



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΜΟΥΣΙΚΩΝ ΣΠΟΥΔΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ «ΜΟΥΣΙΚΟΛΟΓΙΑ»
ΚΑΤΕΥΘΥΝΣΗ: ΜΟΥΣΙΚΗ ΤΕΧΝΟΛΟΓΙΑ

**birdMelos: Δημιουργία εφαρμογής για την αυτόματη
αναγνώριση των φωνημάτων των πτηνών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΕΩΡΓΙΟΥ ΚΟΥΤΣΟΓΙΑΝΝΗ

Επιβλέπουσα: Αναστασία Γεωργάκη
Αναπληρώτρια Καθηγήτρια Ε.Κ.Π.Α.

Αθήνα, Μάρτιος 2019

ΠΙΣΤΟΠΟΙΗΣΗ

Πιστοποιείται ότι η Διπλωματική Εργασία με θέμα

**«birdMelos: Δημιουργία εφαρμογής για την αυτόματη αναγνώριση των
φωνημάτων των πτηνών»**

Του μεταπτυχιακού φοιτητή του Τμήματος Μουσικών Σπουδών

ΚΟΥΤΣΟΓΙΑΝΝΗ ΓΕΩΡΓΙΟΥ του ΣΤΑΥΡΟΥ

Με Αριθμό Μητρώου: 170305

Παρουσιάστηκε δημόσια και εξετάστηκε στο Τμήμα Μουσικών Σπουδών στις

..... / /

Η Επιβλέπουσα Καθηγήτρια

.....

Η Επιτροπή

.....

Ευχαριστίες

Θα ήθελα να ευχαριστήσω ιδιαίτερα την επιβλέπουσα καθηγήτριά μου κα Αναστασία Γεωργάκη για την αμέριστη υποστήριξη και καθοδήγησή της καθ' όλη τη διάρκεια της φοίτησής μου στο ΠΜΣ. Επίσης ευχαριστώ θερμά τους καθηγητές μου κ. Γιώργο Κωστελέτο και κα Αρετή Ανδρεοπούλου για τη γνώση που μου προσέφεραν και την άριστη συνεργασία που είχαμε. Τέλος, ευχαριστώ τον Μάξιμο Καλιακάτσο-Παπακώστα για τις πολύτιμες τεχνικές πληροφορίες που μου έδωσε.

Περίληψη

Το «τραγούδι» των πτηνών έχει επηρεάσει διαχρονικά τους συνθέτες και τους μουσικούς με ποικίλους τρόπους, ενώ παράλληλα έχουν παρατηρηθεί αναλογίες μεταξύ της «μουσικής» των πουλιών και της ανθρώπινης μουσικής. Ένα σημαντικό ζήτημα που προκύπτει είναι εάν μέσα από την οργάνωση των ήχων τους διαφαίνεται κάποια άλλη λειτουργία, πέραν της βιολογικής. Σε τέτοιου είδους ερωτήματα προσπαθεί να δώσει απαντήσεις η επιστήμη της μουσικολογίας μέσα από τους διάφορους κλάδους της: τη βιομουσικολογία, τη ζωομουσικολογία και την ορνιθομουσικολογία. Αυτή η προσπάθεια αποκωδικοποίησης της φωνητικής επικοινωνίας των πτηνών γίνεται μέσα από διάφορα επιστημονικά πεδία, ενώ παράλληλα, αντικείμενο προς συζήτηση αποτελεί η συμβολή της τεχνολογίας σε αυτή. Τα τελευταία χρόνια, εταιρείες σε συνεργασία με εργαστήρια βιοακουστικής έχουν παρουσιάσει λογισμικά ακουστικής ανάλυσης των καλεσμάτων των πτηνών, μέσω αυτόματης αναγνώρισης προτύπων, που ταυτοποιούν τη ρυθμική και μελωδική δομή του «τραγουδιού» τους σύμφωνα με τη βιολογική τους κατάταξη. Παρότι οι εφαρμογές αυτές ακολουθούν διαφορετική μεθοδολογία, παρουσιάζουν κάποιους κοινούς περιορισμούς. Για τους σκοπούς της παρούσας εργασίας δημιουργήθηκε η πρωτότυπη εφαρμογή birdMelos, η οποία έχει ως αντικείμενο να διερευνήσει τις μουσικές παραμέτρους της φωνητικής επικοινωνίας των πτηνών και να προσφέρει ένα εργαλείο για την αναγνώριση του είδους. Ο σχεδιασμός της έγινε στην πλατφόρμα του MATLAB, με χρήση της βιβλιοθήκης MIRtoolbox 1.7.1, και επικεντρώνεται σε τρία είδη των οποίων τα καλέσματα έχουν σαφείς και διακριτές δομές: στον γκιώνη, στον κούκο και στη δεκαοχτούρα. Το birdMelos έχει τη δυνατότητα να πραγματοποιεί ανάλυση ήχου σε οιοδήποτε δείγμα δεδομένων του εισάγεται - στερεοφωνικό ή μονοφωνικό - και να προβαίνει σε αναγνώριση του είδους του πτηνού, ενώ ταυτόχρονα χρησιμοποιεί προσαρμοσμένα φίλτρα μείωσης θορύβου προς βελτίωση της ακρίβειας της αναγνώρισης.

Λέξεις - κλειδιά:

Ακουστική οικολογία, αναγνώριση πτηνών, ανάλυση ήχου, βιομουσικολογία, μηχανική μάθηση, μουσική τεχνολογία, φωνήματα πτηνών.

Abstract

Throughout the ages birdsong has influenced composers and musicians in various ways, while at the same time have been observed analogies between birds' "music" and human music. A major issue is whether in birdsong there is some other use than the biological one. The science of musicology seeks answers through its various fields: biomusicology, zoomusicology and ornithomusicology. The attempt to decode bird vocalizations is done through various scientific fields, whereas the contribution of technology to this effort constitutes a subject of discussion. During the last few years, companies in association with bioacoustics laboratories have introduced particular acoustical analysis software for bird calls, via the so-called automatic pattern recognition method, which identifies with both the rhythmic and the melodic structure of birdsong, always in accordance with the species' biological classification. Although such applications are based on an entirely different methodology, they appear to have some common limitations. For the purposes of this work, the birdMelos novel application was developed to explore the musical parameters of the vocal communication in birds and give a tool for identifying the structure of the birdsong related to the specie. Its design was based on the MATLAB platform, using the MIRtoolbox 1.7.1 library, and focuses on three types whose calls have clear and distinct structures: otus scops, cuculus canorus and streptopelia decaocto. BirdMelos is capable of performing sound analysis on any given sample - stereo or mono - and identifying the species, while also implementing custom noise reduction filters to improve the classification rate performance.

Key Words:

Acoustic ecology, biomusicology, bird vocalization, birdsong recognition, machine learning, music technology, sound analysis.

ΠΕΡΙΕΧΟΜΕΝΑ

Εισαγωγή	9
1. Τα φωνήματα των πτηνών: διεπιστημονικές προσεγγίσεις	11
1.1 Γενικά χαρακτηριστικά των ωδικών πτηνών	12
1.1.2 Τρόποι επικοινωνίας	12
1.2 Το birdsong στο πεδίο της Ακουστικής	14
1.3 Η μελέτη του birdsong στη Γνωσιακή Νευροεπιστήμη και στην Εξελικτική Νευροβιολογία	15
1.4 Η σχέση του birdsong με τη γλώσσα	17
2. Το birdsong υπό το πρίσμα της μουσικολογίας	20
2.1 Αποτελεί το birdsong μουσική;.....	20
2.2 Βιομουσικολογία.....	21
2.3 Ζωομουσικολογία.....	21
2.3.1 François-Bernard Mâche	22
2.4 Ορνιθομουσικολογία.....	23
2.4.1 Olivier Messiaen.....	23
2.5 Birdsong και μουσική δημιουργία.....	25
2.6 Μουσικές δομές στη φωνητική επικοινωνία των πτηνών	26
3. Αυτόματη αναγνώριση του birdsong	31
3.1 Αναγνώριση Προτύπων	31
3.1.1 Προεπεξεργασία – Αφαίρεση Θορύβου	32
3.1.2 Κατάτμηση (Segmentation)	32
3.1.3 Τεχνικές Ανάλυσης - Εξαγωγή Χαρακτηριστικών (Feature Extraction)	33
3.1.4 Τεχνικές Ταξινόμησης	35

3.2 Συνήθη προβλήματα που προκύπτουν κατά την αναγνώριση	38
3.3 Τα λογισμικά	39
3.3.1 Λογισμικά ακουστικής ανάλυσης και αναγνώρισης.....	39
3.3.2 Εφαρμογές αναγνώρισης του birdsong αποκλειστικά για έξυπνα κινητά τηλέφωνα (smartphones)	41
3.4 Βάσεις δεδομένων – βιβλιοθήκες ήχων	42
4. Η δημιουργία της εφαρμογής birdMelos	43
4.1 Προεπεξεργασία	44
4.2 Εξαγωγή χαρακτηριστικών.....	46
4.2.1 Ο κώδικας.....	54
4.3 Η επιλογή του ταξινομητή	55
4.4 Δοκιμές - αποτελέσματα.....	57
4.5 Η δημιουργία του interface.....	59
4.5.1 Άνοιγμα της εφαρμογής.....	59
4.5.2 search file Button.....	60
4.5.3 play / stop Buttons	61
4.5.4 recognise Button.....	61
4.5.5 info Button	62
4.6 Σύνοψη και συμπεράσματα	63
5. Επίλογος.....	64
Βιβλιογραφία	65
Παραρτήματα.....	71

Εισαγωγή

Η παρούσα εργασία αποσκοπεί σε μία συνολική εξέταση του ζητήματος της μελέτης των ήχων των πουλιών από τον άνθρωπο. Αυτό γίνεται μέσα από δύο άξονες, τον άξονα της Μουσικολογίας και τον άξονα της Τεχνολογίας. Εκτός από το αμιγώς θεωρητικό σκέλος, τμήμα της εργασίας αποτελεί και η δημιουργία μίας πρωτότυπης εφαρμογής. Η εργασία απαρτίζεται από τέσσερα κεφάλαια. Το πρώτο κεφάλαιο (*Οι ήχοι των πουλιών: διεπιστημονικές προσεγγίσεις*) αναφέρεται στην μελέτη των ήχων των πουλιών μέσα από διαφορετικά επιστημονικά πεδία. Παρουσιάζονται γενικά χαρακτηριστικά της φωνητικής επικοινωνίας των πουλιών και στη συνέχεια δίνεται έμφαση στην εξέταση των φωνημάτων τους μέσα από το πρίσμα της Ακουστικής, της Νευροεπιστήμης και της Γλωσσολογίας.

