητας ενός ομιλητή. Το «φασματικό σχήμα» μιας ακουστικής ομάδας, μπορεί να αναπαρασταθεί από τον μέσο όρο, μ_i , της πυκνότητας του εκάστοτε συστατικού του μείγματος, και η διαμόρφωση του γενικού, φασματικού σχήματος μπορεί να αναπαρασταθεί από τον πίνακα συνδιακυμάνσεων Σ_i .³⁰

³⁰ Douglas A. Reynolds, Richard C. Rose “*Robust Text-Independent Speaker Identification Using Gaussian Mixture Models*” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1., January 1995 §II.C “Model Interpretations”



Σχ. 7.11 Απεικόνιση της πυκνότητας ενός Γκαουσιανού μείγματος που αποτελείται από M συστατικά. Μια πυκνότητα Γκαουσιανού μείγματος αποτελεί ένα σταθμισμένο άθροισμα επιμέρους Γκαουσιανών πυκνοτήτων όπου το $p_i, i=1, \dots, M$, είναι οι στάθμες του μείγματος και $b_i(\cdot), i=1, \dots, M$, είναι η Γκαουσιανές των συστατικών.

Ο δεύτερος λόγος που χρησιμοποιούμε τις πυκνότητες Γκαουσιανών μειγμάτων, για την αναγνώριση ομιλητή, είναι η εμπειρική διαπίστωση ότι ένας γραμμικός συνδυασμός Γκαουσιανών συναρτήσεων, είναι ικανός να αναπαραστήσει μία μεγάλη ομάδα κατανομών από δείγματα. Ένα από τα πιο σημαντικά γνωρίσματα των *GMM* είναι η ικανότητά τους να εξομαλύνουν πυκνότητες με απροσδιόριστο σχήμα. Το κλασικό Γκαουσιανό μοντέλο ομιλητή, με μία κορυφή, αναπαριστά την κατανομή χαρακτηριστικών ενός ομιλητή με ένα σημείο (μέσο διάνυσμα) και ένα ελλειπτικό σχήμα (πίνακας συνδιακυμάνσεων), ενώ ένα *VQ* μοντέλο αναπαριστά την κατανομή ενός ομιλητή από ένα διακριτό σύνολο προτύπων. Κατά μία έννοια, το *GMM* αποτελεί ένα υβρίδιο μεταξύ των δύο αυτών μοντέλων, χρησιμοποιώντας ένα διακριτό σύνολο από Γκαουσιανές συναρτήσεις που η κάθε μια έχει το δικό της μέσο και πίνακα συνδιακυμάνσεων επιτρέποντας, έτσι, την καλύτερη, δυνατή, μοντελοποίηση.³¹

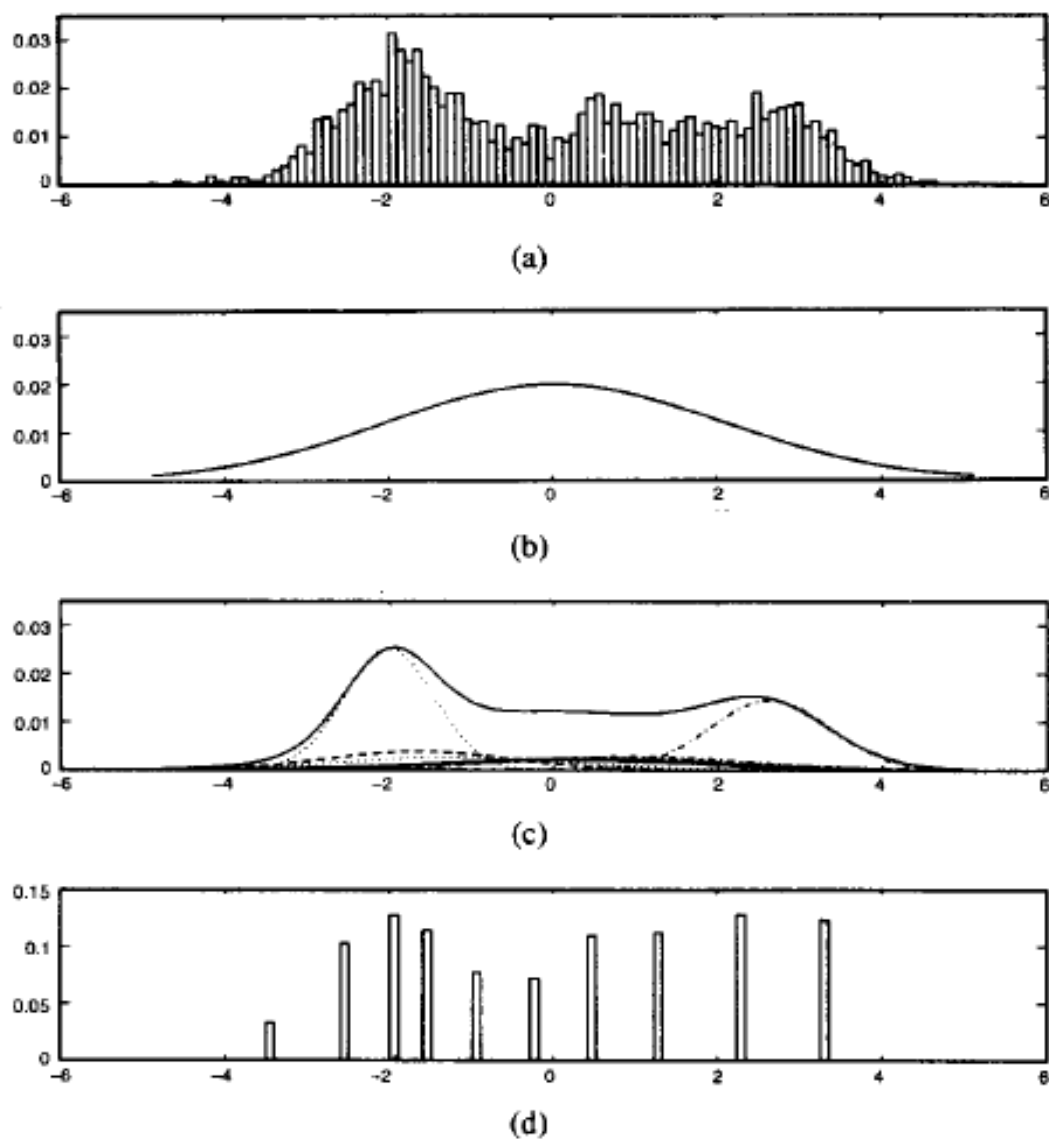
Στο Σχ.7.12 συγκρίνονται οι πυκνότητες που σχηματίζονται χρησιμοποιώντας ένα Γκαουσιανό μοντέλο με μια κορυφή, ένα *GMM* και ένα *VQ* μοντέλο. Το διάγραμμα (a) δείχνει το ιστόγραμμα από έναν *cepstral* συντελεστή, που εξάγεται από ομιλία 25 δευτερολέπτων, το διάγραμμα (b) δείχνει το Γκαουσιανό μοντέλο με μία κορυφή, το διάγραμμα (c) δείχνει το *GMM* και τις δέκα επιμέρους πυκνότητες

³¹ Douglas A. Reynolds, Richard C. Rose “Robust Text-Independent Speaker Identification Using Gaussian Mixture Models” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1, January 1995 §II.C “Model Interpretations”

των συστατικών του και το διάγραμμα (d) δείχνει το ιστόγραμμα των δεδομένων που αντιστοιχίζονται στην περιοχή ενός VQ κυττάρου σε ένα codebook δέκα στοιχείων. Το *GMM* όχι μόνο παρέχει μια ομαλή συνολική κατανομή, αλλά, τα συστατικά του, περιγράφουν επακριβώς την πολυμορφική φύση της πυκνότητας.

Επίσης, εξ' αιτίας του γεγονότος ότι οι επιμέρους Γκαουσιανές ενεργούν συνολικά προκειμένου να μοντελοποιήσουν τη συνολική συνάρτηση πυκνότητας πιθανότητας (pdf), οι πλήρεις πίνακες συνδιακυμάνσεων δεν είναι αναγκαίοι, ακόμα και αν τα χαρακτηριστικά δεν είναι, στατιστικά, ανεξάρτητα. Ο γραμμικός συνδυασμός διαγώνιων Γκαουσιανών συνδιακυμάνσεων, είναι ικανός να μοντελοποιήσει τις συσχετίσεις μεταξύ των στοιχείων των διανυσμάτων χαρακτηριστικών. Το αποτέλεσμα της χρήσης ενός συνόλου από M πλήρεις Γκαουσιανές συνδιακύμανσης μπορεί, ισοδύναμα, να αποκτηθεί χρησιμοποιώντας ένα μεγαλύτερο σύνολο από διαγώνιες Γκαουσιανές συνδιακύμανσης.³²

³² Douglas A. Reynolds, Richard C. Rose “*Robust Text-Independent Speaker Identification Using Gaussian Mixture Models*” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1., January 1995 §II.C “Model Interpretations”



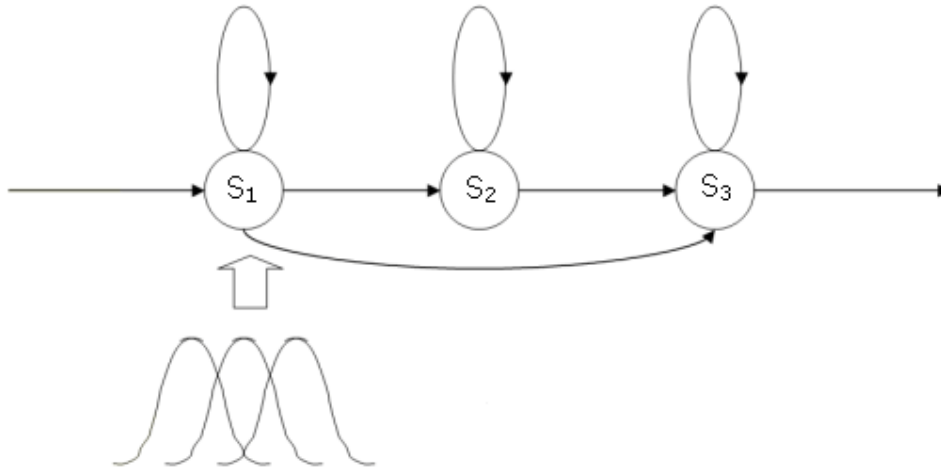
Σχ. 7.12 Σύγκριση μοντελοποιήσεων κατανομών: (a) ιστόγραμμα αποτελούμενο από *cepstral* συντελεστές, που εξάγονται από ομιλία 25 δευτερολέπτων, (b) Γκαουσιανό μοντέλο με μία κορυφή, (c) *GMM* με δέκα επιμέρους πυκνότητες των συστατικών του, (d) ιστόγραμμα των δεδομένων που αντιστοιχίζονται στην περιοχή ενός *VQ* κυττάρου σε ένα codebook δέκα στοιχείων.

Υπολογισμός Μέγιστης Πιθανότητας

Εξάγοντας δεδομένα εκπαίδευσης από μια ομιλία, ο στόχος της εκπαίδευσης του μοντέλου ομιλητή, είναι ο υπολογισμός των λ παραμέτρων του *GMM* που, κατά κάποιο τρόπο, ταιριάζουν περισσότερο με την κατανομή των εκπαιδευτικών διανυσμάτων χαρακτηριστικών. Υπάρχουν διαθέσιμες διάφορες τεχνικές για τον υπολογισμό αυτών των παραμέτρων αλλά η, μακράν, δημοφιλέστερη είναι η μέθοδος του υπολογισμού της μέγιστης πιθανότητας (*maximum likelihood*)

estimation). Ο στόχος αυτής της μεθόδου είναι η ανεύρεση των παραμέτρων του μοντέλου, που μεγιστοποιούν την πιθανότητα των *GMM*, αφού δοθούν τα δεδομένα εκπαίδευσης. Για μία ακολουθία T διανυσμάτων εκπαίδευσης $X = \{\vec{x}_1, \dots, \vec{x}_T\}$, η *GMM* πιθανότητα μπορεί να διατυπωθεί ως εξής³³:

$$p(X | \lambda) = \prod_{t=1}^T p(\vec{x}_t | \lambda) \quad (7.23)$$



Σχ. 7.13 Μοντέλο *GMM*. Κάθε συστατικό μοντελοποιεί μια ακουστική ομάδα.

Αυτή η έκφραση, όμως, είναι μια μη γραμμική συνάρτηση των λ παραμέτρων και δεν είναι δυνατή η άμεση μεγιστοποίηση. Όμως, ο υπολογισμός της μέγιστης πιθανότητας μπορεί να γίνει χρησιμοποιώντας μια ειδική περίπτωση του κυκλικού αλγορίθμου expectation – maximization (EM).

Ξεκινώντας με ένα αρχικό μοντέλο λ , η βασική ιδέα του αλγορίθμου EM, είναι ο υπολογισμός ενός νέου μοντέλου $\bar{\lambda}$, έτσι ώστε $p(X | \bar{\lambda}) \geq p(X | \lambda)$. Έπειτα, το νέο μοντέλο, μετατρέπεται σε αρχικό μοντέλο, για τον επόμενο κύκλο και η διαδικασία επαναλαμβάνεται μέχρι να φτάσουμε σε ένα σημείο σύγκλισης.

Σε κάθε κύκλο εφαρμογής της μεθόδου EM, εφαρμόζονται οι παρακάτω τύποι, οι οποίοι εξασφαλίζουν μια μονοτονική αύξηση της τιμής της πιθανότητας του μοντέλου.

³³ Douglas A. Reynolds, Richard C. Rose “Robust Text-Independent Speaker Identification Using Gaussian Mixture Models” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1, January 1995, §.II D “Maximum Likelihood Parameter Estimation”

Στάθμες Μείγματος:

$$\bar{p}_i = \frac{1}{T} \sum_{t=1}^T p(i | \bar{x}_t, \lambda) \quad (7.24)$$

Μέσοι Όροι:

$$\bar{\mu}_i = \frac{\sum_{t=1}^T p(i | \bar{x}_t, \lambda) \bar{x}_t}{\sum_{t=1}^T p(i | \bar{x}_t, \lambda)} \quad (7.25)$$

Διασπορές:

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T p(i | \bar{x}_t, \lambda) \bar{x}_t^2}{\sum_{t=1}^T p(i | \bar{x}_t, \lambda)} - \bar{\mu}_i^2 \quad (7.26)$$

Τα σ_i^2 , x_t και μ_i αναφέρονται στα αυθαίρετα στοιχεία των διανυσμάτων $\bar{\sigma}_i^2$, \bar{x}_t και $\bar{\mu}_i$ αντίστοιχα.³⁴

Η, εκ των υστέρων, (a posteriori) πιθανότητα για την ακουστική ομάδα i δίνεται από την εξίσωση

$$p(i | \bar{x}_t, \lambda) = \frac{p_i b_i(\bar{x}_t)}{\sum_{k=1}^M p_k b_k(\bar{x}_t)} \quad (7.27)$$

Δύο κρίσιμοι παράγοντες στην εκπαίδευση του *GMM* είναι η επιλογή της τάξης M του μείγματος και η τοποθέτηση των αρχικών τιμών των παραμέτρων του μοντέλου πριν την εφαρμογή του EM αλγορίθμου. Δεν υπάρχει κάποια συγκεκριμένη θεωρία για την τελευταία διαδικασία, έτσι ο προσδιορισμός των παραμέτρων γίνεται πειραματικά.