Στο δεύτερο κεφάλαιο (*Το birdsong υπό το πρίσμα της μουσικολογίας*) το ζήτημα της «μουσικής» των πουλιών εξετάζεται μέσα από την οπτική της μουσικολογίας και των διάφορων κλάδων της: της βιομουσικολογίας, της ζωομουσικολογίας και της ορνιθομουσικολογίας. Αναφέρεται παράλληλα το έργο και η συμβολή των κύριων εκφραστών των πεδίων αυτών. Στη μουσική, το «birdsong» έχει επηρεάσει διαχρονικά τους συνθέτες και τους μουσικούς με ποικίλους τρόπους: είτε έχοντας αποτελέσει πηγή έμπνευσης, είτε έχοντας χρησιμοποιηθεί ο σκοπός τους μέσα σε μουσικά έργα. Μέσα από την εργασία αυτή αναδεικνύεται όλη αυτή η πορεία ιστορικά, και διαφαίνεται ο ρόλος που επιτέλεσε η εξάπλωση της τεχνολογίας στη συνθετική διαδικασία. Επίσης, περιγράφονται οι αναλογίες που υπάρχουν μεταξύ της μουσικής των πουλιών και της ανθρώπινης μουσικής και αναλύονται ζητήματα όπως οι μουσικές δομές στην επικοινωνία των πουλιών, καθώς και τρόποι ανάλυσης και επεξεργασίας των ήχων τους.

Στο τρίτο κεφάλαιο (*Αυτόματη αναγνώριση του birdsong*) εξετάζονται οι τάσεις της έρευνας αναφορικά με τους ήχους των πουλιών, με κεντρικό άξονα τη συμβολή της τεχνολογίας και των νέων μέσων σε αυτή. Βασικοί άξονες της συγκεκριμένης ενότητας είναι: περιγραφή μιας τυπικής διαδικασίας αναγνώρισης προτύπων, τεχνικά ζητήματα που αφορούν την επεξεργασία των δεδομένων, εξαγωγή ακουστικών χαρακτηριστικών από τα δεδομένα αυτά, κατηγοριοποίηση και ταξινόμηση. Ιδιαίτερη έμφαση δίδεται στις πρόσφατες έρευνες που αφορούν εφαρμογές ακουστικής ανάλυσης των καλεσμάτων των πτηνών, οι οποίες μέσω αυτόματης αναγνώρισης προτύπων ταυτοποιούν τη ρυθμική και μελωδική δομή του «τραγουδιού» τους σύμφωνα με τη βιολογική τους κατάταξη. Μέσω συγκριτικής μελέτης παρουσιάζονται οι βασικότερες τεχνικές που χρησιμοποιούν αυτά τα λογισμικά. Βέβαια, όπως αναφέρεται και στο κείμενο, ένα απόλυτα ασφαλές συμπέρασμα από αυτή τη σύγκριση δεν είναι εφικτό, για διάφορους λόγους. Τέλος αξιολογούνται οι υπάρχουσες εφαρμογές αναγνώρισης και

παρουσιάζονται οι διαδικτυακές βιβλιοθήκες ήχων που παρέχουν την δυνατότητα άντλησης δεδομένων από αυτές.

Το τέταρτο κεφάλαιο (*Η δημιουργία της εφαρμογής birdMelos*) αφορά στο πρακτικό σκέλος της εργασίας αυτής: τη δημιουργία μίας πρωτότυπης υπολογιστικής εφαρμογής αναγνώρισης των καλεσμάτων των πουλιών. Πρόκειται για το πιο απαιτητικό τμήμα της εργασίας, την εμβάθυνση δηλαδή στις τεχνικές της μηχανικής μάθησης και της επιστήμης των υπολογιστών. Με γνώμονα προηγούμενες έρευνες στο πεδίο αυτό, χρησιμοποιήθηκαν δεδομένα που ανακτήθηκαν από το διαδίκτυο, αφενός για να δοκιμαστεί η απόδοση διαφόρων αλγορίθμων ταξινόμησης, και αφετέρου για να εξεταστούν διάφορα χαρακτηριστικά του ήχου ως προς την καταλληλότητά τους. Στο τελικό στάδιο, επιλέχθηκε ο αλγόριθμος ταξινόμησης που πέτυχε τα μεγαλύτερα ποσοστά ακρίβειας, και στη συνέχεια δημιουργήθηκε ένα εύχρηστο interface. Μάλιστα ο συνδυασμός των χαρακτηριστικών και του ταξινομητή που χρησιμοποιήθηκε αποτελεί καινοτομία της εφαρμογής. Το τελικό αποτέλεσμα με βάση την ακρίβεια στην αναγνώριση των πουλιών και την ευκολία στη χρήση, κρίνεται επιτυχές. Παράλληλα, δημιουργούνται νέες προοπτικές για την βελτίωση της εφαρμογής και την προσθήκη νέων λειτουργιών και δυνατοτήτων. Συγκεκριμένα η προσθήκη μεγαλύτερου εύρους πουλιών – συμπεριλαμβανομένων και ειδών της ελληνικής ορνιθοπανίδας - αποτελεί αντικείμενο μελλοντικής έρευνας, ενώ ενδιαφέρον επίσης θα παρουσίαζε μια ενδεχόμενη τροποποίηση της εφαρμογής ώστε να μπορεί να εξάγει μουσικά χαρακτηριστικά του ήχου τους, όπως τονικότητα (κλίμακα) και ρυθμικά στοιχεία ή ακόμη να δημιουργεί «αυτόματες χορωδίες πουλιών».

Συνοψίζοντας, οι στόχοι που φιλοδοξούμε να πετύχουμε μέσα από την εργασία αυτή είναι καταρχάς να παρουσιαστούν διεξοδικά και με εύγλωττο τρόπο οι βασικότερες εκφάνσεις της μελέτης των ήχων των πουλιών, με έμφαση στο κομμάτι της Ακουστικής και της Μουσικολογίας. Όσον αφορά ειδικά στο τεχνολογικό κομμάτι, σήμερα – όταν ζητήματα όπως αυτά που άπτονται του πεδίου της Ακουστικής Οικολογίας συζητούνται έντονα, επιτάσσεται η συνεργασία μεταξύ διαφορετικών επιστημονικών κλάδων. Πιο συγκεκριμένα, μέσα από αυτή την έρευνα αναδεικνύεται η σχέση που μπορεί να έχουν αμιγώς μουσικολογικά ζητήματα με άλλα θεματικά πεδία, όπως για παράδειγμα από τον κλάδο των Τηλεπικοινωνιών ή της Πληροφορικής. Είναι γεγονός ότι μόνο το προηγούμενο μουσικολογικό υπόβαθρο ίσως να μην επαρκούσε για την ολοκλήρωση αυτής της εργασίας· συνέβαλαν και άλλοι παράγοντες. Μερικοί από αυτούς είναι εξειδικευμένες γνώσεις και δεξιότητες που αποκτήθηκαν μέσα από τον κύκλο σπουδών του ΠΜΣ Μουσικής Τεχνολογίας, η προηγούμενη ενασχόληση με τον προγραμματισμό αλλά και προσωπικός κόπος και όρεξη για μάθηση.

1. Τα φωνήματα των πτηνών: διεπιστημονικές προσεγγίσεις

Η πρώτη γνωστή καταγραφή του «τραγουδιού» των πουλιών (birdsong¹) έγινε το 1889 από τον Ludwig Koch ενώ η σύγχρονη μελέτη του ξεκίνησε με το έργο του ζωολόγου William Thorpe (Thorpe, 1958). Τα χρόνια που μεσολάβησαν συντελέστηκαν πάρα πολλά στον τομέα αυτό, μέσα από διαφορετικά επιστημονικά πεδία. Στον διεπιστημονικό κλάδο της Ακουστικής Οικολογίας, ο Schafer και οι συνεργάτες του στο πλαίσιο του «World Soundscape Project» (Schafer, 1977) κατέγραψαν τους ήχους που μας περιβάλλουν, συμπεριλαμβανομένου και του κελαηδίσματος των πουλιών. Το διεπιστημονικό πεδίο της Βιοακουστικής έχει ασχοληθεί με τη διερεύνηση της παραγωγής και της πρόσληψης του ήχου στα πουλιά. Αφετηρία για αυτού του είδους τις έρευνες υπήρξαν οι μελέτες του Σλοβένου βιολόγου Ivan Regen, το 1925, σχετικά με τους ήχους των εντόμων. Στην επιστήμη της Βιολογίας η έρευνα και η καταγραφή του τραγουδιού των πουλιών χρησιμοποιείται για την παρακολούθηση του οικοσυστήματος αναφορικά με τον πληθυσμό τους. Τα τελευταία χρόνια τη σκυτάλη για τη μελέτη των καλεσμάτων των πουλιών έχει πάρει από την επιστήμη της Μουσικής Ακουστικής η επιστήμη της Εξελικτικής Νευροβιολογίας όπου ένα από τα κύρια ερευνητικά ενδιαφέροντα της είναι η συγκριτική προσέγγιση της μουσικής συμπεριφοράς που σχετίζεται με τη μάθηση στον ανθρώπινο εγκέφαλο και στον εγκέφαλο των πουλιών. Παράλληλα, η Γνωσιακή Νευροεπιστήμη ασχολείται με την κατανόηση των νευρικών μηχανισμών που εμπλέκονται στην εκμάθηση του κελαηδίσματος στα πουλιά, με στόχο την κατανόηση της εκμάθησης της γλώσσας. Τέλος, στον τομέα της Πληροφορικής, νέα υπολογιστικά συστήματα στοχεύουν στην ανάλυση, αναγνώριση και επεξεργασία του ήχου των πουλιών ενώ οι τεχνικές που χρησιμοποιούνται είναι παρόμοιες

¹ Από εδώ και στο εξής θα προτιμάται η χρήση του όρου «birdsong» γιατί κρίνεται ως ο πιο δόκιμος όλων, αφού άλλωστε χρησιμοποιείται κατά κόρον στη διεθνή βιβλιογραφία. Κάποιοι παρεμφερείς όροι που χρησιμοποιούνται, κρίνονται ως ανεπαρκείς στο να αποδώσουν πλήρη και σωστή περιγραφή. Συγκεκριμένα εάν μιλήσουμε για «ήχους» των πουλιών, θα πρέπει να διευκρινίσουμε εάν αναφερόμαστε μόνο στους φωνητικούς ήχους ή εάν συμπεριλαμβάνουμε και τους μη φωνητικούς ήχους που χρησιμοποιούν τα πουλιά για να επικοινωνούν. Εάν επίσης αναφερθούμε γενικά σε «κάλεσμα» ή «καλέσματα» των πουλιών, περιορίζουμε το σύνολο της φωνητικής επικοινωνίας των πουλιών μέσα σε ένα υποσύνολό της, δηλαδή στα καλέσματα (calls), τα οποία είναι σύντομες, απλές σε δομή συνθέσεις, συχνά έντονες και χωρίς μουσικότητα. Συνεπώς θα έχουμε αποκλείσει τα τραγούδια (songs), τα οποία είναι μεγαλύτερες και πολυπλοκότερες συνθέσεις, και τα οποία είναι τα πιο ενδιαφέροντα στα ωδικά πτηνά. Ακριβώς το αντίθετο συμβαίνει με τον όρο «κελάηδισμα», ο οποίος εμπεριέχει ένα στοιχείο μουσικότητας και επικεντρώνεται κυρίως στα songs. Ενώ μια αυστηρή ερμηνεία του όρου «birdsong» αφορά στα παραγόμενα songs των πουλιών και όχι στα calls, στο πεδίο της Αναγνώρισης του Birdsong (Birdsong Recognition) περιλαμβάνονται και τα calls στην αναγνώριση. Οπότε κρίνεται σωστότερο στην προκειμένη περίπτωση να προτιμηθεί ο ξένος όρος έναντι του ελληνικού. Τέλος, η χρήση των όρων «τραγούδι» και «μουσική» των πουλιών, είναι μια προσπάθεια ερμηνείας βασισμένη στο ανθρώπινο σύστημα αναφοράς και δεν προτιμάται, εκτός εάν οι όροι αυτοί αποδίδονται μέσα σε εισαγωγικά. Όπως αναλύεται και στο 2^ο κεφάλαιο, η παραγωγή και η οργάνωση των φωνημάτων των πουλιών (bird vocalizations) σχετίζεται – σύμφωνα με την επικρατούσα άποψη - με τη μεταφορά πληροφορίας, όπως συμβαίνει και στην ανθρώπινη ομιλία.

με αυτές που χρησιμοποιούνται στην Αυτόματη Αναγνώριση Ομιλίας² στις Τηλεπικοινωνίες.

1.1 Γενικά χαρακτηριστικά των ωδικών πτηνών

Από τα 8.700 είδη πουλιών, περίπου τα 4.000 με 5.000 χαρακτηρίζονται ως «ωδικά πτηνά» (songbirds). Πρόκειται για μία οικογένεια πουλιών τα οποία ανήκουν στην τάξη των Στρουθιόμορφων (Passeriformes) και διακρίνονται για το ιδιαίτερα ανεπτυγμένο φωνητικό τους όργανο, τη «Σύριγγα» (Syrinx), το οποίο τους δίνει τη δυνατότητα παραγωγής μελωδικών ήχων. Η Σύριγγξ παρουσιάζει διαφορές τόσο από το ανθρώπινο φωνητικό όργανο όσο και ανάμεσα στις διαφορετικές κατηγορίες πουλιών. Όσο περισσότερο ανεπτυγμένη είναι, τόσο μεγαλύτερη δυνατότητα έχουν τα πουλιά για παραγωγή πολυπλοκότερων ήχων. Περίπου 200 με 300 είδη ωδικών πτηνών έχουν κινήσει το ενδιαφέρον της επιστημονικής κοινότητας εξαιτίας της ποικιλίας των ηχητικών τους σημάτων (Mâche, 1983).

1.1.2 Τρόποι επικοινωνίας

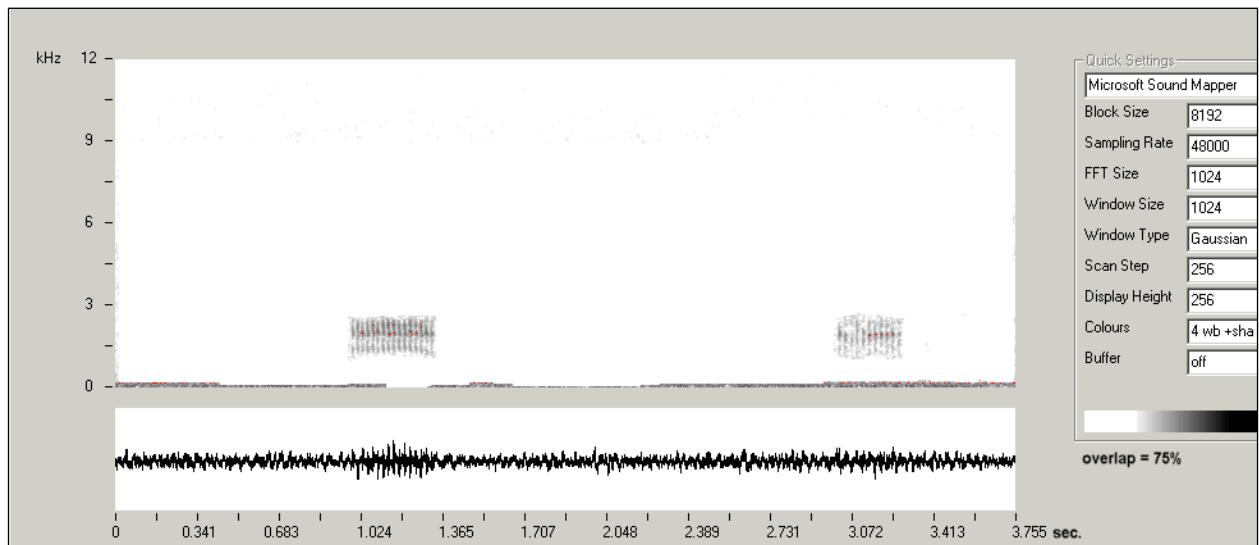
Τα πουλιά επικοινωνούν είτε χρησιμοποιώντας το φωνητικό τους όργανο είτε με μη φωνητικούς τρόπους. Όσον αφορά στη μη φωνητική επικοινωνία, αυτή παρατηρείται συχνότερα σε μη μελωδικά πτηνά που έχουν λιγότερο ανεπτυγμένη Σύριγγα, τα οποία δημιουργούν ρυθμικά μοτίβα και άλλους ήχους χωρίς τη χρήση του φωνητικού τους οργάνου. Τέτοιοι μη φωνητικοί ήχοι είναι, παραδείγματος χάρη, το ρυθμικό χτύπημα του ράμφους του δρυοκολάπτη και το κροτάλισμα του ράμφους της κουκουβάγιας.

Η φωνητική επικοινωνία εξυπηρετεί δύο βασικές λειτουργίες στα πουλιά· είτε χρησιμοποιείται από αρσενικά και θηλυκά πουλιά προς οριοθέτηση μίας περιοχής, είτε χρησιμοποιείται από αρσενικά πουλιά προς προσέλκυση θηλυκών κατά την αναπαραγωγική περίοδο. Μόνο σε λίγα είδη παράγουν birdsong τα θηλυκά πουλιά· στις περιπτώσεις που αυτό συμβαίνει, υπάρχει μεγάλη ποικιλομορφία στην πολυπλοκότητα του birdsong των θηλυκών σε σχέση με εκείνο των ομοειδών αρσενικών. Το τραγούδι στα θηλυκά φτάνει στο μέγιστο βαθμό της απόδοσής του σε ορισμένα τροπικά είδη, στα οποία αρσενικά και θηλυκά συνδράμουν εξίσου στα φωνητικά ντουέτα. Γενικότερα, στα είδη των πουλιών της τροπικής ζώνης το birdsong χρησιμοποιείται συχνά για οριοθέτηση περιοχών, καθ' όλη τη διάρκεια του έτους.

Η φωνητική επικοινωνία στα πουλιά, διακρίνεται σε δύο κατηγορίες: στα καλέσματα (calls) και στα τραγούδια (songs). Τα καλέσματα (βλ. Εικόνα 1) είναι απλές φωνητικές συνθέσεις μικρής διάρκειας και παράγονται από πουλιά, ανεξαρτήτως ηλικίας και φύλου, καθ' όλη τη διάρκεια του έτους. Διακρίνονται με βάση την λειτουργικότητα τους σε καλέσματα

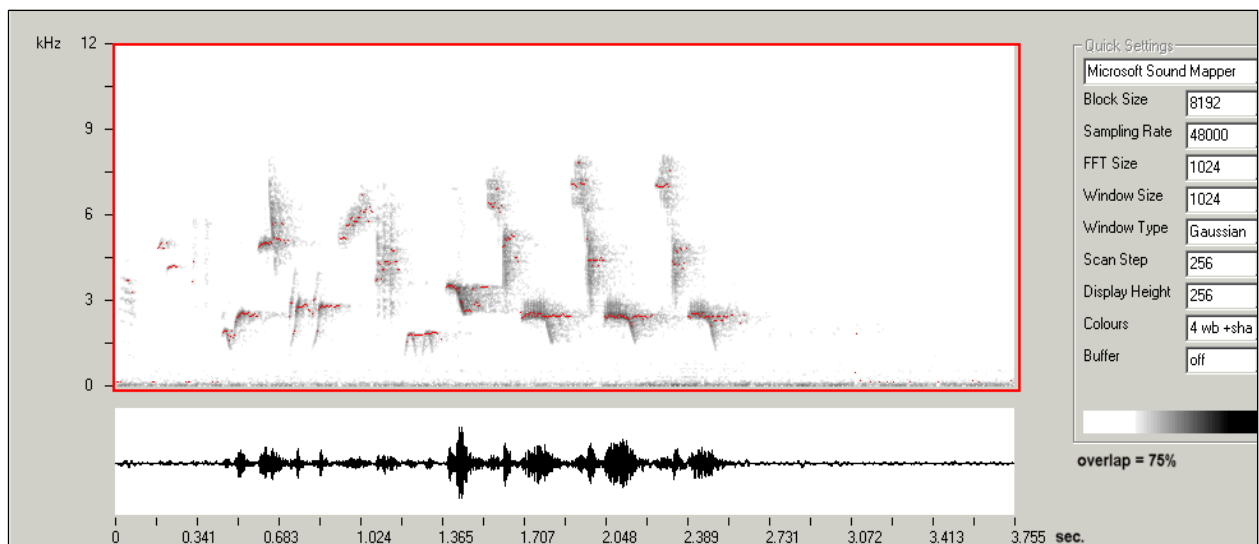
² Βλ. Automatic Speech Recognition (ASR).