³⁴ Douglas A. Reynolds, Richard C. Rose “*Robust Text-Independent Speaker Identification Using Gaussian Mixture Models*” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1., January 1995, §.II D “Maximum Likelihood Parameter Estimation”

Αναγνώριση Ομιλητή

Με την αναγνώριση ομιλητή, ένα σύνολο S ομιλητών, $S = \{1, 2, \dots, S\}$ αναπαρίσταται χρησιμοποιώντας τις παραμέτρους $\lambda_1, \lambda_2, \dots, \lambda_S$ των *GMM*. Ο στόχος είναι η ανεύρεση του μοντέλου ομιλητή που έχει την μέγιστη *a posteriori* πιθανότητα για μια δεδομένη ακολουθία.

Δηλαδή

$$\hat{S} = \arg \max_{1 \leq k \leq S} \Pr(\lambda_k | X) = \arg \max_{1 \leq k \leq S} \frac{p(X | \lambda_k) \Pr(\lambda_k)}{p(X)} \quad (7.28)$$

Η δεύτερη εξίσωση εξάγεται από τον νόμο του Bayes. Στην περίπτωση που έχουμε δύο ομιλητές με ίση πιθανότητα (π.χ. $p_r(\lambda_k) = 1/S$) και η $P(X)$ είναι ίδια για όλα τα μοντέλα ομιλητών, η παραπάνω σχέση απλοποιείται σε³⁵:

$$\hat{S} = \arg \max_{1 \leq k \leq S} p(X | \lambda_k) \quad (7.29)$$

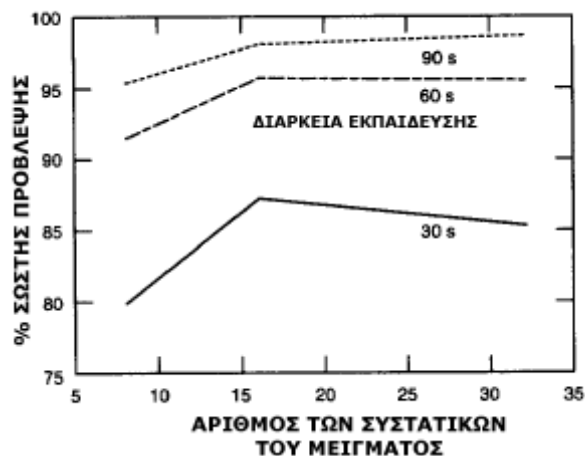
Αποδεικνύεται πειραματικά ότι υπάρχει άμεση συσχέτιση μεταξύ της αποτελεσματικότητας του αυτόματου συστήματος αναγνώρισης ομιλητή και τον αριθμό των συστατικών ενός Γκαουσιανού μείγματος καθώς και της διάρκειας των δειγμάτων εκπαίδευσης και αναγνώρισης.³⁶

Όπως φαίνεται και στα παρακάτω διαγράμματα, κατά την αύξηση της διάρκειας των δειγμάτων που χρησιμοποιούνται κατά την εκπαίδευση του συστήματος, παρουσιάζεται σημαντική αύξηση στο ποσοστό σωστής πρόβλεψης κατά την εξαγωγή του αποτελέσματος. Παρόμοια συσχέτιση υπάρχει μεταξύ του ποσοστού σωστής πρόβλεψης και τον αριθμό των συστατικών του μείγματος, ως ένα σημείο, καθώς παρουσιάζεται μείωση της αποτελεσματικότητας του συστήματος από έναν αριθμό και μετά.

Σχετικά με την διαδικασία αναγνώρισης, η σχέση μεταξύ των τριών προαναφερθέντων μεγεθών είναι παρόμοια.

³⁵ Douglas A. Reynolds, Richard C. Rose “*Robust Text-Independent Speaker Identification Using Gaussian Mixture Models*” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1 . January 1995, §.II E “Speaker Identification”

³⁶ Douglas A. Reynolds, Richard C. Rose “*Robust Text-Independent Speaker Identification Using Gaussian Mixture Models*” IEEE Transactions On Speech And Audio Processing, Vol.3, No.1 . January 1995, §.II C 3) “Speaker Identification”



Σχ. 7.14 Απόδοση του συστήματος αναγνώρισης ομιλητή σε συνάρτηση με τον αριθμό των συστατικών για μοντέλα εκπαιδευμένα με 30, 60 και 90 δευτερόλεπτα ομιλίας. Η ομιλία που αναγνωρίζεται έχει διάρκεια 5 δευτερόλεπτα.



Σχ. 7.15 Απόδοση του συστήματος αναγνώρισης ομιλητή σε συνάρτηση με τον αριθμό των πυκνοτήτων των συστατικών για κάθε μοντέλο ομιλητή, για διάρκειες δείγματος 1, 5 και 10 δευτερόλεπτα.

7.3.4. Κρυφά μαρκοβιανά μοντέλα - Hidden Markov Models (HMM)

Τα Στοιχεία Των Hidden Markov Models

Ένα HMM είναι ένας συνδυασμός δύο στοχαστικών διαδικασιών, (Q, X) , μιας Q κρυφής αλυσίδας Markov (hidden Markov chain), η

οποία περιγράφει τη χρονική μεταβλητότητα (temporal variability), και μιας \underline{X} , φανερής, η οποία περιγράφει τη φασματική μεταβλητότητα (spectral variability).

Για να οριστεί πλήρως ένα HMM, απαιτείται καθορισμός των εξής στοιχείων:

- Αριθμός καταστάσεων (number of states): N
- Πλήθος διακριτών συμβόλων (number of distinct observation symbols), που μπορούν να διακριθούν ανά κατάσταση: M , για διακριτά HMMs ή άπειρο, για συνεχή HMMs, δηλαδή όταν x_t είναι πραγματικός αριθμός ή διάνυσμα πραγματικών.
- Πιθανότητες μετάβασης (state transition probability distribution): η διαδικασία $\{q_t\}$ μοντελοποιείται ως αλυσίδα Markov με πιθανότητες μετάβασης $A = \{a_{ij}\}$ όπου $a_{ij} = P(q_{t+1} = j / q_t = i)$, με $1 \leq i \leq N$.

Με άλλα λόγια :

$$a_{ij} = \frac{\text{αναμενόμενο πλήθος μεταβάσεων από την κατάσταση } i \text{ στην κατάσταση } j}{\text{αναμενόμενο πλήθος μεταβάσεων από την κατάσταση } i \text{ σε οποιαδήποτε κατάσταση}}$$

- Κατανομές εξόδου (output distribution) σε μία κατάσταση j . Σε κάθε χρονική στιγμή δημιουργείται μία παρατήρηση (τυχαίο διάνυσμα ή τυχαία μεταβλητή (διακριτή ή συνεχής), με βάση μία κατανομή, που εξαρτάται από την κατάσταση στην οποία βρισκόμαστε. Για διακριτά HMMs είναι $B = \{b_j(x_t)\}$, όπου το μέγεθος $b_j(x_t) = P(x_t | q_t = j)$ είναι η κατανομή εξόδου, με $1 \leq j \leq N$ και $1 \leq x_t \leq M$.

Με άλλα λόγια :

$$b_j(x_t) = \frac{\text{αναμενόμενη συχνότητα παραμονής στην κατάσταση } j \text{ παρατηρώντας τα σύμβολα } x_t}{\text{συνολική συχνότητα παραμονής στην κατάσταση } j}$$

- Αρχικές πιθανότητες : $\Pi = \{\pi_i\}$, όπου $\pi_i = P(q_0 = i)$ με $1 \leq i \leq N$, για την ακολουθία καταστάσεων : $q_0, q_1, q_2, \dots, q_t, \dots$ όπου $q_t \in \{1, 2, \dots, N\}$.

Με άλλα λόγια :

π_i = αναμενόμενη συχνότητα (αριθμός φορών) στην κατάσταση i τη χρονική στιγμή $t=1$

Έτσι, με δεδομένες κατάλληλες τιμές των μεγεθών N , M , A , B και π , το *HMM* μπορεί να χρησιμοποιηθεί σαν μία γεννήτρια που παράγει ακολουθίες εξόδου της μορφής :

$$\underline{X} = [x_1, x_2, \dots, x_T]$$

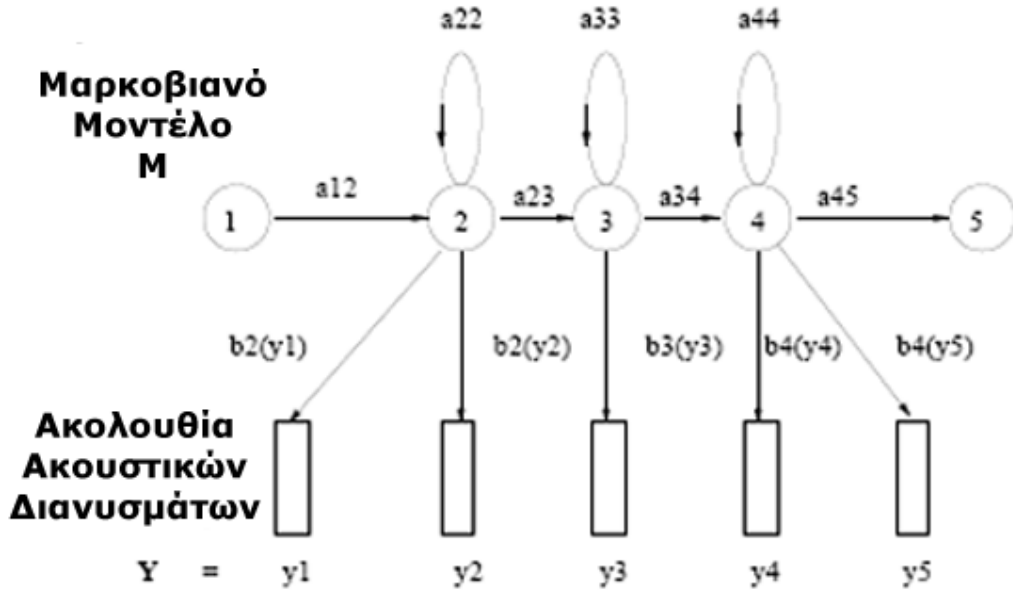
όπου x_t είναι η παρατήρηση ενός συμβόλου και T είναι το πλήθος των παρατηρήσεων στην συγκεκριμένη ακολουθία.

Η παραπάνω διαδικασία μπορεί να χρησιμοποιηθεί, ταυτόχρονα, ως γεννήτρια παρατηρήσεων, αλλά και ως ένα μοντέλο του τρόπου, σύμφωνα με τον οποίο, μία δοσμένη ακολουθία παρατηρήσεων έχει παραχθεί από ένα κατάλληλο HMM. Από αυτή την ανάλυση φαίνεται ότι για τον πλήρη καθορισμό ενός HMM απαιτούνται οι παράμετροι N και M καθώς και ο καθορισμός του συνόλου των συμβόλων παρατήρησης και των τριών πιθανοτικών μεγεθών: A , B , π . Για συντομία χρησιμοποιούμε τον πιο συμπαγή συμβολισμό:

$$\lambda = (A, B, \pi)$$

για την συνολική περιγραφή των παραμέτρων ενός HMM.

Το HMM αναπαριστά το φώνημα που είναι η μονάδα μοντελοποίησης. Ένα HMM έχει μία κατάσταση εισόδου, μία κατάσταση εξόδου και 3 ενδιάμεσες καταστάσεις. Οι καταστάσεις εισόδου-εξόδου χρησιμεύουν στην ένωση πολλών διαδοχικών HMMs, τα οποία σχηματίζουν ένα σύνθετο HMM, το οποίο μπορεί να αναπαραστήσει μεγαλύτερες μονάδες, όπως λέξεις ή μια ολόκληρη πρόταση. Το HMM μπορεί να θεωρηθεί σαν μια γεννήτρια ακολουθίας συμβόλων (στη περίπτωση μας, ακουστικών διανυσμάτων) μοντελοποιώντας παράλληλα μια κρυφή ακολουθία καταστάσεων (αποτελούμενων από τις 3 ενδιάμεσες καταστάσεις). Η πιθανότητα μετάβασης από την κατάσταση i στην κατάσταση j καθορίζεται από την διακριτή πιθανότητα a_{ij} , ενώ η πιθανότητα παραγωγής ενός συμβόλου στη κατάσταση j , καθορίζεται από τη πιθανότητα εξόδου b_j .



Σχ. 7.9 HMM μοντέλο. Ένα HMM μπορεί να θεωρηθεί σαν μια γεννήτρια ακολουθίας συμβόλων (στη περίπτωση μας, ακουστικών διανυσμάτων) μοντελοποιώντας παράλληλα μια κρυφή ακολουθία καταστάσεων. Μεταβαίνει μεταξύ των καταστάσεων βάσει των πιθανοτήτων μετάβασης και κάθε χρονική περίοδο παράγει ένα καινούριο ακουστικό διάνυσμα σύμφωνα με την κατανομή εξόδου της παρούσας κατάστασης..

Στο Σχήμα 7.9 φαίνεται ένα παράδειγμα της παραπάνω διαδικασίας, όπου το μοντέλο μεταβαίνει μέσω της ακολουθίας καταστάσεων $X = 1, 2, 2, 3, 4, 4, 5$, για να παράγει την ακολουθία y_1 έως y_5 . Η συνδυασμένη πιθανότητα μιας ακολουθίας διανυσμάτων Y και μιας ακολουθίας καταστάσεων X , δεδομένου κάποιου μοντέλου M (πιθανότητες μετάβασης και εξόδου), υπολογίζεται σαν το γινόμενο των πιθανοτήτων μετάβασης και των πιθανοτήτων εξόδου. Για την ακολουθία καταστάσεων X του σχήματος 7.9 έχουμε ότι:

$$P(Y, X | M) = a_{12}b_1(y_1)a_{22}b_2(y_2)a_{23}b_3(y_3)... \quad (7.30)$$

Στην γενική περίπτωση που $X = x(1), x(2), x(3), \dots, x(T)$, η παραπάνω πιθανότητα γίνεται :

$$P(Y, X | M) = a_{x(0)x(1)} \prod_{t=1}^T b_{x(t)}(y_t) a_{x(t)x(t+1)} \quad (7.31)$$

όπου $x(0)$ είναι η κατάσταση εισόδου του μοντέλου και $x(T+1)$ η κατάσταση εξόδου.

Στην πραγματικότητα, εμείς γνωρίζουμε μόνο την ακολουθία Y , ενώ η ακολουθία X μας είναι άγνωστη, για αυτό και μιλάμε για κρυφές Μαρκοβιανές ακολουθίες. Για τον υπολογισμό, λοιπόν, της $P(Y|M)$, αρκεί να αθροίσουμε την 1.3 για όλες τις πιθανές μεταβάσεις καταστάσεων. Ένας αποτελεσματικός τρόπος για να γίνει αυτό είναι ο *forward - backward* αλγόριθμος.

Είδη HMMs

Η επιλογή της κατανομής εξόδου, παίζει σημαντικό ρόλο αφού μοντελοποιεί την μεταβλητότητα του φάσματος της φωνής, σε αντίθεση με το μοντέλο μετάβασης καταστάσεων το οποίο έχει να κάνει με την διάρκεια. Ανάλογα λοιπόν με το πως μοντελοποιούνται οι κατανομές εξόδου, έχουμε τρεις κατηγορίες ακουστικών μοντέλων.

Διακριτά HMMs³⁷

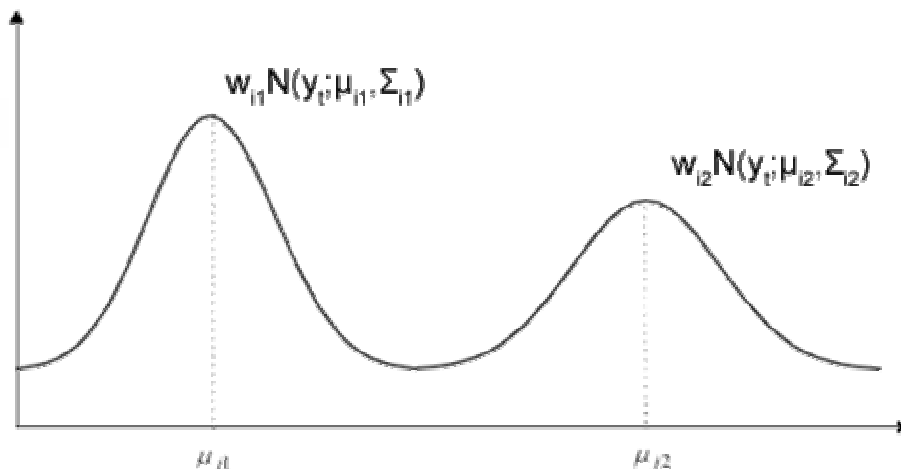
Τα πρώτα συστήματα αναγνώρισης φωνής χρησιμοποιούσαν διακριτές κατανομές εξόδου σε συνδυασμό με διανυσματικό κβαντιστή (*Vector Quantizer - VQ*). Έτσι, κάθε ακουστικό διάνυσμα αντικαθίσταται από το *index* (δείκτη) του πιο κοντινού σε αυτού διανύσματος από ένα σύνολο τέτοιων διανυσμάτων (*codebook*), ενώ οι κατανομές εξόδου ισοδυναμούν με *look-up tablets* που περιέχουν τις πιθανότητες για κάθε πιθανό δείκτη. Για τη κατανομή εξόδου $b_j(y_t)$ στην κατάσταση j , με y_t το ακουστικό διάνυσμα τη χρονική στιγμή t θα ισχύει, δεδομένου της απεικόνισης του y_t στο διάνυσμα y_m του *codebook*:

$$\sum_{m=1}^M b_j(y_m) = 1 \quad (7.32)$$

Το μεγάλο πλεονέκτημα, όσον αφορά τα μοντέλα, είναι ότι, υπολογιστικά, αυτός ο τρόπος είναι αρκετά αποτελεσματικός, από άποψη αναγνώρισης, αφού η πιθανότητα εξόδου δεν χρειάζεται να υπολογιστεί, αλλά ανακτάται μέσω ενός *look-up table*. Αντίθετα, το μεγαλύτερο μέρος του χρόνου σπαταλείται στον κβαντισμό, όπου πρέπει να υπολογιστεί η ελάχιστη απόσταση του διανύσματος εισόδου με τα πρότυπα ακουστικά διανύσματα που περιέχονται στο *codebook*. Βλέπουμε λοιπόν ότι ο μηχανισμός *VQ* δρα σαν προεπεξεργαστής αναγνώρισης (*recognition preprocessor*).

³⁷ Περακάκης Μανόλης, «Κωδικοποίηση φωνής για αναγνώριση σε ενσύρματα και ασύρματα δίκτυα δεδομένων (Διαδίκτυο, GPRS) μέσα από διάλογο 2 Kbps», Πολυτεχνείο Κρήτης, τμήμα ηλεκτρολόγων μηχανικών και μηχανικών υπολογιστών, εργαστήριο τηλεπικοινωνιών πληροφορίας και δικτύων, Διπλωματική εργασία, Δεκέμβριος 2000, §1.5

Όμως εκ των πραγμάτων, η ανάλυση του ακουστικού χώρου είναι μικρή (αφού έχουμε πεπερασμένο αριθμό συμβόλων εξόδου) με επιπτώσεις στην ακρίβεια της αναγνώρισης, οι οποίες ενισχύονται και λόγω του κβαντισμού ο οποίος εισάγει επιπλέον θόρυβο. Επίσης η χρήση VQ με είσοδο «ολόκληρο» το διάνυσμα ως έχει, απαιτεί πολύ μεγάλα codebooks για να καλύψει το ακουστικό χώρο, ακόμα και όταν αυτός παραμένει σχετικά μικρός.



Σχ. 7.10. Γκαουσιανές κατανομές εξόδου σε συνεχές HMM

Βέβαια για τις πρώτες εφαρμογές αναγνώρισης φωνής το μοντέλο αυτό λειτουργούσε ικανοποιητικά, αφού εκείνη την εποχή η διαθεσιμότητα σε πόρους, όπως μνήμη και υπολογιστική ισχύς, ήταν περιορισμένη, ενώ και τα λεξιλόγια ήταν σχετικά μικρά. Καθώς περνούσαν τα χρόνια, με την εισαγωγή όλο και ισχυρότερων υπολογιστικών συστημάτων αλλά κυρίως με τις απαιτήσεις εφαρμογών για πολύ μεγαλύτερο λεξιλόγιο, έγινε φανερό ότι το συγκεκριμένο μοντέλο δεν ήταν πια κατάλληλο.

Συνεχή HMMs – Gaussian mixture HMMs³⁸

Έτσι, τα σύγχρονα συστήματα αναγνώρισης χρησιμοποιούν παραμετρικές, συνεχείς κατανομές εξόδου που μοντελοποιούν τα ακουστικά διανύσματα απευθείας. Για το σκοπό αυτό, συνήθως, χρησιμοποιούνται μείγματα Γκαουσιανών κατανομών και η πιθανότητα εξόδου (*output probability*) είναι η εξής :

³⁸ Περακάκης Μανόλης, «Κωδικοποίηση φωνής για αναγνώριση σε ενσύρματα και ασύρματα δίκτυα δεδομένων (Διαδίκτυο, GPRS) μέσα από διάλογο 2 Kbps», Πολυτεχνείο Κρήτης, τμήμα ηλεκτρολόγων μηχανικών και μηχανικών υπολογιστών, εργαστήριο τηλεπικοινωνιών πληροφορίας και δικτύων, Διπλωματική εργασία, Δεκέμβριος 2000, §1.5

$$b_j(y_t) = \sum_{jm}^M w_{jm} N(y_t; \mu_{jm}, \Sigma_{jm}) \quad (7.33)$$

όπου w_{jm} είναι το βάρος (*weight*) του μείγματος, m , στην κατάσταση j και $N(y_t; \mu_{jm}, \Sigma_{jm})$ είναι η πολυδιάστατη (*multivariate*) Γκαουσιανή, με μέση τιμή μ και πίνακα συνδιακυμάνσεων Σ για το μείγμα m στην κατάσταση j . Συνήθως, «παρόμοιες» καταστάσεις μεταξύ *HMMs* χρησιμοποιούν κοινές Γκαουσιανές, έτσι ώστε να μειώνεται ο συνολικός αριθμός τους. Παρ' όλα αυτά, ο χρόνος που χρειάζεται τόσο για την εκπαίδευση των παραμέτρων τους, όσο και κατά την διαδικασία της αναγνώρισης, είναι αρκετά μεγαλύτερος σε σχέση με τους αντίστοιχους χρόνους με διακριτά *HMM* μοντέλα.

8 ΚΕΦΑΛΑΙΟ

8. ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

8.1. ΕΙΣΑΓΩΓΗ

Το πειραματικό μέρος, της διπλωματικής μας εργασίας, χωρίζεται σε δύο ξεχωριστά στάδια πειραματικών διαδικασιών. Το πρώτο στάδιο σχετίζεται με την «Αναγνώριση Ομιλητή» (Speaker Identification) και το δεύτερο με την «Επιβεβαίωση Ομιλητή» (Speaker Verification).

Στο στάδιο Αναγνώρισης Ομιλητή, ο αντικειμενικός σκοπός μας, είναι η αναγνώριση της ηχητικής εκπομπής (ταυτοποίηση) ενός εντόμου, ανάμεσα από ένα σύνολο ηχητικών εκπομπών – φωνήσεων εντόμων, διαφορετικού είδους, συμπεριλαμβανομένου και της προνύμφης του ρυγχωτού κανθάρου,.

Στο δε στάδιο της Επιβεβαίωσης Ομιλητή, ο στόχος μας είναι η επιβεβαίωση, ή μη, της ταυτότητας του ρυγχωτού κανθάρου, σε σχέση με το υπόλοιπο του συνόλου των εντόμων, που λαμβάνουν μέρος στην πειραματική διαδικασία.

Περιγραφή του λογισμικού (software)

Το λογισμικό που χρησιμοποιήσαμε για την διεξαγωγή των πειραματικών διαδικασιών είναι μια εφαρμογή που έχει υλοποιηθεί στο περιβάλλον του Matlab, της Mathworks Inc.

Η διαδικασία η οποία ακολουθείται, προκειμένου να εξαχθεί η απόφαση σχετικά με την αναγνώριση ή την επιβεβαίωση ομιλητή, στην συγκεκριμένη εφαρμογή, αναλύεται στα παρακάτω στάδια.

- Προέμφαση (βλέπε 7.1.1.)
- Εφαρμογή Παραθύρου (Windowing) (βλέπε 7.1.2)
- Fast Fourier Transformation (FFT)³⁹
- Δημιουργία *Mel-Frequency* κλίμακας (βλέπε 7.2.3)

³⁹ Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992). “*Numerical recipes in C: the art of scientific computing*”. Cambridge University Press.

- Εξαγωγή του Cepstrum (βλέπε 7.2.2)
- Μοντελοποίηση *GMM* (βλέπε 7.3.4)

8.2. ΠΕΙΡΑΜΑΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ

8.2.1. Αναγνώριση εντόμου

Η Αναγνώριση εντόμου συνοψίζεται τα εξής στάδια:

Συγκέντρωση ηχητικών δειγμάτων προς εκπαίδευση

Σε αυτό το στάδιο, δημιουργήσαμε δύο σύνολα ηχητικών δειγμάτων. Το πρώτο περιλαμβάνει 12 ηχητικά δείγματα ακουστικών εκπομπών εντόμων, που δρουν στο ευρύτερο περιβάλλον του ρυγχωτού κανθάρου, όπως γρύλοι, ακρίδες, τζίτζικια κ.λ.π..(Πείραμα 1.1)

Το δεύτερο σύνολο ηχητικών δειγμάτων, που δημιουργήσαμε, περιλαμβάνει 11 αρχεία ήχου, τα οποία αποτελούν ακουστικές εκπομπές εντόμων που βρίσκονται στο στάδιο της προνύμφης καθώς και τερμιτών που παρασιτούν, κυρίως, μέσα σε ξύλα, κορμούς δέντρων, ρίζες φυτών και αποθηκευμένα γεωργικά προϊόντα. Τα συγκεκριμένα είδη επιλέχθηκαν εξ' αιτίας των ομοιοτήτων που παρουσιάζουν οι ηχητικές τους εκπομπές, λόγω των κοινών τρόπων παραγωγής τους. Με αυτόν τον τρόπο θέλαμε να διαπιστώσουμε την αποτελεσματικότητα του αυτόματου συστήματος αναγνώρισης, στην περίπτωση που πρέπει να αναγνωριστεί ένα είδος ανάμεσα σε ένα σύνολο εντόμων που παρουσιάζουν φασματικές ομοιότητες στις ηχητικές τους εκπομπές. (Πείραμα 1.2)

Όλα τα ηχητικά δείγματα που εισήχθησαν και αναλύθηκαν σε αυτό το στάδιο, σχηματίζουν μία βάση δεδομένων που περιλαμβάνει τα ονόματα των δειγμάτων καθώς και έναν κωδικό αριθμό για κάθε ένα από αυτά.

Εκπαίδευση συστήματος και δημιουργία μοντέλων

Σε αυτό το στάδιο κάθε ηχητικό δείγμα εισάγεται στο σύστημα αναγνώρισης προκειμένου να εξαχθούν οι *MFCC* παράμετροι και να δημιουργηθούν τα *GMM* μοντέλα. Με αυτόν τον τρόπο γίνεται η εκπαίδευση του συστήματος.

Εισαγωγή ηχητικού δείγματος του ρυγχωτού κανθάρου για αναγνώριση

Για το στάδιο της αναγνώρισης εισάγουμε ένα ηχητικό δείγμα ακουστικής εκπομπής του ρυγχωτού κανθάρου για να διαπιστώσουμε αν το σύστημα το αναγνωρίζει. Εξάγονται οι *MFCC* παράμετροι και δημιουργείται ένα *GMM* μοντέλο, το οποίο θα τεθεί υπό σύγκριση με τα *GMM* μοντέλα που βρίσκονται στην βάση δεδομένων.

Εξαγωγή αποτελέσματος

Εφόσον το σύστημα πραγματοποιήσει την επεξεργασία του σήματος και προβεί στις σχετικές συγκρίσεις, μας δίνει τον κωδικό αριθμό του είδους με το οποίο αντιστοιχήθηκε το εισαχθέν, προς αναγνώριση, ηχητικό δείγμα.

Σημειώνουμε ότι τα δύο τελευταία στάδια επαναλήφθηκαν για δέκα διαφορετικά δείγματα του ρυγχωτού κανθάρου, τα οποία εισήχθησαν προς αναγνώριση. Με αυτόν τον τρόπο μπορέσαμε να εξάγουμε το ποσοστό επιτυχίας του αυτόματου συστήματος αναγνώρισης.

Τα ηχητικά δείγματα, με τα οποία εκπαιδεύσαμε το σύστημα, έπρεπε να έχουν την ίδια συχνότητα δειγματοληψίας και το ίδιο εύρος κωδικοποίησης (bit rate). Εξ' αιτίας του γεγονότος ότι τα δείγματα που είχαμε στην διάθεσή μας δεν πληρούσαν αυτήν την προϋπόθεση, μετατρέψαμε τις συχνότητες δειγματοληψίας και τα bit rates, όλων των δειγμάτων, προς μία συχνότητα δειγματοληψίας και ένα bit rate αναφοράς, τα οποία συμφωνήθηκε να είναι: 44100 KHz και 16-bit αντίστοιχα. Αυτή διαδικασία μετατροπής διεκπεραιώθηκε με το software Wavelab 5 της Steinberg.

Μία άλλη παράμετρος, η οποία έπρεπε να ρυθμιστεί, ήταν το μέγεθος των δειγμάτων που θα χρησιμοποιούσαμε για την εκπαίδευση του συστήματος. Σύμφωνα με όσα έχουμε αναφέρει μέχρι στιγμής, η αποτελεσματικότητα του συστήματος αναγνώρισης εξαρτάται από τον αριθμό των συστατικών του Γκαουσιανού μίγματος καθώς και από την διάρκεια των δειγμάτων, που χρησιμοποιούνται για εκπαίδευση και αναγνώριση. Ως εκ τούτου, αποφασίστηκε να δοθεί η μέγιστη δυνατή διάρκεια στα δείγματα, η οποία συνέπιπτε με τη διάρκεια του μικρότερου δείγματος που είχαμε στη διάθεσή μας, η οποία ισοδυναμούσε με χρόνο 30 sec. Επίσης, η διατήρηση της ίδιας διάρκειας, για όλα τα δείγματα, κρίθηκε απαραίτητη για την δημιουργία ισοδύναμων πιθανοτικών μοντέλων *GMM*.

Το στάδιο της αναγνώρισης επαναλήφθηκε για διαφορετικές διάρκειες δειγμάτων εκπαίδευσης καθώς και για διαφορετικές διάρκειες δειγμάτων αναγνώρισης, με σκοπό να αποφανθούμε για το πώς ανταποκρίνεται κατά την μεταβολή της διάρκειας των δειγμάτων

(πείραμα 2.1, για το 1^ο σύνολο δειγμάτων και πείραμα 2.2 για το 2^ο σύνολο δειγμάτων).

8.2.2. Ταυτοποίηση εντόμου

Στη φάση της επιβεβαίωσης ομιλητή, το ζητούμενο είναι η επιβεβαίωση της ταυτότητας του ρυγχωτού κανθάρου ανάμεσα από όλα τα υπόλοιπα είδη εντόμων, που παίρνουν μέρος στο πείραμα. Επομένως, χρειάζεται να δημιουργήσουμε δύο μοντέλα πιθανότητας *GMM*. Το πρώτο θα αφορά την ακουστική εκπομπή του ρυγχωτού κανθάρου και το δεύτερο τις ακουστικές εκπομπές των υπολοίπων εντόμων.