προειδοποίησης («alarm calls»), καλέσματα φωλιάς («nest calls»), καλέσματα τροφής («feeding calls») κ.ά. Τα τραγούδια (βλ. Εικόνα 2) είναι πολύπλοκες φωνητικές συνθέσεις μεγαλύτερης διάρκειας, με σαφείς μουσικές δομές και μεγαλύτερο τονικό και ρυθμικό εύρος, και παράγονται σχεδόν αποκλειστικά από αρσενικά πουλιά. Συνήθως ξεκινούν με εισαγωγικές νότες τις οποίες διαδέχονται ένα ή δύο μοτίβα, τα οποία είναι επαναλαμβανόμενες νότες ή συλλαβές³. Το τραγούδι παρατηρείται κυρίως κατά την άνοιξη αλλά και κατά την αρχή του καλοκαιριού, ιδίως τις πρωινές ώρες και τεχνικά προκύπτει μόνο από τα ωδικά πτηνά.



Εικόνα 1: Φασματογράφημα και κυματομορφή ενός καλέσματος «απειλής» (αηδόνι)

(Πηγή: www.birdsongs.it/songs/luscinia_megarhynchos/spectro2.html)



Εικόνα 2: Φασματογράφημα και κυματομορφή από τραγούδι ενήλικου αρσενικού πουλιού (αηδόνι)

(Πηγή: www.birdsongs.it/songs/luscinia_megarhynchos/spectro3.html)

³ Ως «συλλαβή» ορίζεται ένας αδιάλειπτος ήχος, ο οποίος εμπεριέχει ένα ή περισσότερα ηχητικά ίχνη χρόνου-συχνότητας τα οποία αποκαλούνται «νότες» (Berwick et al., 2011).

Ενδιαφέρουσες ομοιότητες και αναλογίες προκύπτουν αναφορικά με τη χρήση ηχητικών σημάτων ανάμεσα στα πουλιά και στους ανθρώπους. Παραδείγματος χάρη τα καλέσματα «κινδύνου» κάποιων πουλιών αποτελούνται από ήχους που μεταβαίνουν από υψηλότερες συχνότητες σε χαμηλότερες, ενώ τα καλέσματα που παράγονται κατά την εποχή του ζευγαρώματος χρησιμοποιούν κυρίως υψηλές συχνότητες. Αντίστοιχα, παρόμοιες μορφές ηχητικής έκφρασης παρατηρούνται στην έκφραση της λύπης και της απόλαυσης στους ανθρώπους (Schafer, 1977).

1.2 Το birdsong στο πεδίο της Ακουστικής

Στην Ακουστική, και δη στη Μουσική Ακουστική, έχει εξεταστεί η λειτουργία της Σύριγγας και του ακουστικού συστήματος των πουλιών και έχουν προκύψει ενδιαφέρουσες παρατηρήσεις. Ιδιαίτερα όσον αφορά στη Σύριγγα έχει πλέον αποδειχτεί ότι δεν προσομοιάζει με κανένα μουσικό όργανο, αλλά ούτε και με την ανθρώπινη φωνή. Τα τραγούδια που παράγει είναι τόσο περίτεχνη δομή και περιλαμβάνουν τόσο απότομες μεταβολές του πλάτους, ώστε καμία ανθρώπινη επίδοση δεν μπορεί να πετύχει τέτοια πολυπλοκότητα φράσεων (Σπυρίδης, 1996). Ανάλογα συμπεράσματα προκύπτουν και για την ακοή των πουλιών.

Η Ακουστική Οικολογία είναι ένας διεπιστημονικός κλάδος, ο οποίος εξετάζει τη σχέση ανάμεσα στους ανθρώπους και στο περιβάλλον τους μέσω του ήχου. Οι μελέτες στο πεδίο αυτό ξεκίνησαν κατά τα τέλη της δεκαετίας του 1960 από τον Καναδό συνθέτη, συγγραφέα και μουσικοπαιδαγωγό R. Murray Schafer και την ομάδα του, στο Πανεπιστήμιο Simon Fraser. Ο Schafer προσπάθησε να κατανοήσει και να καταγράψει τις επιπτώσεις του αυξανόμενου θορύβου στο περιβάλλον της πόλης του Βανκούβερ, στον Καναδά, και η μέθοδος που χρησιμοποίησε γι' αυτή την καταγραφή ήταν οι ηχητικοί περίπατοι (soundwalks)⁴. Κατά την μελέτη αυτή, που ονομάστηκε «The Vancouver Soundscape», καταγράφηκαν τα επίπεδα θορύβου σε διάφορα σημεία της πόλης. Η καταγραφή του ηχοτοπίου του Βανκούβερ αποτελεί τμήμα μίας ευρύτερης μελέτης με τίτλο «World Soundscape Project», με την οποία εισάγεται επίσημα το πεδίο της Ακουστικής Οικολογίας (Schafer, 1977). Επίσης, ένα ζήτημα που απασχόλησε ιδιαίτερα τον Schafer ήταν η σχέση των παιδιών με τον κόσμο μέσω του ήχου, και οι συνέπειες των αλλαγών στα ηχητικά τοπία των βιομηχανικών περιοχών στα παιδιά. Τέλος, η Ακουστική Οικολογία έχει βρει εφαρμογή σε πολλούς διαφορετικούς τομείς. Κάποιες από τις ευρύτερες εφαρμογές της αφορούν στις ηχητικές επιπτώσεις της κατασκευής δρόμων και αεροδρομίων, στην καταγραφή των ήχων των ωκεανών κ.ά.

⁴ Ο «ηχητικός περίπατος» είναι ένας περίπατος με εστίαση στην ακρόαση του περιβάλλοντος. Ο όρος χρησιμοποιήθηκε για πρώτη φορά από μέλη του World Soundscape Project στη δεκαετία του '70. Ο Hildegard Westerkamp, από την ίδια ομάδα καλλιτεχνών, ορίζει το soundwalking ως «κάθε εκδρομή που έχει ως κύριο στόχο την ακρόαση του περιβάλλοντος».

Η Βιοακουστική είναι ένα διεπιστημονικό πεδίο που συνδυάζει τη Βιολογία και την Ακουστική, το οποίο σχετίζεται με τη διερεύνηση της παραγωγής, της διασποράς και της πρόσληψης του ήχου στα ζώα - συμπεριλαμβανομένων και των ανθρώπων. Αυτό σχετίζεται με τη νευροφυσιολογική και ανατομική βάση της παραγωγής και ανίχνευσης του ήχου, και τη σχέση των ακουστικών σημάτων με το μέσο με το οποίο διαχέονται. Τα ευρήματα παρέχουν ενδείξεις σχετικά με την εξέλιξη των ακουστικών μηχανισμών και, κατά συνέπεια, την εξέλιξη των ζώων που τους χρησιμοποιούν. Η Βιοακουστική ως επιστημονικός κλάδος δημιουργήθηκε από τον Σλοβένο βιολόγο Ivan Regen, ο οποίος κατά το 1925 άρχισε να μελετά συστηματικά τους ήχους των εντόμων. Μία συζήτηση σχετικά με τα είδη των πουλιών μέσα από το πρίσμα της Βιοακουστικής αφορά στα αυξημένα επίπεδα θορύβου, ως συνέπεια της αυξημένης αστικοποίησης και ανάπτυξης. Είναι γεγονός ότι ο θόρυβος μπορεί να μεταβάλει το ακουστικό περιβάλλον των υδάτινων και χερσαίων οικοτόπων. Η ποικιλία των πτηνών έχει παρουσιάσει μείωση λόγω των χρόνιων επιπέδων θορύβου στις πόλεις και κατά μήκος των οδικών αρτηριών. Επίσης, ορισμένα είδη πουλιών αναγκάστηκαν να αλλάξουν τις συχνότητες του birdsong τους προκειμένου να προσαρμοστούν σε αυτά τα επίπεδα θορύβου. Συμπερασματικά, ενώ ο ανθρωπογενής θόρυβος είναι ένα πρόσφατο φαινόμενο στην ιστορία της εξέλιξης, έχει αποδειχτεί ότι έχει τη δυνατότητα να μεταβάλλει τη συμπεριφορά των ζώων, να αλλάξει τη φυσιολογία τους, ακόμη και να αναδιαρθρώσει τις κοινότητές τους.