Για την δημιουργία του δεύτερου μοντέλου, προέκυψαν δύο εκδοχές. Η πρώτη εκδοχή (Πείραμα 3.1) περιλαμβάνει μια μίξη ακουστικών εκπομπών από έντομα όπως ακρίδες, γρύλοι, τζίτζικια κ.λ.π., ώστε να προκύψει ένα ενιαίο αρχείο ήχου. Η δεύτερη εκδοχή (Πείραμα 3.2) περιλαμβάνει την μίξη των ακουστικών εκπομπών προνυμφών και τερμιτών, το ακουστικό φάσμα των οποίων μοιάζει με εκείνο του ρυγχωτού κανθάρου. Έτσι, τα στάδια της επιβεβαίωσης ομιλητή συμπίπτουν με εκείνα της αναγνώρισης ομιλητή, μόνο που σε αυτήν την φάση, τα ηχητικά δείγματα που χρησιμοποιήθηκαν για την εκπαίδευση του συστήματος είναι η ηχητική εκπομπή του ρυγχωτού κανθάρου καθώς και το ενιαίο αρχείο ήχου που, εμείς, δημιουργήσαμε.

Εξαγωγή αποτελέσματος

Το αποτέλεσμα που θα μας δώσει το αυτόματο σύστημα αναγνώρισης είναι ο κωδικός αριθμός ενός εκ των 2 αρχείων ήχου που εισήγαμε για την εκπαίδευση των μοντέλων *GMM*.

Και αυτή η φάση διεξήχθη επαναλαμβανόμενη 10 φορές, με σκοπό να μπορέσουμε να υπολογίσουμε το ποσοστό επιτυχίας τους αυτόματου συστήματος αναγνώριση, στην επιβεβαίωση της ταυτότητας του ρυγχωτού κανθάρου.

8.3. ΠΑΡΟΥΣΙΑΣΗ ΚΑΙ ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

8.3.1. Πείραμα 1.1 Αναγνώριση του R.P.W. από πλήθος ακμαίων εντόμων

Όπως φαίνεται και από τα αποτελέσματα του πίνακα, στο συγκεκριμένο πείραμα, είχαμε απόλυτη επιτυχία, με ποσοστό, σωστής πρόβλεψης, 100%. Σε αυτήν την περίπτωση, μπορούμε να

πούμε ότι το σύστημα δεν συνάντησε καμία δυσκολία, αφού ήταν ικανό να ανιχνεύσει, σε κάθε επανάληψη, το ακουστικό φάσμα, της εκπομπής του ρυγχωτού κανθάρου, ανάμεσα από τα φάσματα των υπολοίπων εντόμων.

Επαναλήψεις		1	2	3	4	5	6	7	8	9	10
Είδος											
Ακρίδες	Acridid										
	Tsukutsukuboushi										
	Trimerotropis pallidipennis										
Γρύλοι	Teleogryllus										
	Velarifictorus ornatus										
	Yamato-hibari										
Σκαθάρια	Longhorn beetle										
Πεταλούδες	Tosena melanoptera										
	Aburazemi s										
Τζιτζίκια	Cicada										
Αλογάκι της Παναγίας	Katydid										
Ρυγχωτός Κάνθαρος	Rhynchophorus ferrugineus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Πίνακας 8.1: Αποτελέσματα πειράματος 1.1

Επαναλήψεις		1	2	3	4	5	6	7	8	9	10
Είδος											
Coleoptera	Jewel Beetles										
	Hylotrupes Bajulus										
	Otiorhynchus Sulcatus										
	Diaprepes Abbreviatus			✓							
	Sitophilus Oryzae										
Τερμίτες	Drepanotermes										
	Dry Wood Termite Solders										
Γαιοσκώληκας	Lumbricidae										
Lepidoptera	Plodia Interpunctella										
Πεταλούδες	Euzophera Magnolialis Capps										
Ρυγχωτός Κάνθαρος	Rhynchophorus Ferrugineus	✓	✓		✓	✓	✓	✓	✓	✓	✓

Πίνακας 8.2: Αποτελέσματα πειράματος 1.2

8.3.2. Πείραμα 1.2 Αναγνώριση του R.P.W. από πλήθος προνυμφών

Στη συγκεκριμένη περίπτωση, του πειράματος 1.2, το αυτόματο σύστημα αναγνώρισης, δεν κατάφερε να αναγνωρίσει το ηχητικό δείγμα του ρυγχωτού κανθάρου, σε όλες τις επαναλήψεις. Το ποσοστό επιτυχίας ήταν 90%, καθώς το ηχητική εκπομπή του ρυγχωτού

κανθάρου αντιστοιχήθηκε, 2 φορές, σε λάθος έντομο. Το αποτέλεσμα του πειράματος ήταν, ως ένα βαθμό, αναμενόμενο, εξαιτίας του γεγονότος ότι τα φάσματα των ακουστικών εκπομπών, των εντόμων που έλαβαν μέρος στην πειραματική διαδικασία, παρουσίαζαν πολλές ομοιότητες και κοινά φασματικά μοτίβα.

8.3.3. Πειράματα αναγνώρισης του R.P.W. με μεταβλητή διάρκεια δειγμάτων 2.1, 2.2

Σύμφωνα με τα συγκριτικά πειράματα, που έγιναν, και τα αποτελέσματα που παρουσιάζονται στους παραπάνω πίνακες, το ποσοστό σωστής πρόβλεψης εξαρτάται άμεσα από τη διάρκεια των δειγμάτων. Μεγαλύτερη βαρύτητα, δε, έχει η διάρκεια των δειγμάτων που χρησιμοποιούνται για την εκπαίδευση του συστήματος.

Μέγεθος Δειγμάτων προς εκπαίδευση (sec)	Μέγεθος Δειγμάτων προς αναγνώριση (sec)	
	5	10
10	80%	80%
20	80%	90%
30	100%	100%

Πίνακας 8.5 Αποτελέσματα πειράματος 2.1

Μέγεθος Δειγμάτων προς εκπαίδευση (sec)	Μέγεθος Δειγμάτων προς αναγνώριση (sec)	
	5	10
10	60%	70%
20	70%	80%
30	90%	90%

Πίνακας 8.6 Αποτελέσματα πειράματος 2.2

8.3.4. Πείραμα 3.1 Ταυτοποίηση του R.P.W. από πλήθος ακμαίων εντόμων

Κατά την διεξαγωγή του πειράματος 2.1 παρουσιάστηκε, επίσης, ποσοστό σωστής πρόβλεψης 100%. Εδώ, πρέπει να σημειώσουμε ότι, εξαιτίας της μεγάλης διάρκειας των ηχητικών δειγμάτων, που χρησιμοποιήθηκαν κατά την εκπαίδευση του συστήματος, που συνεπάγεται τη δημιουργία περισσότερων, επί μέρους, Γκαουσιανών, για κάθε μοντέλο, η διαδικασία της αναγνώρισης έγινε περισσότερο χρονοβόρα και το υπολογιστικό κόστος αυξήθηκε σημαντικά.

Επαναλήψεις	1	2	3	4	5	6	7	8	9	10
Είδος										
Rhynchophorus Ferrugineus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Μίξη Ακουστικών Εκπομπών (1 ^η Εκδοχή)										

Πίνακας 8.3: Αποτελέσματα πειράματος 3.1

8.3.5. Πείραμα 3.2 Ταυτοποίηση του R.P.W. από πλήθος προνυμφών

Το ποσοστό σωστής πρόβλεψης του πειράματος 3.2 ήταν 90% καθώς το σύστημα δεν στάθηκε ικανό να επιβεβαιώσει την ταυτότητα του ρυγχωτού κανθάρου, σε 1 περίπτωση. Η αιτία της ανεπιτυχούς ανίχνευσης στηρίζεται, πάλι, στην δυσκολία, που αντιμετωπίζει το αυτόματο σύστημα αναγνώρισης, στο να διακρίνει το φάσμα της ακουστικής εκπομπής του ρυγχωτού κανθάρου, από τα φάσματα των υπολοίπων εντόμων, με τα οποία, παρουσιάζει ομοιότητες.

Επαναλήψεις	1	2	3	4	5	6	7	8	9	10
Είδος										
Rhynchophorus Ferrugineus	✓	✓	✓	✓	✓	✓	✓		✓	✓
Μίξη Ακουστικών Εκπομπών (2 ^η Εκδοχή)								✓		

Πίνακας 8.4: Αποτελέσματα πειράματος 3.2

9 ΚΕΦΑΛΑΙΟ

9. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΟΠΤΙΚΕΣ

Σε αυτήν την εργασία, και σύμφωνα με τα αποτελέσματα των πειραματικών διαδικασιών, μελετήσαμε και καταλήξαμε στο συμπέρασμα ότι είναι δυνατό να ανιχνευθεί ο *Rhynchophorus Ferrugineus* με τη χρήση ενός αυτόματου συστήματος, το οποίο θα αναγνωρίζει, τον τελευταίο, μέσα από ένα πλήθος άλλων εντόμων-παρασίτων.

Διαπιστώσαμε ότι τα ήδη υπάρχοντα αυτόματα συστήματα αναγνώρισης ανθρώπινης ομιλίας, μπορούν να χρησιμοποιηθούν σε εφαρμογές ανίχνευσης και αναγνώρισης ακουστικών εκπομπών εντόμων και ειδικότερα προνυμφών, όπως εκείνης του ρυγχωτού κανθάρου. Τα δε ποσοστά επιτυχίας ήταν αρκετά ενθαρρυντικά ώστε να μπορούμε να εμπιστευτούμε τέτοιου είδους συστήματα, για βιοακουστικές εφαρμογές ενώ η αύξηση του ποσοστού σωστής πρόβλεψης μπορεί να επιτευχθεί με την βελτίωση των αλγορίθμων, πάνω στους οποίους βασίζεται ένα σύστημα αυτόματης αναγνώρισης.

Ένας άλλος παράγοντας που είναι άρρηκτα συνδεδεμένος με την βελτίωση της απόδοσης του συστήματος, είναι η διάρκεια των δειγμάτων που χρησιμοποιούνται, τόσο για την εκπαίδευση όσο και για την αναγνώριση. Όπως προκύπτει από τα πειράματα 2.1 και 2.2, το ποσοστό σωστής πρόβλεψης αυξάνεται καθώς μεγαλώνει η διάρκεια των δειγμάτων. Αυτό μπορεί να αιτιολογηθεί από το γεγονός ότι τα πιθανοτικά μοντέλα καλύπτουν περισσότερες φασματικές διαφοροποιήσεις και γίνονται πιο συγκεκριμένα καθώς αυξάνεται η πληροφορία που καλούνται να μοντελοποιήσουν.

Επίσης, παρατηρήσαμε ότι η αποτελεσματικότητα του συστήματος και η βέλτιστη λειτουργία του απαιτεί την χρήση μονάδων επεξεργαστών με μεγάλη υπολογιστική ισχύ.

Η ανίχνευση του Red Palm Weevil, όπως έχει εξεταστεί και στα υποκεφάλαια 4.4 και 5.1, είναι δυνατή όταν το έντομο βρίσκεται στο στάδιο της προνύμφης. Αυτό συνεπάγεται ότι ο φοίνικας έχει προσβληθεί από το έντομο.

Έχοντας υπόψη ότι τα έντομα, με την ταλάντωση μερών του σώματος τους, όπως τα φτερά (μέλισσες, σφήκες, κουνούπια, μύγες, κ.λ.π.) παράγουν ήχο, θεωρούμε ότι μπορεί να γίνει μελέτη γύρω από τον τρόπο που δονούνται τα φτερά του ακμαίου R.P.W., το συχνοτικό περιεχόμενο του ήχου και κατά πόσο διαφέρει από το συχνοτικό περιεχόμενο άλλων εντόμων, ώστε να μπορεί να ανιχνευτεί από συστήματα αναγνώρισης.

Ολοκληρωμένα συστήματα παρακολούθησης πουλιών, που αποτελούνται από πολλούς σταθμούς, διάσπαρτους σε εκατοντάδες στρέμματα, παρακολουθούν και καταγράφουν, επί εικοσιτετραώρου, τις φωνήσεις των πουλιών, τις ταξινομούν και τις αποστέλλουν αυτόματα μέσω δικτύου (internet audio streaming).

Έχοντας, μελλοντικά, την γνώση γύρω από τον ήχο που παράγεται από τα φτερά του R.P.W., θα ήταν δυνατό να στηθεί ένα ολοκληρωμένο σύστημα παρακολούθησης, όπως αυτό των πουλιών που περιγράφεται παραπάνω, τοποθετώντας σταθμούς, οι οποίοι θα ανιχνεύουν το έντομο από τον ήχο που εκπέμπει με τη δόνηση των φτερών του.

Με αφορμή τα παραπάνω, περί ολοκληρωμένων συστημάτων παρακολούθησης και βάση των δικών μας μελετών για την προνύμφη του *Rhynchophorus Ferrugineus*, πιστεύουμε, επίσης, ότι θα ήταν δυνατή η εγκατάσταση σταθμών, οι οποίοι θα είναι εφοδιασμένοι με πλήθος από αισθητήρες, που θα εισάγονται στο δέντρο (βλ.4.4.2) και θα είναι μόνιμα τοποθετημένοι, στα υψηλότερα σημεία των φοινίκων, όπου και συνηθίζει να γεννάει, το έντομο, τα αυγά του. Αυτό, όμως, πέραν του μεγάλου αριθμού από αισθητήρες που χρειάζονται, προϋποθέτει ότι πρέπει να γίνει μελέτη και να κατασκευαστεί κάποιο περίβλημα, το οποίο θα προστατεύει το εισαγόμενο τμήμα του αισθητήρα, έτσι ώστε να μην ενσωματωθεί μέσα στο δέντρο, στην προσπάθεια του φοίνικα να επουλώσει την οπή η οποία ανοίχθηκε. Επίσης, θα πρέπει, το περίβλημα, να κατασκευαστεί με τέτοιο τρόπο ώστε να μην μειώνει την δυνατότητα ανίχνευσης του αισθητήρα.

ΠΑΡΑΡΤΗΜΑ A

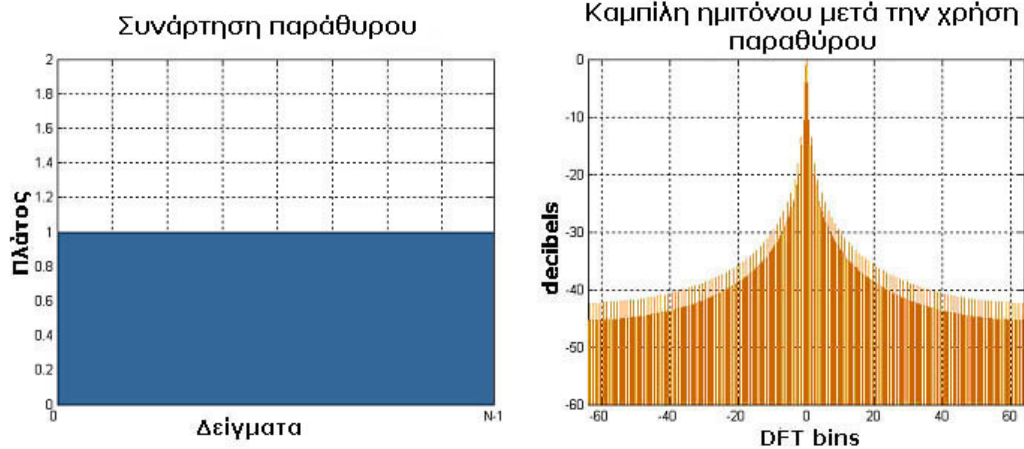
Τύποι Παραθύρων

N = το εύρος, σε δείγματα, μιας συνάρτησης παραθύρου, διακριτού χρόνου. Τυπικά, είναι μία ακέραιη δύναμη του 2 (π.χ. $2^{10} = 1024$)

n = ένας ακέραιος αριθμός που ανήκει στο πεδίο τιμών $0 \leq n \leq N - 1$

B = το ισοδύναμο εύρος ζώνης θορύβου σε DFT bins

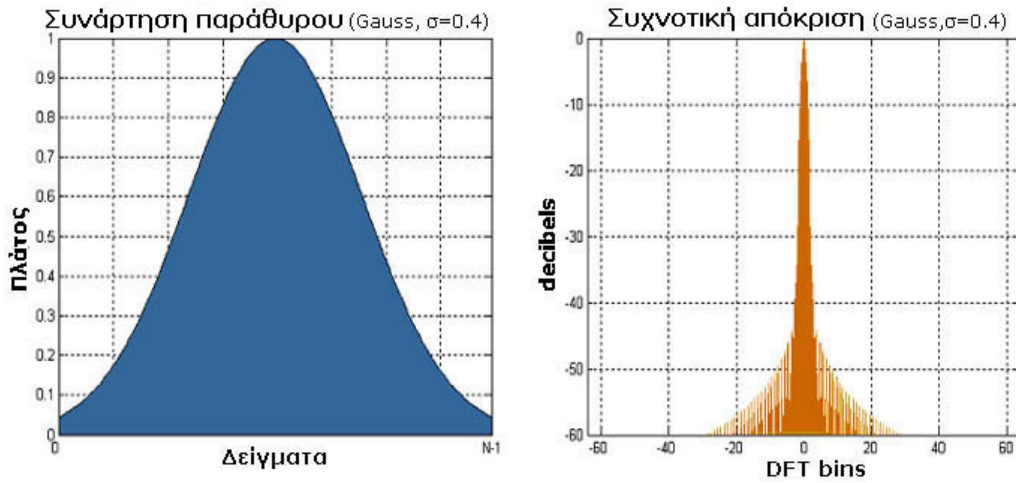
Ορθογώνιο (Rectangular)



Σχ. 1

Εξίσωση: $w(n) = 1$, $B = 1$

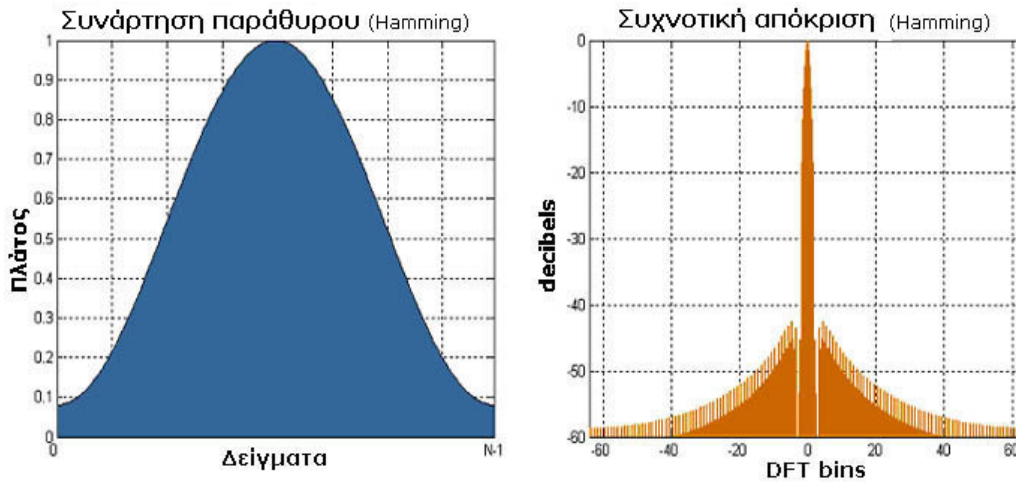
Gauss



Σχ. 2

Εξίσωση: $w(n) = e^{-\frac{1}{2} \left(\frac{n-(N-1)/2}{\sigma(N-1)/2} \right)^2}$, $\sigma \leq 0.5$, $B = 1,45$

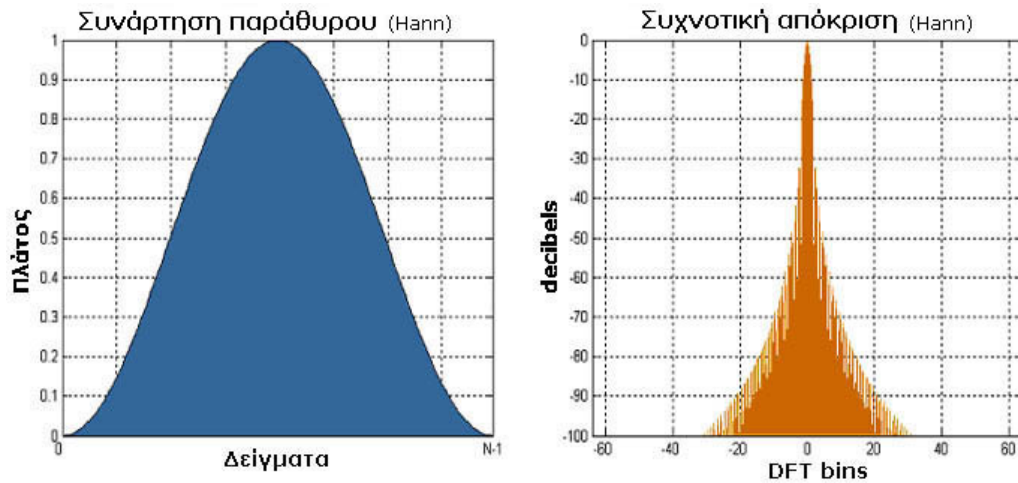
Hamming



Σχ. 3

Εξίσωση: $w(n) = 0,53836 - 0,46164 \cdot \cos\left(\frac{2\pi n}{N-1}\right)$, $B = 1,37$

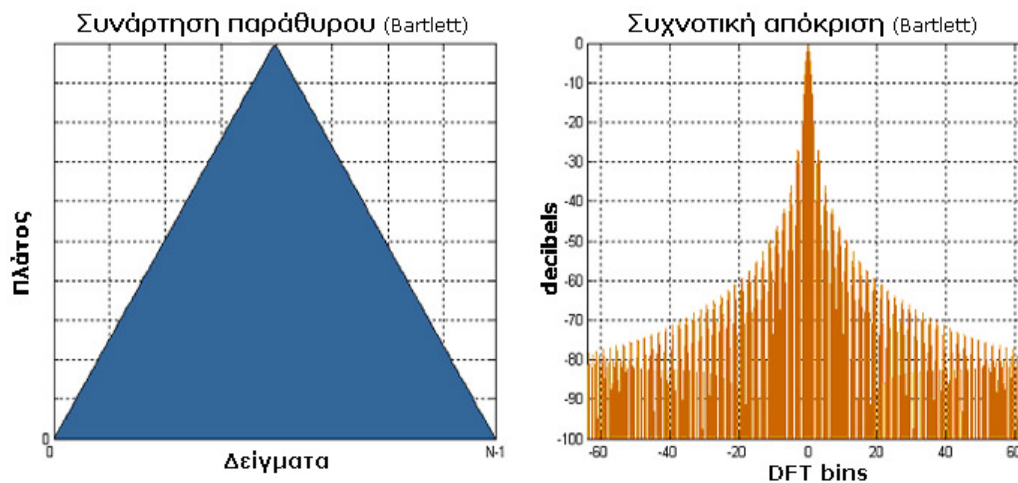
Hanning



Σχ. 4

$$\text{Εξίσωση: } w(n) = 0.5 \cdot \left(1 - \cos\left(\frac{2\pi n}{N-1}\right) \right), \quad B = 1,50$$

Bartlett window (με μηδενισμένα άκρα)

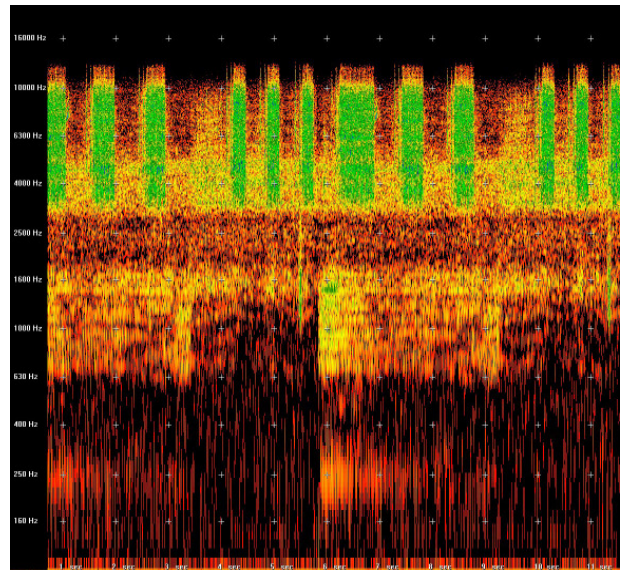


Σχ. 5

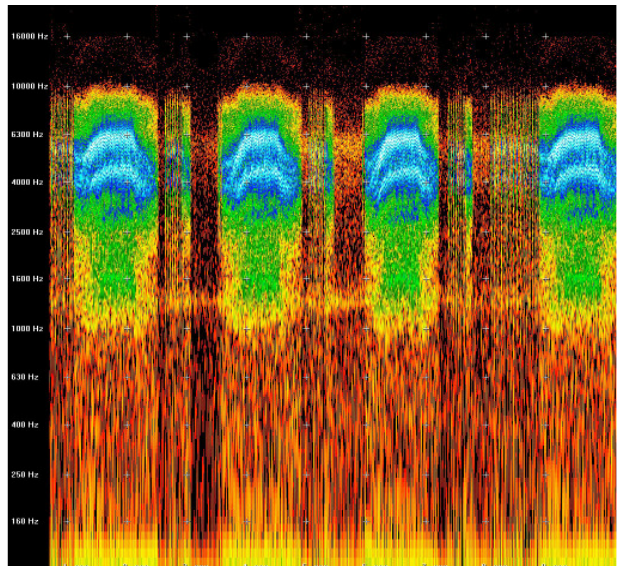
$$\text{Εξίσωση: } w(n) = \frac{2}{N-1} \cdot \left(\frac{N-1}{2} - \left| n - \frac{N-1}{2} \right| \right), \quad B = 1,33$$

ΠΑΡΑΡΤΗΜΑ Β

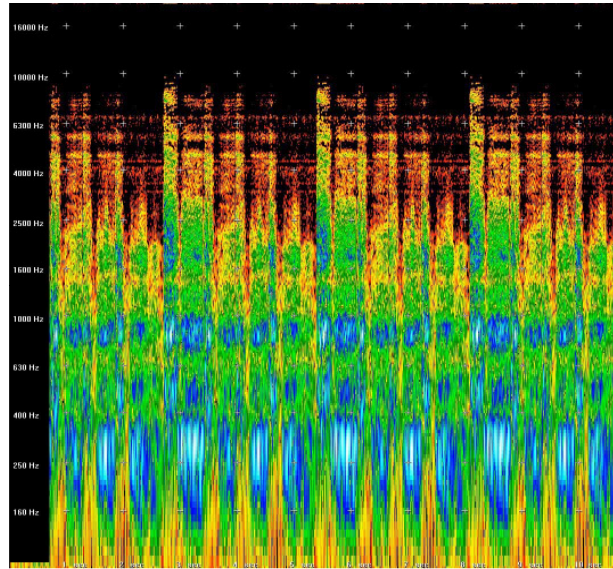
Το σύνολο των εντόμων που χρησιμοποιήθηκαν στα πειράματα 1.1 και 3.1. Αριστερά από το καθένα, το φασματογράφημα δήγματος του ήχου που παράγουν.



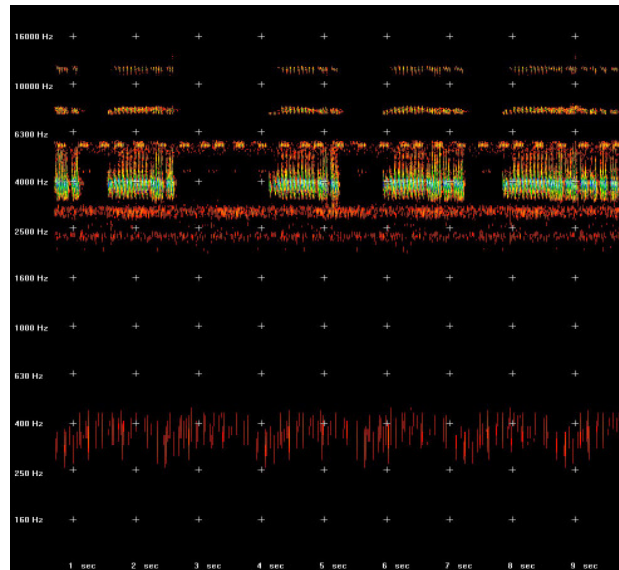
Εικ. 1 Η ακρίδα Acridid της οικογένειας των Acrididae Grasshoppers



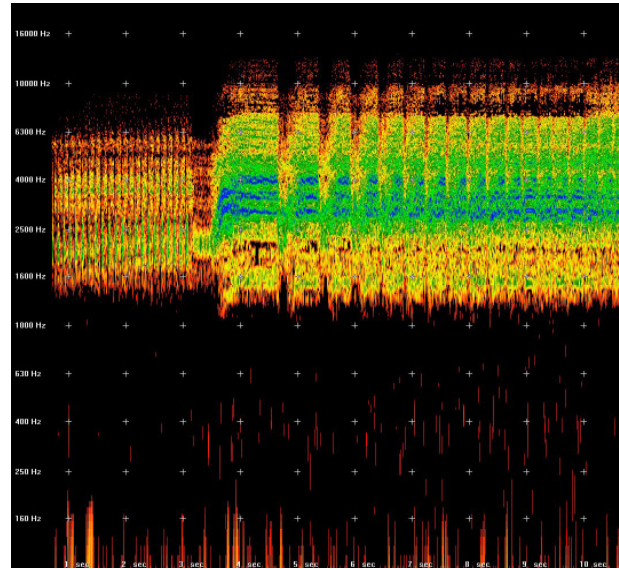
Εικ. 2 Τζιτζίκι (Cicada) το οποίο ανήκει στα Hemiptera και υπάγεται στην οικογένεια των Cicadoidea



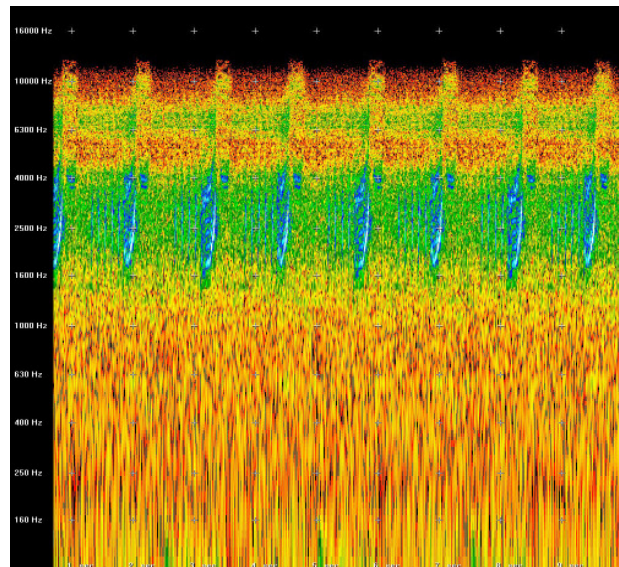
Εικ. 3 Το longhorn beetle ή long-horned beetles της οικογένειας των Cerambycidae