1.3 Η μελέτη του birdsong στη Γνωσιακή Νευροεπιστήμη και στην Εξελικτική Νευροβιολογία

Παλαιότερες μελέτες (Thorpe, 1958) έδειξαν ότι νεαροί σπίνοι που εκτράφηκαν σε εργαστήρια, απομονωμένοι από ενήλικα αρσενικά πουλιά του ίδιου είδους, παρήγαγαν πολύ αφύσικα τραγούδια. Όταν όμως αυτά τα νεαρά πουλιά υποβλήθηκαν σε ακρόαση ηχογραφημένου birdsong από άγριους σπίνους, κατάφεραν να παραγάγουν birdsong παρεμφερές με αυτό των ηχογραφήσεων. Αποδείχτηκε επιστημονικά, για πρώτη φορά, ότι τα νεαρά πουλιά πρέπει να μαθαίνουν το χαρακτηριστικό τραγούδι του είδους τους ακούγοντας ώριμα, ομοειδή πουλιά. Ο Peter Marler, μαθητής του Thorpe, και οι συνάδελφοί του (Marler & Peters 1987, Marler 1997) επεσήμαναν την ύπαρξη διαφορετικών «διαλέκτων» στα πουλιά, ανά γεωγραφική περιοχή. Υπογράμμισαν επίσης ότι η εκμάθηση του birdsong διαδραματίζεται σε μία πρώιμη, ευαίσθητη περίοδο της ζωής των πουλιών και ότι τα πουλιά έχουν μία έμφυτη προδιάθεση στη μάθηση του birdsong του είδους τους. Ο Nottebohm - μαθητής και αυτός με τον Marler - και οι συνάδελφοί του εντόπισαν νευρωνικά κυκλώματα στον εμπρόσθιο εγκέφαλο των πτηνών που ελέγχουν την τραγουδιστική συμπεριφορά τους, μία σημαντική ανακάλυψη, η οποία άνοιξε το δρόμο για επόμενες έρευνες (Nottebohm et al. 1976, Nottebohm 1987, Nottebohm et al. 1990).

Κατά τους Brenowitz et al. (1997), το σύστημα του birdsong προσφέρει πολλά πλεονεκτήματα ως μοντέλο για τον εντοπισμό νευρικών μηχανισμών που βρίσκονται πίσω από βιολογικά παρεμφερή συμπεριφορά και τα συμπεράσματα που προκύπτουν από την μελέτη του είναι τα εξής:

α) Το birdsong είναι μια μαθησιακή συμπεριφορά η οποία ελέγχεται από διακριτά νευρικά κυκλώματα.

β) Οι Thorpe και Marler απέδειξαν ότι υπάρχουν ξεχωριστές φάσεις στην ανάπτυξη του birdsong, με σαφώς καθορισμένες ευαίσθητες περιόδους. Κατά τη διάρκεια μιας αρχικής φάσης απομνημόνευσης, τα νεαρά πουλιά αποκτούν ένα αισθητηριακό μοντέλο του «τραγουδιού», με βάση το «τραγούδι» των ώριμων αρσενικών. Αυτό το πρώτο στάδιο ανάπτυξης του birdsong αναφέρεται ως «απόκτηση μνήμης» (memory acquisition) ή «αισθητηριακή απόκτηση» (sensory acquisition). Αργότερα, τα νεαρά πουλιά αρχίζουν να μετατρέπουν αυτή την αισθητηριακή μνήμη σε ένα κινητικό μοτίβο παραγωγής «τραγουδιού» στην αισθητηριοκινητική φάση της μάθησης του τραγουδιού. Διακρίνονται τρία στάδια σε αυτήν τη φάση. Αρχικά, τα πουλιά παράγουν «subsong», το οποίο είναι σιγανό, ημιτελές και ποικίλλει σε μορφή. Μέσω εξάσκησης, τα νεαρά πουλιά το εξελίσσουν σε «plastic song», το οποίο είναι πιο δυναμικό και ολοκληρωμένο, αλλά ακόμα ποικίλλει σε μορφή. Στο τέλος, το τραγούδι «κρυσταλλοποιείται» σε δομή, καθώς τα πουλιά παράγουν μια στερεότυπη μορφή του αισθητηριακού μοντέλου στο οποίο είχαν εκτεθεί νωρίτερα.

γ) Το birdsong είναι προϊόν στερεότυπων κινητικών προγραμματισμών, με ιεραρχική οργάνωση των προκινητικών και κινητικών πυρήνων.

δ) Η τραγουδιστική συμπεριφορά και τα σχετιζόμενα νευρωνικά κυκλώματα είναι διαφοροποιούνται στα περισσότερα είδη με βάση το φύλο του πουλιού.

ε) Οι γοναδικές στεροειδείς ορμόνες έχουν έντονη επίδραση στην ανάπτυξη και τη λειτουργία των κυκλωμάτων ελέγχου του birdsong στα ώριμα πουλιά, καθώς και στη τραγουδιστική συμπεριφορά.

στ) Υπάρχει εκτεταμένη πλαστικότητα του συστήματος τραγουδιού στα ώριμα πουλιά.

ζ) Υπάρχουν διαφοροποιήσεις ανάμεσα στα διάφορα είδη σχετικά με την περίοδο εκμάθησης του birdsong, την παραγωγή του birdsong με βάση το φύλο του πουλιού, τον αριθμό «τραγουδιών» που μπορεί να μάθει ένα πουλί, και την τραγουδιστική συμπεριφορά με βάση την εποχή. Συγκεκριμένα ως προς την περίοδο εκμάθησης του birdsong, υπάρχουν είδη που αναφέρονται ως «age limited learners» ή «sensitive period learners», όπου η απομνημόνευση του birdsong περιορίζεται στο πρώτο έτος της ζωής τους και κατά την ωρίμανση τους δεν προστίθενται νέα «τραγούδια». Σε αυτήν την κατηγορία ανήκουν ο Σπίνος Ζέβρα («taeniopygia guttata») και το Λευκόφρυδο Σπουργίτι («zonotrichia leucophrys»). Από την άλλη πλευρά,

υπάρχουν είδη πουλιών τα οποία μπορούν να αναπτύξουν νέα μοτίβα birdsong πέραν από το πρώτο έτος της ζωής τους, τα οποία αναφέρονται ως «open-ended learners». Σε αυτά ανήκουν το Καναρίνι («serinus canaria») και το ευρωπαϊκό Ψαρόνι («sturnus vulgaris»). Η ανάπτυξη νέων «τραγουδιών» από ώριμα πτηνά open-ended learners συμβαίνει κατά μία συγκεκριμένη εποχή και όχι καθ' όλη τη διάρκεια του έτους. Δεν είναι γνωστό εάν η ικανότητα εκμάθησης νέων «τραγουδιών» συνεχίζεται σε ολόκληρη την ενήλικη ζωή των ειδών open-ended learners.

1.4 Η σχέση του birdsong με τη γλώσσα

Για να ερμηνευτεί και να αναδειχθεί μία ενδεχόμενη σχέση του birdsong με την ανθρώπινη γλώσσα, θα πρέπει καταρχάς να αποσαφηνιστούν μερικοί βασικοί όροι της Γλωσσολογίας (Linguistics), δηλαδή της επιστημονικής μελέτης της ανθρώπινης γλώσσας, η οποία δίνει απαντήσεις σε ερωτήματα σχετικά με τη φύση της γλώσσας, τη δομή και τη σχέση της με τον ανθρώπινο νου, αλλά και με την κοινωνία. Η γλωσσολογική ανάλυση γίνεται σε έξι διακριτά γλωσσολογικά επίπεδα: το φωνητικό, το φωνολογικό, το μορφολογικό, το συντακτικό, το σημασιολογικό και το πραγματολογικό.

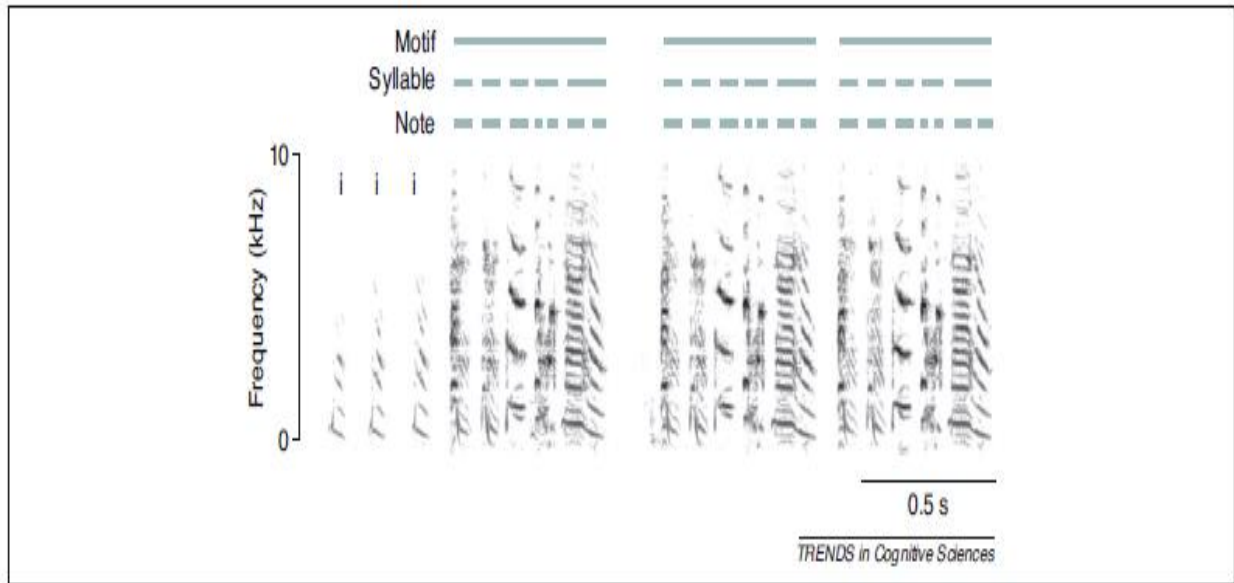
Ως Φωνητική (Phonetics) ορίζεται η μελέτη και η πρόσληψη των γλωσσικών ήχων από αντικειμενική σκοπιά. Συνδέεται με τους ήχους της γλώσσας, τον τρόπο με τον οποίο αρθρώνονται και τον τρόπο με τον οποίο τους αντιλαμβάνεται ο ακροατής. Επιπρόσθετα, χρησιμοποιεί τις ίδιες μεθόδους ανάλυσης και επεξεργασίας του ήχου που χρησιμοποιεί η Ακουστική. Η επιστήμη της Φωνητικής εξετάζει την παραγωγή των ήχων (Αρθρωτική Φωνητική), την πρόσληψή τους (Ακουστική Φωνητική) και την κατανόησή τους (Αντιληπτική Φωνητική). Η Φωνολογία (Phonology) αποτελεί επιστημονικό κλάδο της γλωσσολογίας που ασχολείται με τη μελέτη των ήχων από λειτουργική σκοπιά, δηλαδή με τους ήχους που έχουν διακριτική / διαφοροποιητική λειτουργία για το νόημα του γλωσσικού σήματος. Επίσης, ο όρος «φωνολογία» μπορεί να αναφέρεται στο φωνολογικό (ηχητικό) σύστημα μίας συγκεκριμένης γλώσσας, ένα από τα βασικά συστήματα από τα οποία απαρτίζεται μια γλώσσα, όπως είναι το συντακτικό και το λεξιλόγιο. Η Μορφολογία (Morphology) είναι η μελέτη του σχηματισμού μίας λέξης και της δομής της, και μελετά τον τρόπο κατά τον οποίο συντίθενται οι λέξεις από τα μικρότερα συστατικά τους (τα «μορφήματα»⁵), καθώς και τους κανόνες που ελέγχουν αυτή τη διαδικασία. Η Σύνταξη (Syntax) είναι η μελέτη των κανόνων και των ιεραρχικών σχέσεων που ρυθμίζουν τον τρόπο με τον οποίο συνδυάζονται οι λέξεις για το σχηματισμό προτάσεων στις φυσικές ανθρώπινες γλώσσες. Η Σημασιολογία (Semantics) είναι ο κλάδος της Γλωσσολογίας που εξετάζει τη σημασιολογική δομή μίας γλώσσας, και συνδέεται με την περιγραφή της

⁵ Ως μόρφημα ορίζεται η μικρότερη μονάδα-φορέας σημασίας σε μια δεδομένη γλώσσα.

αναπαράστασης του νοήματος μιας λέξης στον ανθρώπινο νου και με το πώς χρησιμοποιείται αυτή η αναπαράσταση για την κατασκευή προτάσεων. Τέλος, η Πραγματολογία είναι η μελέτη της χρήσης της φυσικής γλώσσας σε πραγματικές συνθήκες ανθρώπινης επικοινωνίας, και πιο συγκεκριμένα, η μελέτη της επίδρασης που έχει το περιβάλλον, είτε είναι γλωσσικό είτε εξωγλωσσικό, στην ερμηνεία μιας πρότασης, όπως αυτή εκφέρεται μέσα σε συγκεκριμένο χώρο και χρόνο.

Όσον αφορά στη σχέση του birdsong με την ανθρώπινη φωνολογία, με δεδομένο ότι το birdsong δεν εμπεριέχει συνδυασμούς ηχητικών στοιχείων μέσα σε μορφήματα, δεν εμπεριέχει μορφήματα μέσα σε λέξεις, και δεν εμπεριέχει λέξεις μέσα σε φράσεις ή προτάσεις, πρέπει να θεωρείται ότι φέρει μόνο «φωνολογική σύνταξη». Αυτό οφείλεται στο γεγονός ότι το birdsong δεν διαθέτει σημασιολογία, δηλαδή τα ηχητικά στοιχεία του «τραγουδιού» δεν συνδυάζονται για να σχηματίσουν νέες έννοιες· αποτελεί ένα σύστημα επικοινωνίας, το οποίο έχει τη δυνατότητα να μεταφέρει μόνο ένα περιορισμένο σύνολο προθέσεων, που αφορούν είτε προσέλκυση συντρόφων είτε εδαφική υπεράσπιση (Berwick et al., 2011). Αντίθετα, η ανθρώπινη σύνταξη συνδέεται στενά με το νοητικό σύστημα, μία διάκριση που δημιουργεί θεμελιώδεις διαφορές μεταξύ του birdsong και της ανθρώπινης φωνολογίας. Τέλος, εν αντιθέσει με το birdsong, οι προτάσεις στην ανθρώπινη γλώσσα είναι δυνητικά άπειρες σε μήκος και δομή και περιορίζονται μόνο από εξωγενείς παράγοντες, όπως η βραχυπρόθεσμη μνήμη ή η χωρητικότητα των πνευμόνων (Berwick et al., 2011).

Στη φωνολογία, τα φωνήματα αποτελούν τα ελάχιστα στοιχεία/μονάδες μίας γλώσσας, που παρέχουν διακριτική/διαφοροποιητική λειτουργία στο φωνητικό επίπεδο, για το νόημα του γλωσσικού ήχου. Αν τα μορφήματα είναι οι μονάδες της πρώτης άρθρωσης της γλώσσας, τα φωνήματα είναι οι μονάδες της δεύτερης άρθρωσης. Δεν είναι φορείς σημασίας αλλά χρησιμεύουν στο να διαφοροποιούν σημασιολογικά τα μορφήματα. Ομοίως και στο birdsong, τα φωνήματα είναι τα ελάχιστα δομικά συστατικά που τοποθετούνται σε σειρά προκειμένου να σχηματιστεί το «τραγούδι». Αυτό γίνεται αντιληπτό με πιο εύγλωττο τρόπο στην Εικόνα 3, όπου απεικονίζεται το φασματογράφημα ενός Σπίνου Ζέβρα, με σημειωμένες διακριτά τις νότες, τις συλλαβές και τα μοτίβα που σχηματίζονται. Το «τραγούδι» ξεκινάει με τρεις εισαγωγικές νότες (οι οποίες σημειώνονται με «i»), εν συνεχεία ακολουθούν 5 συλλαβές που αποτελούνται από 7 νότες, και οι οποίες σχηματίζουν ένα μοτίβο. Το μοτίβο αυτό επαναλαμβάνεται 2 φορές.



Εικόνα 3: Φασματογράφημα ενός Σπίνου Ζέβρα, στο οποίο απεικονίζεται η ιεραρχική δομή του birdsong (Πηγή: Berwick et al., 2011).

2. Το birdsong υπό το πρίσμα της μουσικολογίας

Τη σχέση του birdsong με τη μουσική διερευνά η επιστήμη της Μουσικολογίας μέσα από τους διάφορους κλάδους της. Πιο συγκεκριμένα η Βιομουσικολογία - όρος που πρωτοεμφανίζεται από τον Nils L. Wallin (Wallin, 1991) - ασχολείται με τη μελέτη της μουσικής από βιολογική άποψη, ενώ η Ζωομουσικολογία - όρος που χρησιμοποιείται αρχικά από τον Γάλλο συνθέτη François-Bernard Mâche (Mâche, 1983) - είναι τομέας της Μουσικολογίας και της Ζωολογίας ο οποίος μελετά τις μουσικές πτυχές του παραγόμενου και προσλαμβανόμενου ήχου των ζώων. Τέλος, η Ορνιθομουσικολογία – η οποία αποτελεί κλάδο της Ζωομουσικολογίας - εστιάζει στα μελωδικο-ρυθμικά μοτίβα των πουλιών, μέσα από την παρατήρηση, την ανάλυση και την καταγραφή της μουσικής τους, με βασικό εργαλείο τη χρήση φασματογραφημάτων και μουσικής σημειογραφίας. Πρωτοπόρος του συγκεκριμένου κλάδου υπήρξε ο συνθέτης και ορνιθολόγος Olivier Messiaen, δάσκαλος του F.B. Mâche.

2.1 Αποτελεί το birdsong μουσική;

Πάνω στο ερώτημα εάν τα πουλιά παράγουν ή αναγνωρίζουν μουσική έχουν γραφεί πολλά. Είναι γεγονός ότι ο ήχος που παράγουν έχει υποτυπώδη δομή και οργάνωση, και συγκεκριμένη λειτουργία. Επιπρόσθετα, διαθέτουν από τη φύση τους τον εξοπλισμό για να δημιουργήσουν μουσική (τη Σύριγξ) αλλά και την απαραίτητη ακοή⁶. Ορισμένα είδη πουλιών έχουν τη δυνατότητα ακόμη και να αναγνωρίζουν μουσικά διαστήματα, ανεξαρτήτως από το εάν μπορούν να τα παραγάγουν ή όχι (Araya-Salas, 2012). Ο ήχος περιέχει πολλά ακουστικά χαρακτηριστικά τα οποία όμως δεν είναι όλα κατάλληλα ή απαραίτητα για οποιαδήποτε εργασία. Ενώ η μουσική μελωδία βασίζεται αποκλειστικά στο τονικό ύψος (pitch), η ομιλία το χρησιμοποιεί ως επικουρικό στοιχείο στη λειτουργία της. Συνεπώς, παρόλο που τα «τραγουδία» των πουλιών παρουσιάζουν μουσικότητα, το τονικό ύψος αυτών των «τραγουδιών» δεν αποτελεί δομικό τους στοιχείο.

Συγκριτικές μελέτες έχουν αποδείξει ότι η μουσική εμπλέκει τους ίδιους νευρικούς μηχανισμούς του συναισθήματος ανάμεσα σε πουλιά και ανθρώπους. Ωστόσο, ζώα όπως οι πίθηκοι – οι οποίοι βιολογικά παρουσιάζουν κοινή καταγωγή με τους ανθρώπους - επικοινωνούν κυρίως μέσω χειρονομιών και εκφράσεων του προσώπου, και όχι μέσω ηχητικών σημάτων. Αυτό το γεγονός αποτελεί ένα σημαντικό τεκμήριο της διαφορετικής λειτουργίας της μουσικής ανάμεσα στα πουλιά και στον άνθρωπο. Όπως συμβαίνει και στην ανθρώπινη ομιλία, το τονικό

⁶ Η διάκριση των συχνοτήτων του ακουστικού συστήματος των πουλιών μοιάζει με αυτήν του ανθρώπινου ακουστικού συστήματος, αντίθετα η χρονική διάκριση είναι πολύ βελτιωμένη στα πουλιά (Σπυρίδης, 1996).