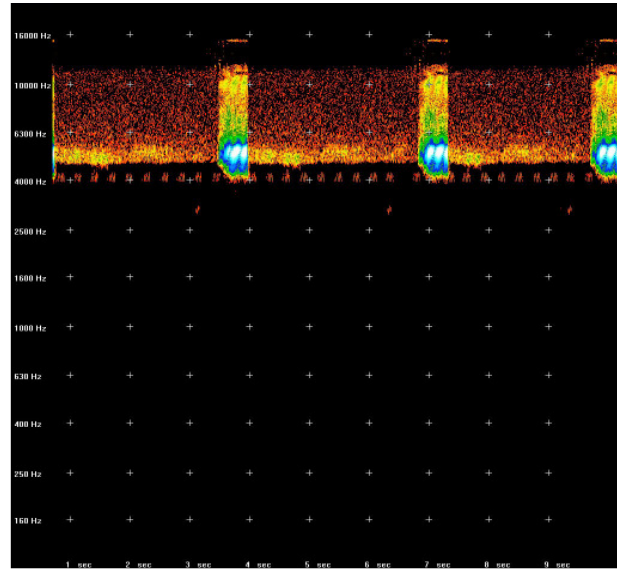
Εικ. 4 Ο γρύλος *Teleogryllus*



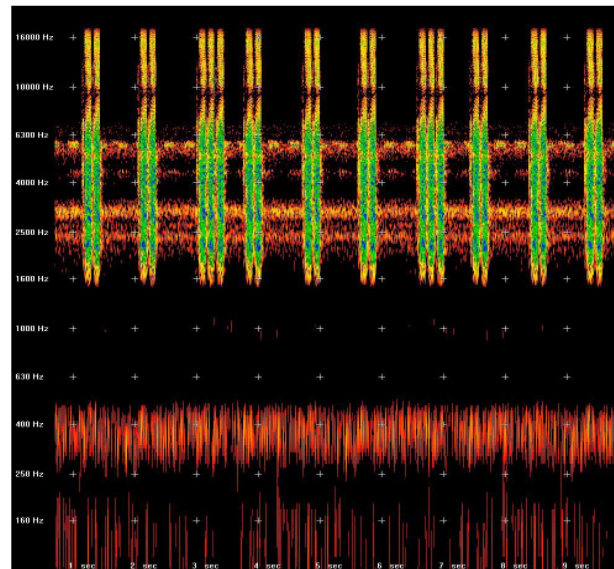
Εικ. 5 *Tosena melanoptera*



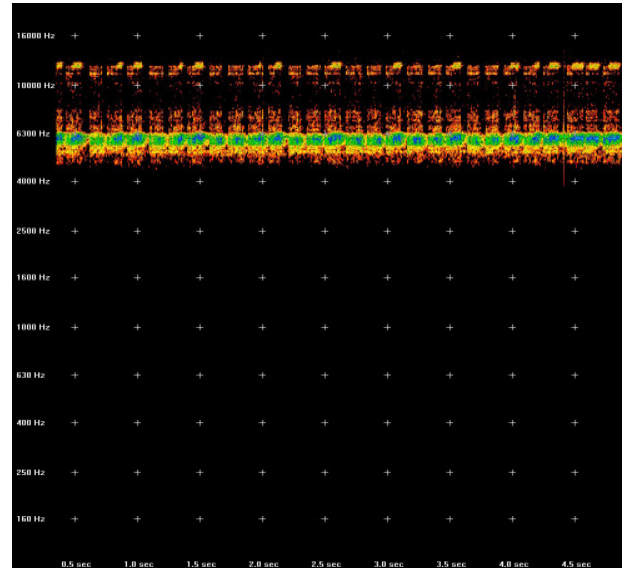
Εικ. 6 *Tsukutsukuboushi* της οικογένειας των Acrididae Grasshoppers



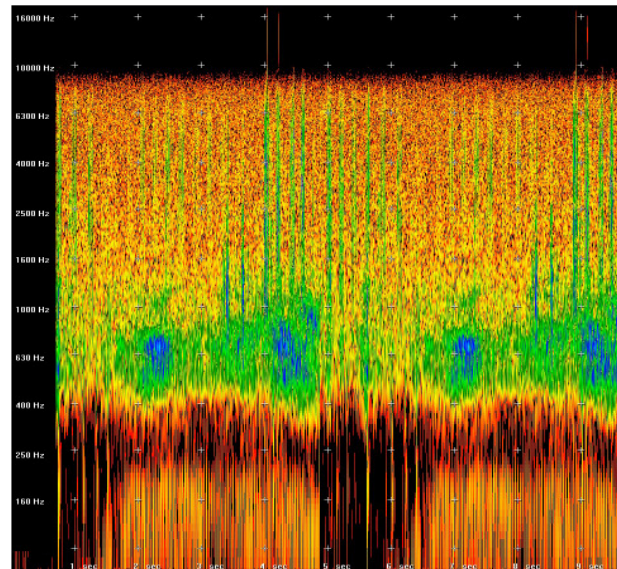
Εικ. 7 Γρύλος, *Velarifictorus ornatus*



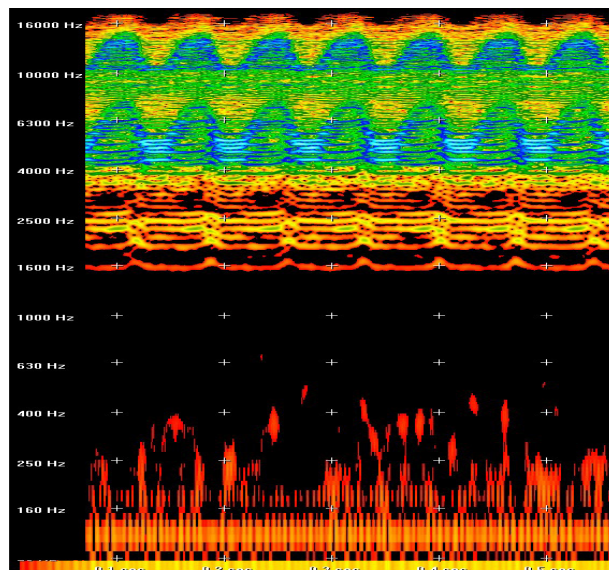
Εικ. 8 Katydid γνωστό και ως «αλογάκι της Παναγίας» της οικογένειας Tettigoniidae



Εικ. 9 Γρύλος, Yamato-hibari

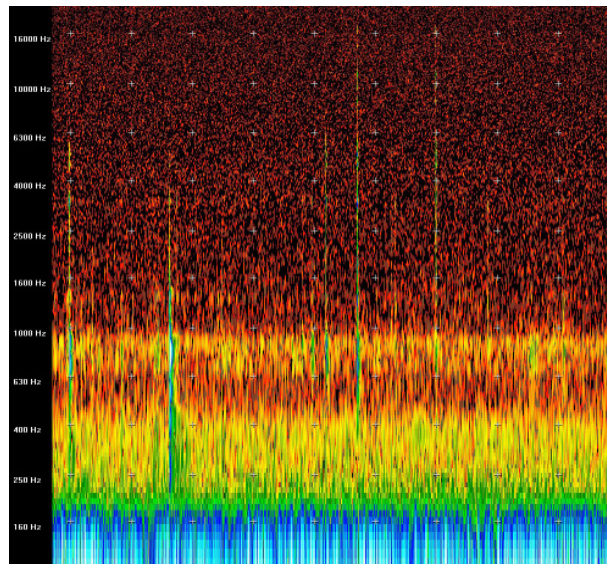


Εικ. 10. Η ακρίδα *Trimerotropis pallidipennis* (pallid-winged) της οικογένειας Acrididae

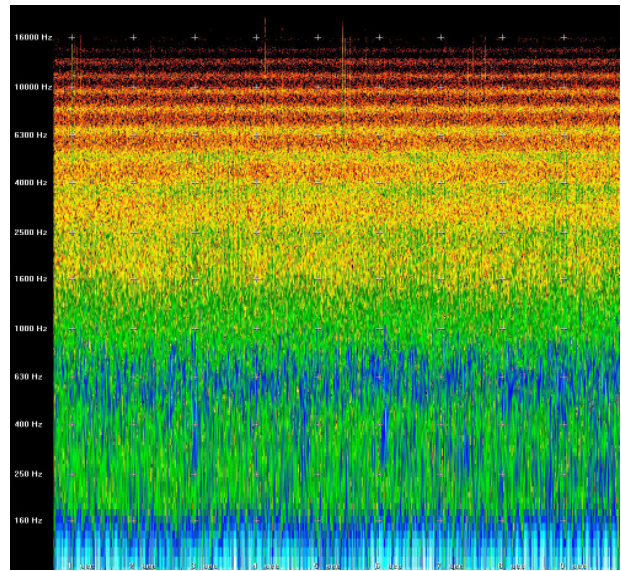


Εικ. 11 Η πεταλούδα *Aburazemi s* οποία συναντάται στις ευρωπαϊκές χώρες και κυρίως στη μεσόγειο.

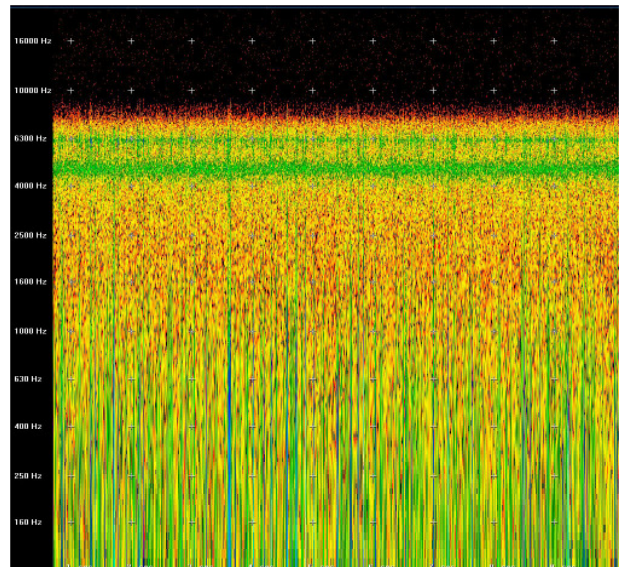
Το σύνολο των προνυμφών και τερμιτών που χρησιμοποιήθηκαν στο πείραμα 1.2 και 3.2. Αριστερά από το καθένα, το φασματογράφημα δείγματος του ήχου που παράγουν.



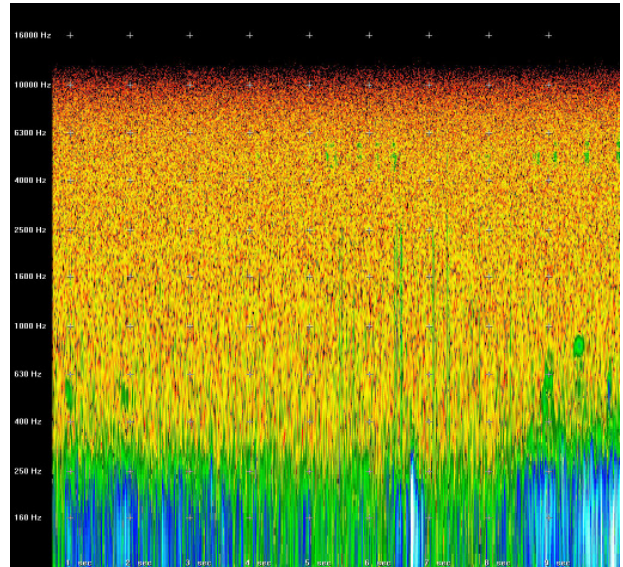
Εικ. 12 Η προνύμφη αυτή προέρχεται από το έντομο jewel beetles ή metallic wood-boring beetles το οποίο ανήκει στα Coleoptera και υπάγεται στην οικογένεια των Buprestidae.



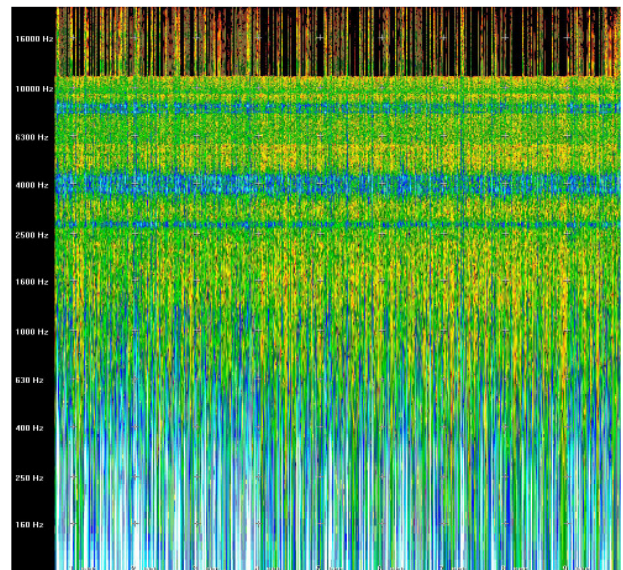
Εικ. 13 Τερμίτες *Drepanotermes* οι οποίοι είναι Pterygota της οικογένειας Isoptera



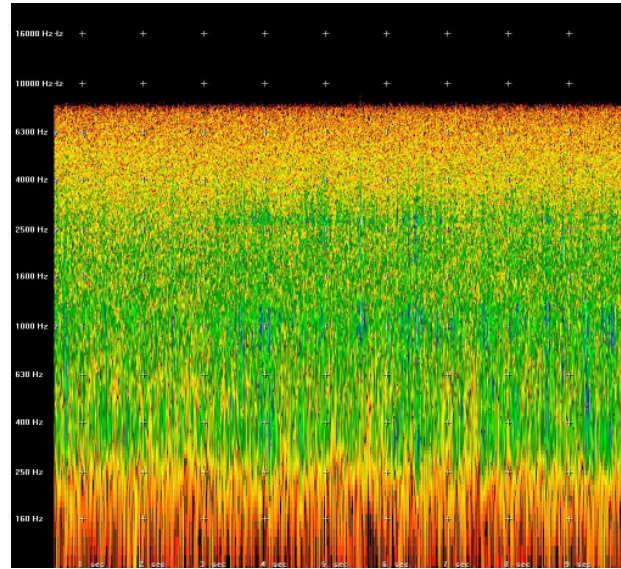
Εικ. 14. Dry wood termite soldiers



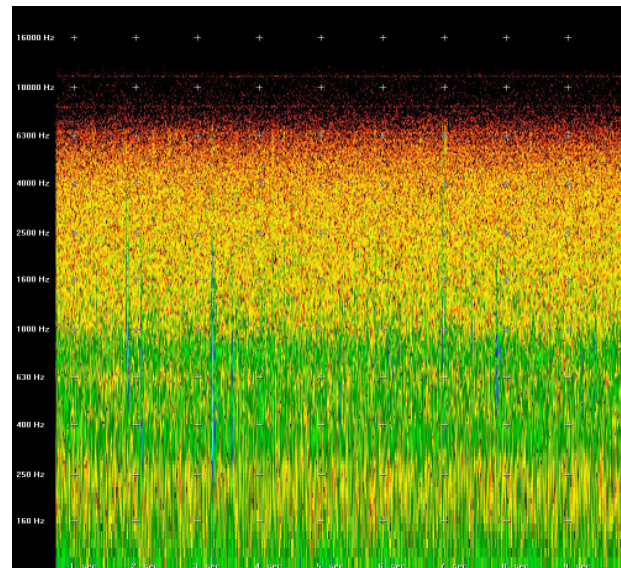
Εικ. 15 Η πεταλούδα *Euzophera magnolialis* Carps στο στάδιο της νύμφωσης. Το φασματογράφημα είναι από την εκπομπή της προνύμφης.