ύψος στα πουλιά χρησιμοποιείται για να μεταφέρει πληροφορίες σχετικά με την ταυτότητα του ομιλητή, να προσδώσει έμφαση ή να αποδώσει συναισθηματικό περιεχόμενο. Έχει μάλιστα διαπιστωθεί από τους ηθολόγους ότι το κελάηδισμα των πουλιών μεταδίδει πληροφορίες, που αφορούν διάφορα επίπεδα της κοινωνικής τους οργάνωσης. Μπορούν δηλαδή τα πουλιά, με βάση το κελάηδισμα, να αναγνωρίζονται αναμεταξύ τους είτε σε επίπεδο είδους, είτε σε γεωγραφικό επίπεδο («διαλέκτου») είτε σε επίπεδο γειτονίας (Σπυρίδης, 1996). Εν τέλει, με ανθρωπομορφικούς όρους θα μπορούσε να λεχθεί πως τα πουλιά εάν ήθελαν να οργανώσουν το τραγούδι τους μέσα σε αρμονικά πλαίσια, θα είχαν την δυνατότητα να το κάνουν. Ωστόσο κάτι τέτοιο δεν θα εξυπηρετούσε κανέναν απολύτως φυσικό σκοπό. Πρέπει πάντως να τονιστεί ότι αυτό το ζήτημα είναι ακόμη υπό διερεύνηση από την επιστημονική κοινότητα.

2.2 Βιομουσικολογία

Η Βιομουσικολογία σχετίζεται με διάφορους κλάδους της Μουσικολογίας και της Μουσικής Ψυχολογίας (ή «Ψυχολογίας της Μουσικής»), που στοχεύει στην κατανόηση της μουσικής συμπεριφοράς και εμπειρίας. Οι βασικότεροι από αυτούς τους κλάδους είναι τρεις. Ο κλάδος της Εξελικτικής Μουσικολογίας μελετά τις ρίζες της μουσικής, το «τραγούδι» των ζώων, καθώς και την εξέλιξη του ανθρώπου και της μουσικής από βιολογικής πλευράς. Ο κλάδος της Νευρομουσικολογίας (ή «Γνωσιακής Νευροεπιστήμης της Μουσικής») ασχολείται με τις περιοχές του εγκεφάλου που εμπλέκονται στην επεξεργασία της μουσικής, του νευρικού συστήματος και τις γνωσιακές διαδικασίες της μουσικής επεξεργασίας, καθώς και την οντογένεση των μουσικών ικανοτήτων και μουσικών δεξιοτήτων. Τέλος, ο κλάδος της Συγκριτικής Μουσικολογίας (ή «Εθνομουσικολογίας») ασχολείται με τις λειτουργίες και τις χρήσεις της μουσικής, καθώς και τα παγκόσμια χαρακτηριστικά των μουσικών συστημάτων και της μουσικής συμπεριφοράς (Wallin, 1991).

2.3 Ζωομουσικολογία

Η Ζωομουσικολογία είναι τομέας της Μουσικολογίας και της Ζωολογίας ο οποίος μελετά τις μουσικές πτυχές των ήχων που παράγονται και προσλαμβάνονται από τα ζώα (πέραν από τον άνθρωπο). Ο όρος αυτός χρησιμοποιείται αρχικά από τον François- Bernard Mâche, πατέρα της Ζωομουσικολογίας, κατά τον οποίο το πεδίο αυτό ασχολείται με την ιδέα ότι τα πουλιά έχουν μία δική τους ψυχική αντίληψη της μουσικής, την οποία προβάλλουν στον ήχο τους. Στην Ζωομουσικολογία οι παραγόμενοι ήχοι των ζώων εξετάζονται βάσει δύο κεντρικών κατευθύνσεων: ως προς την ακουστική φύση - μέσω της ανατομίας και της φυσιολογίας, και ως προς την έννοια της επικοινωνίας - μέσω των ηθολογικών μελετών (Mâche, 1983). Χρησιμοποιώντας εργαλεία και γνώσεις μέσα από διαφορετικά επιστημονικά πεδία, οι

ζωομουσικολόγοι διερευνούν το ζήτημα της «παγκόσμιας μουσικής», ενώ παράλληλα αναλύουν και αξιοποιούν τις αισθητικές ποιότητες των ήχων των ζώων. Πέραν του Mâche, αντίστοιχοι ορισμοί του πεδίου της Ζωομουσικολογίας έχουν δοθεί και από άλλους. Ο John Blacking ορίζει τη μουσική των ζώων ως «ήχους οι οποίοι παράγονται από άλλα είδη, τους οποίους μπορούμε να ακούσουμε ως οργανωμένους», ενώ η Kathleen Higgins προσθέτει ότι οι άνθρωποι πρέπει επίσης να είναι σε θέση να τους κατανοήσουν. Ο ζωομουσικολόγος Dario Martinelli περιγράφει το αντικείμενο της ζωομουσικολογίας ως «αισθητική χρήση της ηχητικής επικοινωνίας μεταξύ των ζώων». Ο μουσικολόγος Marcello Sorce Keller αποδίδει μουσικές ιδιότητες στους ήχους των ζώων, ιδιαίτερα στα «τραγούδια» των φαλαινών και των πουλιών, δηλώνοντας ότι μπορούν να βρεθούν τοπικές παραλλαγές που μοιάζουν με πολιτισμικά χαρακτηριστικά της ανθρώπινης μουσικής.

2.3.1 François-Bernard Mâche

Ο συνθέτης François-Bernard Mâche, γεννημένος το 1935 στο Κλερμόν-Φεράν της Γαλλίας, ανέπτυξε μία δική του μεθοδολογία σύνθεσης την οποία εφήρμοσε σε πολλά από τα έργα του. Βάσει αυτής, η ανάλυση και οργάνωση του birdsong - αλλά και άλλων ήχων της φύσης, στηρίζεται στα φασματογραφήματα, ενώ παράλληλα αναδεικνύει την παγκόσμια μουσική μέσα στα αρχετυπικά μουσικά μοντέλα που διέπουν την γλώσσα και τη μουσική (Mache, 1983). Σημαντικό ρόλο στη διαμόρφωση της προσωπικότητάς τους συνετέλεσε το πλούσιο υπόβαθρό του, αφού γεννήθηκε σε μια οικογένεια μουσικών, ενώ ταυτόχρονα υπήρξε μαθητής του Émile Passani και του Olivier Messiaen. Το 1957 απέκτησε δίπλωμα στην ελληνική αρχαιολογία και το 1958 πιστοποιητικό διδασκαλίας (Agrégation de Lettres classiques). Επίσης, διετέλεσε μέλος του Groupe de Recherches Musicales στο Παρίσι την περίοδο 1958-1963 και της Académie des Beaux-Arts από το 2002.

Ο Mâche έχει μία πλούσια εργογραφία, η οποία περιλαμβάνει ηλεκτροακουστικά έργα, ορχηστρικά έργα, χορωδιακά έργα, φωνητικά έργα και έργα για πιάνο. Κάποια σημαντικά έργα του είναι τα *Kassandra for large ensemble and tape* (1979), *Eridan* (1986), *Moires* (1994), *Kengir* (1991) και *Manuel de résurrection* (1998). Δεδομένου ότι το birdsong χρησιμοποιήθηκε ευρέως από τον Messiaen, ο Mâche απέφυγε να χρησιμοποιήσει ηχογραφήσεις πουλιών στο έργο του μέχρι το 1969. Το έργο *Naluan* (1974) διερευνά συστηματικά το μοντέλο των πουλιών, ενώ στο έργο *Sopiana* (1980) χρησιμοποιείται birdsong παράλληλα με φλάουτο και πιάνο. Επίσης, ηχογραφήσεις φυσικών ήχων, συμπεριλαμβανομένων και ήχων πουλιών, χρησιμοποιήθηκαν στο έργο *Le printemps du serpent* (2001).

Σημείο σταθμός στην Ζωομουσικολογία είναι το βιβλίο του Mâche «Musique, mythe, nature, ou les Dauphins d'Arion» (1983), στο οποίο υποστηρίζει μία στροφή της σύνθεσης στη

«μυθική σκέψη» (Georgaki, 2017) και περιλαμβάνει μια μελέτη της «ορνιθομουσικολογίας» που αναδεικνύει ότι το birdsong είναι οργανωμένο σύμφωνα με μια αρχή επανάληψης-μετασχηματισμού. Το βιβλίο περιγράφει επίσης τη διαδικασία της σύλληψης και της σύνθεσης των κομματιών από τον Mâche: «ως σημείο εκκίνησης δανείζεται ένα ηχητικό μοντέλο από την πραγματικότητα, το οποίο υποβάλλει στη συνέχεια σε μια πολύ περίπλοκη διαδικασία που οδηγεί στην αφαίρεση».

2.4 Ορνιθομουσικολογία

Η Ορνιθομουσικολογία αποτελεί κλάδο της Ζωομουσικολογίας ο οποίος εστιάζει στα μελωδικό-ρυθμικά μοτίβα των πουλιών, μέσα από την παρατήρηση, την ανάλυση και την καταγραφή της μουσικής τους. Ευρεία χρήση στην Ορνιθομουσικολογία βρίσκουν τα φασματογραφήματα και η μουσική σημειογραφία. Το «τραγούδι» των ωδικών πτηνών έχει απασχολήσει ιδιαίτερος τους ορνιθολόγους, που ασχολούνται όμως, κατά κύριο λόγο με ερωτήματα που αφορούν στην εξέλιξη και στη βιολογία. Κατά τις προηγούμενες δεκαετίες δεν είχαν ακόμη πραγματοποιηθεί αρκετές μελέτες ως προς τη μουσική δομή του birdsong και τη σχέση του με την μουσική, και οι ορνιθολόγοι δεν απέκλειαν μια ενδεχόμενη αισθητική χρήση του ήχου τουλάχιστον από τα ωδικά πτηνά (βλ. Mache 1983).⁷

2.4.1 Olivier Messiaen

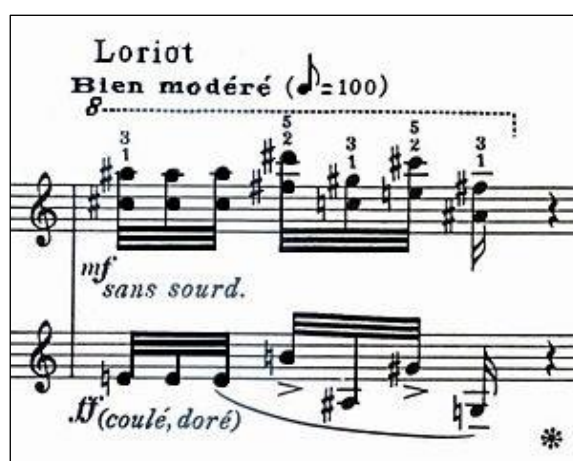
Ο συνθέτης και ορνιθολόγος Olivier Messiaen γεννήθηκε το 1908 στην Αβινιόν της Γαλλίας και υπήρξε ένας από τους πρώτους ορνιθομουσικολόγους. Θεωρείται ένας από τους σημαντικότερους συνθέτες του 20ου αιώνα και ο πλέον γνωστός συνθέτης που ασχολήθηκε με τη μελέτη του birdsong και τη χρησιμοποίησή του στη συνθετική διαδικασία. Τα έργα του «Reveil des Oiseaux» και «Oiseaux Exotiques» περιέχουν πολλές μεταγραφές μελωδικών πολλά προτύπων των πουλιών· ο ίδιος θεωρούσε ότι τα πουλιά «είναι πιθανότατα οι καλύτεροι μουσικοί που κατοικούν στον πλανήτη μας».