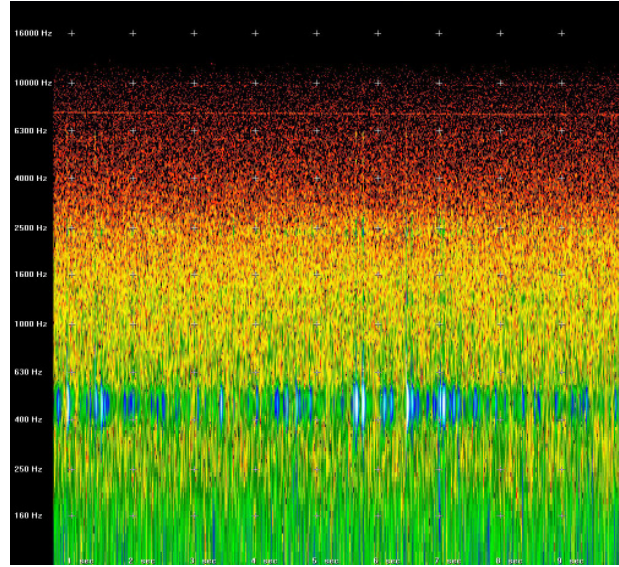
Εικ. 16 Η προνύμφη του *Hylotrupes bajulus* το οποίο ανήκει στα Coleoptera και υπάγεται στην οικογένεια των Cerambycidae.



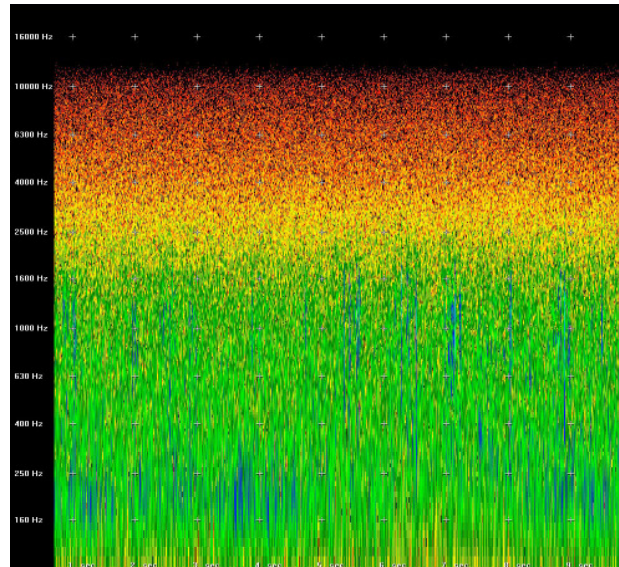
Εικ. 17 Γαιοσκώληκας Lumbricidae το οποίο ανήκει στα Harlotaxida και υπάγεται στην οικογένεια των Lumbricidae.



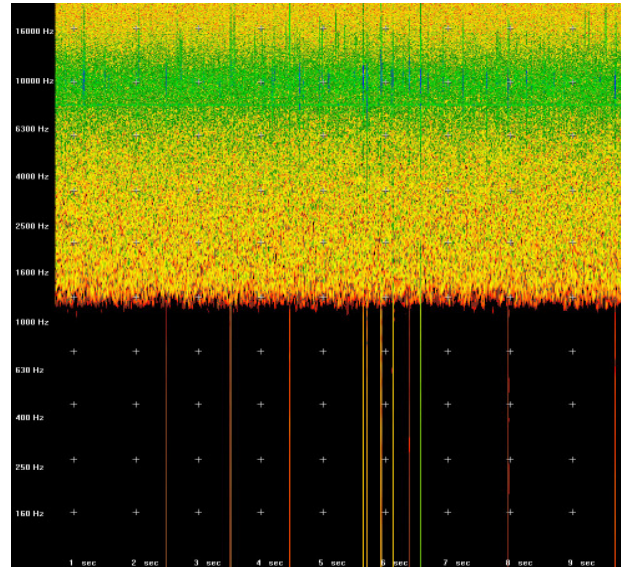
Εικ. 18 Προνύμφη του *Otiorhynchus sulcatus* το οποίο ανήκει στα Coleoptera και υπάγεται στην οικογένεια των Curculionidae



Εικ. 19 Προνύμφη του *Plodia interpunctella* το οποίο ανήκει στα *Lepidoptera* και υπάγεται στην οικογένεια των *Pyalidae*



Εικ. 20 Προνύμφη του *Diaprepes abbreviatus*(citrus root weevil)το οποίο ανήκει στα *Co-*
leoptera και υπάγεται στην οικογένεια των *Curculionidae*



Εικ. 21 Προνύμφη του *Sitophilus oryzae* το οποίο ανήκει στα Coleoptera και υπάγεται στην οικογένεια των Curculionidae



Προτάσεις εξάλειψης θορύβου

Είναι εύκολο να αντιληφθεί κανείς ότι για την εξαγωγή σωστών και επιτυχών αποτελεσμάτων στην αναγνώριση και την επιβεβαίωση ομιλητή, είναι σημαντικό οι ηχογραφήσεις να μην εμπεριέχουν θόρυβο. Το περιεχόμενο της ηχογράφησης δεν πρέπει να επικαλύπτεται από άχρηστη πληροφορία, όπως το θόρυβο από το βάθος πεδίου ή τον μηχανικό θόρυβο.

Η διαφορά στάθμης db ανάμεσα στη μέγιστη στάθμη σήματος και την στάθμη του θορύβου, ονομάζεται λόγος σήματος προς το θόρυβο (signal to noise ratio S/N). Όσο μεγαλύτερος είναι ο λόγος σήματος προς τον θόρυβο τόσο καθαρότερο ήχο έχουμε.

Η αναλογία σήματος προς το θόρυβο πρέπει να είναι τουλάχιστον της τάξης των 40 db. Όσο οι αναλογίες γίνονται μικρότερες της τελευταίας, τόσο μειώνεται το ποσοστό επιτυχούς απόφασης από το σύστημα.

Υπάρχουν πολλές τεχνικές με τις οποίες μπορούμε να κρατήσουμε την αναλογία S/N σε υψηλά επίπεδα. Κάποιες από αυτές είναι:

A) Αρχικά ο πιο απλός τρόπος είναι να χρησιμοποιήσουμε μια πύλη θορύβου (Noise Gate) και να αφαιρέσουμε τον θόρυβο που πιθανότατα θα υπάρχει στον λαμβανόμενο ήχο. Αυτό όμως δεν έχει αποτέλεσμα, όταν η στάθμη του θορύβου κυμαίνεται στα επίπεδα του εκπεμπόμενου σήματος του παρασίτου, γιατί στην προσπάθεια μας να ρυθμίσουμε το Noise Gate να απομακρύνει το θόρυβο, θα απομακρύνει και πληροφορία που θέλουμε.

B) Μια άλλη τεχνική είναι να ηχογραφήσουμε ένα σήμα, με ένα αισθητήρα μέσα στον δέντρο ή χώμα κ.λ.π., χωρίς την παρουσία του εντόμου που θέλουμε να ηχογραφήσουμε. Αυτό θα έχει σαν αποτέλεσμα η ηχογράφηση αυτή να εμπεριέχει μόνο τον μηχανικό θόρυβο άλλα και τον θόρυβο βάθους του υλικού στο οποίο ηχογραφούμε. Έχοντας αυτήν την ηχογράφηση, γνωρίζουμε το συχνοτικό περιεχόμενο του θορύβου (συνήθως είναι χαμηλόσυχνος) και έτσι εύκολα, με την χρήση φίλτρων, μπορούμε να τον απομακρύνουμε. Το ίδιο αποτελέσματα και ίσως καλύτερο μπορούμε να έχουμε με την χρήση equalizer και compressor.

Εισάγοντας το σήμα θορύβου στο EQ και ενισχύοντας τις συχνοτικές περιοχές που περιέχει, το σήμα εισάγεται στην Side Chain είσοδο του compressor ενώ στο IN εισάγεται το σήμα που περιέχει και το

παράσιτο. Αυτό θα έχει σαν αποτέλεσμα στην έξοδο να έχουμε μόνο το σήμα εκπομπής του παρασίτου αφού ο compressor θα έχει συμπίεσει το ενισχυμένο από το EQ σήμα θορύβου.

Γ) Άλλος ένας τρόπος είναι να χρησιμοποιήσουμε ένα σύνολο από Band Pass φίλτρα τα οποία θα επιτρέπουν να περνάνε μόνο οι συχνότητες στις οποίες εκπέμπει το εκάστοτε έντομο-παράσιτο που ηχογραφούμε. Σε περιπτώσεις, όμως, που ενδέχεται να εκπέμπονται ήχοι, οι οποίοι περιέχουν, σχεδόν, όλο το εύρος των συχνοτήτων (σύνηθες φαινόμενο των φωνήσεων του R.P.W.), ίσως είναι καταστροφικό, για την ταυτοποίηση του εντόμου, να μην εμπεριέχονται στο δείγμα που ηχογραφείται. Για τον λόγο αυτό, ίσως, να μην ενδείκνυται η χρήση των Band Pass φίλτρων σε τέτοιες περιπτώσεις αλλά, απλά, να γίνεται μια ενίσχυση στους υπόλοιπους ήχους που ταυτοποιούν το εκάστοτε είδος (όπως για παράδειγμα στα σχήματα 5.5 και 5.6).

ΕΙΚΟΝΟΓΡΑΦΙΑ

ΚΕΦΑΛΑΙΟ 3

- Εικ 3.1** : http://www.ivia.es/~aurbaneja/picudo_web/REVERSO.jpg
- Εικ 3.2** : <http://www.redpalmweevil.com/wpe17.gif>
- Εικ 3.3** : <http://www.fairchildgarden.org/palmguide/images.php?family= ARECA-CEAE&genus=Phoenix>
- Εικ 3.4** : <http://www.fao.org/DOCREP/006/Y4360E/y4360e3i.jpg>
- Εικ 3.5** : <http://portal.paaf.gov.kw/paaf/ershad/palm%20images/A-2/B-2/C-11/Rferrugineus4.jpg>
- Εικ 3.6** : <http://www.redpalmweevil.com/Closeup/Eggs/Picture%201201-1.jpg>
- Εικ 3.7** : : Κοντοδήμας, Δ. - Οικονόμου , Δ. - Θυμάκης, Ν. - Χ.Μεντή και Ανάγνου – Βερονίκη, Μ. "Νέος σοβαρός εχθρός των φοινικοειδών, το κολεόπτερο *Rhynchophorus Ferrugineus* (Olivier) (*Curculionidae: Dryophorinae*)")"
- Εικ 3.8** : Κοντοδήμας, Δ. - Οικονόμου , Δ. - Θυμάκης, Ν. - Χ.Μεντή και Ανάγνου – Βερονίκη, Μ. "Νέος σοβαρός εχθρός των φοινικοειδών, το κολεόπτερο *Rhynchophorus Ferrugineus* (Olivier) (*Curculionidae: Dryophorinae*)")"
- Εικ 3.9** : http://www.palms.org/palmsjournal/2002/ferry_3.jpg
- Εικ 3.10** : <http://www.redpalmweevil.com/Closeup/antmale/Picture%20879-1.jpg>
- Εικ 3.11** : <http://www.redpalmweevil.com/Closeup/Head/Picture%201000-1.jpg>
- Εικ 3.12** : <http://www.redpalmweevil.com/Closeup/antmale/Picture%20977-1.jpg>
- Εικ 3.13** : <http://www.redpalmweevil.com/Closeup/antmale/Picture%20971-1.jpg>
- Εικ 3.14** : <http://www.redpalmweevil.com/Closeup/antennae/Picture%20992-1.jpg>
- Εικ 3.15** : <http://www.redpalmweevil.com/Closeup/antennae/Picture%20885-1.jpg>
- Εικ 3.16** : http://www.plantapalm.com/vpe/photos/Species/Pics/sabal_causiarum_juv.JPG
- Εικ 3.17** : http://www.hirts.com/commercesuite_products/1129660506.jpg
- Εικ 3.18** : <http://www.azwebsitedesign.us.vu/palms/whitetank/images/129.jpg>
- Εικ 3.19** : http://users.bigpond.net.au/filejest/00128_Livistonia_australis.jpg
- Εικ 3.20** : 1),3: Κοντοδήμας, Δ. - Οικονόμου , Δ. - Θυμάκης, Ν. - Χ.Μεντή και Ανάγνου – Βερονίκη, Μ. "Νέος σοβαρός εχθρός των φοινικοειδών, το κολεόπτερο *Rhynchophorus Ferrugineus* (Olivier) (*Curculionidae: Dryophorinae*)")"
- 2)http://www.palms.org/palmsjournal/2002/ferry_4.jpg
- Εικ 3.21** : <http://www.fao.org/DOCREP/006/Y4360E/y4360e3h.jpg>
- Εικ 3.22** : <http://portal.paaf.gov.kw/paaf/ershad/palm%20images/A-2/B-2/C-11/Rferrugineus5.jpg>
- Εικ 3.23** : http://www.seea.es/divulgac/galima/Plagas/Rhynchophorus_ferrugineus/Rferrugineus1_th.jpg
- Εικ 3.24** : http://www.palms.org/palmsjournal/2002/ferry_2.jpg
- Εικ 3.25** : http://www.ivia.es/~aurbaneja/picudo_web/REVERSO.jpg
- Εικ 3.26** : http://www.na.fs.fed.us/fhp/alb/slides_images/slides/larva.htm

ΚΕΦΑΛΑΙΟ 4

- Εικ 4.1** : <http://www.aeconsulting.com/main.html>
- Εικ 4.2** : <http://www.aeconsulting.com/main.html>
- Εικ 4.3** : <http://www.aeconsulting.com/main.html>

- Εικ 4.4** : <http://www.aeconsulting.com/main.html>
Εικ 4.5 : <http://www.avisoft-saslab.com/usg/microphones.htm>
Εικ 4.6 : http://www.posthorn.com/Images/Schoeps/S_CMIT_2.gif
Εικ 4.7 : <http://www.amazing1.com/Graphics/PM60.gif>
Εικ 4.8 :
<http://www.ira.cnr.it/Medicinaweb/Nc/Immagini%20sito/DESCRIPTION%20PARABOLIC%20PROFILE%20FIGURE.JPG>
Εικ 4.9 : <http://www.aeconsulting.com/main.html>
Εικ 4.10 : <http://www.vericomcomputers.com/Support/HowItWorks.htm>
Εικ 4.11 : <http://www.aeconsulting.com/main.html>
Εικ 4.12 : <http://www.aeconsulting.com/main.html>
Εικ 4.13 : <http://www.aeconsulting.com/main.html>
Εικ 4.14 : <http://www.aeconsulting.com/main.html>
Εικ 4.15 : <http://www.aeconsulting.com/main.html>
Εικ 4.16 : <http://www.aeconsulting.com/main.html>
Εικ 4.17 : <http://www.aeconsulting.com/main.html>
Εικ 4.18 : <http://www.aeconsulting.com/main.html>
Εικ 4.19 : http://www.laartech.biz/laartechagriculture/html/tech_notes.html
Εικ 4.20 : <http://www.laartech.biz/data/pdf/WDsystemint.PDF>
Εικ 4.21 : <http://www.laartech.biz/data/pdf/WDsystemint.PDF>
Εικ 4.22 : <http://www.laartech.biz/data/pdf/WDsystemint.PDF>
Εικ 4.23 : <http://www.avisoft.de/usg/usg416.htm>
Εικ 4.24 : <http://www.avisoft.de/usg/usg416.htm>
Εικ 4.25 : <http://www.avisoft.de/usg/usg416.htm>
Εικ 4.26 : <http://www.avisoft.com/usg/usg.pdf>
Εικ 4.27 : <http://www.avisoft.com/usg/usg.pdf>

ΚΕΦΑΛΑΙΟ 5

- Εικ 5.1** : <http://www.laartech.biz/data/pdf/WD60Flyer.pdf>

ΚΕΦΑΛΑΙΟ 7

- Σχ. 7.1** : <http://www.altera.com/products/ip/altera/images/m-alt-fir-1.gif>
Σχ. 7.2 : Dr. Simon "*Speaker Verification Tutorial for Sony*" Lucey Advanced Multimedia Processing (AMP) Labs Carnegie Mellon University 9/5/2002-9/6/2002 (San Jose) (Presentation) p.32
Σχ. 7.3 : Dr. Simon "*Speaker Verification Tutorial for Sony*" Lucey Advanced Multimedia Processing (AMP) Labs Carnegie Mellon University 9/5/2002-9/6/2002 (San Jose) (Presentation) p.59
Σχ. 7.4 : Janet Marques and Pedro J. Moreno "*A Study of Musical Instrument Classification Using Gaussian Mixture Models and Support Vector Machines*" June 1999, Figure 1 "Linear prediction model for speech and music production".
Σχ. 7.6 : <http://labrosa.ee.columbia.edu/doc/HTKBook21/img165.gif>
Σχ. 7.7 : http://lcavwww.epfl.ch/~minhdo/asr_project/image20.gif
Σχ. 7.8 : <http://www.data-compression.com/lbgvq.gif.080>
Σχ. 7.9 : Διγαλάκης Βασίλης. Σημειώσεις μαθήματος: «*Εισαγωγή στην επεξεργασία Φωνής*» Πανεπιστήμιο Κρήτης
Σχ. 7.10 : Διγαλάκης Βασίλης. Σημειώσεις μαθήματος: «*Εισαγωγή στην επεξεργασία Φωνής*» Πανεπιστήμιο Κρήτης

Σχ. 7.11 : Douglas A. Reynolds, Richard C. Rose "Robust Text-Independent Speaker Identification Using Gaussian Mixture Models" *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROSSESING*, VOL.3, No.1 . JANUARY 1995, fig.2

Σχ. 7.12 : Douglas A. Reynolds, Richard C. Rose "Robust Text-Independent Speaker Identification Using Gaussian Mixture Models" *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROSSESING*, VOL.3, No.1 . JANUARY 1995, fig.3

Σχ. 7.13 : CWJ "Speaker Identification using Gaussian Mixture Model" 2000/05/03 (Presentation)

Σχ. 7.14 : Douglas A. Reynolds, Richard C. Rose "Robust Text-Independent Speaker Identification Using Gaussian Mixture Models" *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROSSESING*, VOL.3, No.1 . JANUARY 1995, fig.4

Σχ. 7.15 : Douglas A. Reynolds, Richard C. Rose "Robust Text-Independent Speaker Identification Using Gaussian Mixture Models" *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROSSESING*, VOL.3, No.1 . JANUARY 1995, fig.5

ΠΑΡΑΡΤΗΜΑ Α

Σχ. 1:

http://upload.wikimedia.org/wikipedia/commons/b/bd/Window_function_%28rectangular%29.png

Σχ. 2 :

http://upload.wikimedia.org/wikipedia/commons/f/f2/Window_function_%28gauss%29.png

Σχ. 3 :

http://upload.wikimedia.org/wikipedia/commons/7/7c/Window_function_%28hamming%29.png

Σχ. 4 :

http://upload.wikimedia.org/wikipedia/commons/f/ff/Window_function_%28hann%29.png

Σχ. 5 :

http://upload.wikimedia.org/wikipedia/commons/5/58/Window_function_%28bartlett%29.png

ΠΑΡΑΡΤΗΜΑ Β

Εικ. 1 : <http://www.birdingbraziltours.com/img/acridid-grasshopper.jpg>

Εικ. 2 : http://www.oznet.ksu.edu/dp_hfrr/HortImage/17-year%20Periodical_Cicada.jpg

Εικ. 3 : <http://www.brown-spruce-longhorn-beetle.com/images/brown-spruce-longhorn-beetle-male.jpg>

Εικ. 4 : <http://k41.pbbase.com/v3/77/520077/1/44284276.DSCN7014.JPG>

Εικ. 5 : <http://www.insectcompany.com/gallery/images/cic-melanoptera.gif>

Εικ. 6 : http://www.mokichi.net/insect/image/tsukutsukuboushi0831_sanagi.jpg

Εικ. 7 : <http://mushinone.cool.ne.jp/English/Eenma.htm>

Εικ. 8 : <http://www.hiltonpond.org/images/KatydidBushFProfile01.jpg>

Εικ. 9 : <http://mushinone.cool.ne.jp/contribution/koorogi.htm>

Εικ. 10 : <http://www.insectimages.org/images/3072x2048/1368050.jpg>

Εικ. 11 : http://www.geocities.jp/taihaku6258/photo/samuineru/kontyuu-yatyou/sonota/aburazemi_S.jpg

Εικ. 12. : http://www.entomology.cornell.edu/Extension/Woodys/ALB_Mahaffey/Images/ Agrilus-borerlarvaethumb.jpg

Εικ. 13. : www.utoronto.ca/forest/termite/dreppage.htm

Εικ. 14. : <http://www.allaroundthehouse.com/drywood-termite-solders.jpg>

Εικ. 15. : <http://www.origins.tv/entomology/moths/300/5997/5997PurplishEuzophera090904b.jpg>

Εικ. 16 : <http://www.ipmimages.org/images192x1281435189.jpg>

Εικ. 17. : <http://www.aquaria.com.ua/corm.images/Lumbricidae.jpg>

Εικ. 18. : <http://www.koleopterologie.de/gallery/FHL11/otiorhynchus-sulcatus-larve-foto-koehler.jpg>

Εικ. 19. :<http://cache.eb.com/eb/image?id=3664>

Εικ. 20. :http://www.ars-grin.gov/ars/SoAtlantic/fp/sti/lapointe/drw_iridovirus/larvae.jpg

Εικ. 21. :<http://www.ipm.iastate.edu/ipm/icm/files/images/alfalfa-weevil-larva.jpg>

BIBΛΙΟΓΡΑΦΙΑ

- Aaron E. Rosenberg "**Automatic Speaker Verification: A Review**", Proc. IEEE, vol. 64, p.p. 475 – 487, April 1976
- Andrew W. Moore "**Clustering with Gaussian Mixtures**", School of Computer Science, Carnegie Mellon University, course slideshow Copyright © 2001, 2004 <http://www.autonlab.org/tutorials>
- Benedikt von Laar, Inc. "**The bioacoustic detection of the Red Palm Weevil**" April 2002, technical report.
<http://www.laartech.biz/data/pdf/Red%20Palm%20Weevil.PDF>
- Benedikt von Laar, Gut Klein Goernow, "**Report of the studies on a transfered sample trunk infested with Red Palm Weevil (Rhyncophorus ferrugineus)**" **part 1** 2004, technical report.
<http://www.laartech.biz/data/pdf/Red%20Palm%20Weevil.PDF>
- Bishnu S. Atal "**Automatic Recognition of Speakers from Their Voices**", Proceedings of the IEEE, 64:460 - 475, 1976.
- Carlo Tomasi, "**Estimating Gaussian Mixture Densities with EM – A Tutorial**", Duke University, course notes
<http://www.cs.duke.edu/courses/spring05/cps296.1/handouts/EM/tomasiEM.pdf>
- Carlos Guestrin "**Unsupervised learning or Clustering (cont.) – K-means Gaussian mixture models**", course slideshow, April 5th, 2006
<http://www.cs.cmu.edu/~guestrin/Class/10701/Slides/clustering.pdf>
- Chai Wutiw WATCHAI, Sutat Sae-tang, and Chularat Tanprasert "**Thai Text-Dependent Speaker Identification by ANN with Two Different Time Normalization Techniques**", Proceedings of 5th Natural Language Processing Pacific Rim Symposium (NLPRS 99) & 1st Workshop on Natural Language Processing and Neural Networks (NLPNN 99)
- Douglas A. Reynolds, Richard C. Rose "**Robust Text-Independent Speaker Identification Using Gaussian Mixture Models**" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROSSESING, VOL.3, No.1 . JANUARY 1995
- Douglas A. Reynolds, Thomas F. Quatieri, Robert B. Dunn "**Speaker Verification Using Adapted Gaussian Mixture Models**", Digital Signal Processing, 10:19 - 41, 2000.
- Dr. Joseph Picone "**Fundamentals Of Speech Recognition: A Short Course**", Institute for Signal and Information Processing, Department of Electrical and Computer Engineering, Mississippi State University
http://www.clsp.jhu.edu/ws98/presentations/preliminary/picone/signal_modeling.pdf
- Eric C. Rouchka "**Pattern Matching Techniques and Their Applications to Computational Molecular Biology - A Review**" March 4, 1999, Technical Report, Washington University Department of Computer Science, WUCS-99-09.
<http://kbrin.a-bldg.louisville.edu/brg/papers/WUCS-99-09.pdf>
- Fang Qian, Mingjing Li, Lei Zhang, Hong-Jiang Zhang, Bo Zhang "**Gaussian Mixture Model For Relevance Feedback In Image Retrieval**", 2002 IEEE International Conference on Multimedia and Expo, ICME'02, 2002, I:229-232

- Frank K. Soong **"On the Use of Instantaneous and Transitional Spectral Information in Speaker Recognition"** June 1998, *EEE Trans. Acoustics, Speech Signal Process.* 36 (6), 871-879
- Gerik Alexander von Graevenitz **"About Speaker Recognition Technology"** Bergdata Biometrics GmbH, Germany
- Gray R.M. **"Vector quantization"**, *IEEE ASSP Mag.*, vol.1, April 1984.
- Harsh Gupta, Ville Hautamäki, Tomi Kinnunen, Pasi Fränti **"Field Evaluation of Text-Dependent Speaker Recognition in an Access Control Application"**, *Int. Conf. on Speech and Computer (SPECOM'05)*, Patras, Greece, 551-554, October 2005.
- Janet Marques and Pedro J. Moreno **"A Study of Musical Instrument Classification Using Gaussian Mixture Models and Support Vector Machines"** Compaq Corporation, Cambridge Research laboratory, Technical Report Series CRL 99/4, June 1999
- Jayant M. Naik **"Speaker Verification: A Tutorial"** *IEEE Communications Magazine* January 1990
- Jean-Michel Marin, Kerrie Mengersen, Christian P. Robert **"Bayesian Modelling and Inference on Mixtures of Distributions"** Eds. *Handbook of Statistics Volume 25*. Elsevier Sciences. Christian (2005)
- Joseph A. Kogan, Danie Margoliash **"Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study"**, *Journal of the Acoustical Society of America* 103 (4), 2185-2196, January 1998
- Joseph P. Campbel **"Speaker Recognition"**, Ph.D. Dissertation, Oklahoma State University, 1992.
- Joseph P. Campbell **"Speaker Recognition: A Tutorial"**, *Proc. of IEEE* , Vol. 85, No. 9, pp. 1437-1462, 1997
- Khaled Assaleh **"A Speaker Recognition Tutorial"** The English international symposium on signal processing and its applications, Australia, August 2005
- Kristian Kroschel, Dirk Bechler **"Demonstrator for Automatic Text-independent Speaker Identification"**, *Deutsche Jahrestagung für Akustik (DAGA)*, Braunschweig, März 2006
- Niko Brümmner, Johan du Preez **"Application-Independent Evaluation of Speaker Detection"**, *Odyssey 2004*, The Speaker and Language Recognition Workshop, Toledo, Spain, May 2004
- Oliver Deroo **"A short introduction to speech recognition"**, <http://www.babeltech.com/download/SpeechRecoIntro.pdf>
- Petros A. Maragos **"Two-Dimensional Linear Prediction and Its Application to Adaptive Predictive Coding of Images"**, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.32, no.6, pp.1213-1229, December 1984
- Richard W. Mankin, Robert L. Crooked, Kathy L. Flanders and Jeffrey P. Shapiro **"Acoustic Detection and Identification of Insects in Soil"**, *Proceedings of the 16th International Congress of Acoustics and the 135th Annual Meeting of the Acoustical Society of America*, pp. 685-686, 1998.
- Richard W. Mankin, Stephen L. Lapointe **"Listening To The Larvae: Acoustic Detection Of Diaprepes Abbreviatus (L.)"**, *Proc. Fla. Soc.*, Vol.116, 2003

- V. Soroker, Y. Nakache, U. Landau, A. Mizrach, A. Hetzroni and D. Gerling **"Utilization of Sounding Methodology to Detect Infestation by Rhynchophorus ferrugineus on Palm Offshoots"**, Phytoparasitica, Vol.32, Part 1, p.p. 6-8, 2004
- Vlad M. Trifa, **"A framework for bird songs detection, recognition and localization using acoustic sensor networks"**, Master thesis, UCLA, February 2006
- Xianxian Zhang, John H.L.Hansen, Pongtep Angkititrakul, Kazuya Takeda **"Speaker Verification Using Gaussian Mixture Models Within Changing Real Car Environments"**, Interspeech, p.p.2021-2024,2005,
- Διγαλάκης Βασίλης. Σημειώσεις μαθήματος: **«Εισαγωγή στην επεξεργασία Φωνής»**, Τμήμα Ηλεκτρολόγων Μηχανικών Και Μηχανικών Υπολογιστών. Πανεπιστήμιο Κρήτης
- Θυμάκης Ν., Οικονόμου Δ., Χ. Μεντή, Δ. Κοντοδήμας και Μ. Ανάγου – Βερονίκη. **«Είδη της οικογένειας Curculionidae (Coleoptera) που προσβάλλουν τα φοινικοειδή.»**, Πρακτικά Ημερίδας «Επιζήμια Έντομα και ακάρεα αστικού πρασίνου», Κτήμα Συγγρού, Μαρούσι 15 δεκ.2005
- Κοντοδήμας, Δ. - Οικονόμου , Δ. - θυμάκης, Ν. - Χ.Μεντή και Ανάγνου – Βερονίκη, Μ. **«Νέος σοβαρός εχθρός των φοινικοειδών, το κολεόπτερο Rhynchophorus Ferrugineus (Olivier) (Curculionidae: Dryophorinae)»**, Γεωργία και κτηνοτροφία, τεύχος 1, 54-58, 2006
- Κωνσταντίνος Γ. Βολουδάκης Κωνσταντίνος Α. Μαρμαροκόπος **«Μοντελοποίηση και Αναγνώριση ανθρώπινων Χειρονομιών στην Επικοινωνία Ανθρώπου - Υπολογιστή»** Διπλωματική εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή ηλεκτρολόγων μηχανικών και μηχανικών υπολογιστών, τομέας τεχνολογίας πληροφορικής και υπολογιστών, Αθήνα, Ιούλιος 2005
- Μπλέτσας Α.Α, **"Ελληνικό Σύστημα Αυτόματης Εκφώνησης Κειμένου (Text To -Speech System)"**, Διπλωματική εργασία, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Πολυτεχνική Σχολή, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Νοέμβριος 1998
- Περακάκης Μανόλης, **«Κωδικοποίηση φωνής για αναγνώριση σε ενσύρματα και ασύρματα δίκτυα δεδομένων (Διαδίκτυο , GPRS) μέσα από δίαυλο 2 Kbps»** Πολυτεχνείο Κρήτης, τμήμα ηλεκτρολόγων μηχανικών και μηχανικών υπολογιστών, εργαστήριο τηλεπικοινωνιών πληροφορίας και δικτύων, Διπλωματική εργασία, Δεκέμβριος 2000

ΙΣΤΟΣΕΛΙΔΕΣ

- <http://www.aeconsulting.com>
- <http://www.avisoft.com>
- <http://www.bioacoustic.info>
- http://www.eppo.org/quarantine/Albert_List/insects/rhyefe.htm
- <http://www.laartech.biz>
- <http://www.redpalmweevil.com>
- <http://www.wikipedia.org>
- <http://www.pelagosinstitute.gr/gr/research/dinamiko.html>