Ο Messiaen ακολούθησε μουσικές σπουδές από μικρή ηλικία, στο Ωδείο των Παρισίων, και έλαβε διάφορες διακρίσεις. Το 1930 κέρδισε το πρώτο βραβείο σύνθεσης και το 1931 διορίστηκε μόνιμος οργανίστας της Εκκλησίας της Αγίας Τριάδος στο Παρίσι. Το 1966 ανέλαβε την τάξη σύνθεσης στο Ωδείο των Παρισίων, ενώ κάποιοι φημισμένοι μαθητές του υπήρξαν οι Iannis Xenakis, Pierre Boulez, Karlheinz Stockhausen και François-Bernard Mâche. Περαιτέρω διακρίσεις περιλαμβάνουν την εκλογή στο Institut de France το 1967 και την Académie des

⁷ Σημ. σήμερα οι απόψεις δίστανται σχετικά με αυτό το ζήτημα, με την πλάστιγγα να γέρνει – όπως προαναφέρθηκε – σε καθαρά λειτουργική χρήση. Ωστόσο, υπάρχουν και διαφορετικές απόψεις, όπως π.χ. ότι το birdsong αποτελεί στοιχείο πολιτισμού («culture»), αφού μεταβιβάζεται από γενεά σε γενεά στα πουλιά, όπως δηλαδή συμβαίνει και στον ανθρώπινο πολιτισμό.

beaux-arts το 1968, το Erasmus Prize το 1971, το βραβείο Gold Metal της Royal Philharmonic Society και το βραβείο Ernst von Siemens Music το 1975, το βραβείο Sonning Award 1977, το Βραβείο Wolf Prize in Arts το 1982. Αναζήτησε έμπνευση από διαφορετικές μουσικές παραδόσεις, ενώ σημαντικό στοιχείο στο έργο του αποτέλεσε η σχέση του ήχου με το χρώμα. Στην πολύ μεγάλη εργογραφία του περιλαμβάνονται και σειραϊκά έργα, ενώ παράλληλα στη μουσική του υπάρχουν επιρροές από άλλες μουσικές παραδόσεις πέραν της δυτικής. Τέλος, υπήρξε από τους πρώτους συνθέτες που χρησιμοποίησαν ηλεκτρονικά όργανα (ondes Martenot).

Το birdsong κίνησε το ενδιαφέρον του Messiaen από νεαρή ηλικία και σε αυτό συνέβαλε ο δάσκαλός του Paul Dukas. Κατά την διάρκεια της ζωής του, στα ταξίδια που πραγματοποίησε, έκανε λεπτομερή καταγραφή των ήχων των πουλιών, τους οποίους χρησιμοποίησε μεταγράφοντάς τους στη μουσική του. Επιπρόσθετα, συμπεριέλαβε ήχους πουλιών σε μερικές από τις πρώτες κίολας συνθέσεις του (π.χ. το «*L'abîme d'oiseaux*» από το «*Quatuor pour la fin du temps*»), ενώ μεταγενέστερα έργα του είναι εξ ολοκλήρου δημιουργημένα από ήχους πουλιών. Χαρακτηριστικές περιπτώσεις είναι τα «*Le réveil des oiseaux*», το οποίο είναι μια χορωδία της αυγής για ορχήστρα, και «*Epode*» (από το «*Chronochromie*»), το οποίο είναι γραμμένο για δεκαοκτώ βιολιά, όπου το καθένα από αυτά παίζει διαφορετικό birdsong. Στις παρτιτούρες αυτών των έργων ο Messiaen χρησιμοποιεί μουσική σημειογραφία για να αποδώσει το birdsong (βλ. Εικόνα 1). Οι συνθέσεις αυτές δεν αποτελούν απλές μεταγραφές, αφού ακόμη και έργα με τίτλους αμιγώς εμπνευσμένους από τα πουλιά, όπως τα «*Catalog d'oiseaux*» και «*Fauvette des jardins*», είναι ηχητικά ποιήματα που ανακαλούν το τοπίο, τα χρώματα και την ατμόσφαιρά του (Kraft, 2013).



Εικόνα 1: «*Le Lorient*» («*Catalog d'oiseaux*»)

Το πρωτοποριακό έργο του Messiaen έχει αποτελέσει αντικείμενο συζήτησης κατά το παρελθόν και έχουν υπάρξει αντικρουόμενες απόψεις ως προς αυτό. Ορισμένες από τις

καταγραφές του θεωρούνται προβληματικές από τους ορνιθολόγους διότι το είδος του πτηνού το οποίο «απεικονίζεται» δεν είναι αναγνωρίσιμο ενώ αντίθετα, οι μουσικοί και οι μουσικολόγοι είναι βέβαιοι για την σύνδεση μεταξύ των μουσικών προτύπων του Messiaen και των αντίστοιχων πτηνών διότι δεν μελετούν μόνο τα επιφανειακά χαρακτηριστικά αλλά προχωρούν σε μια βαθύτερη μουσική δομή (Marler & Slabbekoorn, 2004). Ενδιαφέρουσα παρατήρηση για το έργο του Messiaen είναι αυτή του Paul Griffiths, στην οποία αναφέρει ότι «ο Messiaen ήταν ένας περισσότερο συνειδητοποιημένος ορνιθολόγος από οποιονδήποτε προηγούμενο συνθέτη και ένας περισσότερο μουσικός - παρατηρητής του birdsong από οποιονδήποτε προηγούμενο ορνιθολόγο» (Griffiths, 1985).

2.5 Birdsong και μουσική δημιουργία

Το birdsong έχει επηρεάσει τους συνθέτες με διάφορους τρόπους: είτε έχει αποτελέσει πηγή έμπνευσης για αυτούς, είτε έχουν μιμηθεί τον σκοπό του στη σύνθεση μουσικών έργων, είτε έχουν ενσωματωθεί ολόκληρες ηχογραφήσεις πουλιών σε έργα ή , τέλος, έχουν υπάρξει ακόμα και ντουέτα ανθρώπων και πουλιών.

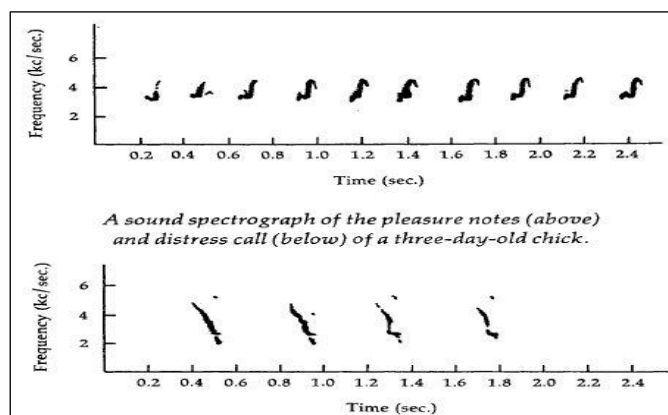
Στην περίπτωση της μίμησης του σκοπού του birdsong, στην κλασική μουσική δύο ιδιαίτερα δημοφιλή πουλιά είναι το αηδόνι και ο κούκος. Το birdsong του αηδονιού έχει χρησιμοποιηθεί από διάφορους συνθέτες όπως ο Mendelssohn, ο Liszt και ο Ravel. Επίσης έχει χρησιμοποιηθεί από τον Handel στην άρια «*Sweet bird*», στο «*Nightingale chorus*» και στο κοντσέρτο για όργανο No 13 «*Ο κούκος και το αηδόνι*», καθώς και από τον Beethoven στην «*Τρίτη Συμφωνία*» του. Το birdsong του κούκου χρησιμοποιείται από τον Beethoven στην «*Ποιμενική Συμφωνία*», από τον Handel στο «*The Cuckoo and the Nightingale*», από τον Rimsky-Korsakov στο «*Snow Maidens*», από τον Vivaldi στο «*Concerto in A, The Cuckoo*» και από τον Gustav Mahler στη «*Πρώτη Συμφωνία*» του. Οι συνθέτες αντλούν έμπνευση από τα πουλιά και το birdsong, τα οποία τα χρησιμοποιούν είτε κυριολεκτικά, μιμούμενοι δηλαδή τους ήχους τους κατά τον ίδιο τρόπο με τον οποίο ο Προκόφιεφ χρησιμοποιεί ένα όμποε για να μιμηθεί τους ήχους μιας πάπιας στο «*ο Πέτρος και ο Λύκος*», είτε συμβολικά (Campbell & Lack, 2011). Μεταξύ των συνθετών του εικοστού αιώνα, πέραν από τον Messiaen, ο Béla Bartók έκανε επίσης εκτεταμένη χρήση του βορειοαμερικανικού birdsong στο «*Κοντσέρτο Νο. 3 για Πιάνο*» ενώ ο Jean Sibelius μιμείται με κλαρινέτα τους ήχους του γερανού στο «*Scene with Cranes*». Όσον αφορά σε άλλες μουσικές παραδόσεις, στη τζαζ ο Paul Winter και ο Jeff Silverbush χρησιμοποίησαν ήχους που θυμίζουν birdsong. Τέλος, χρήση birdsong έχουμε σε πολιτισμούς με πλούσιο φυσικό και ηχητικό περιβάλλον (π.χ. Νέα Γουινέα) ενώ γνωστή είναι η συμβολική χρήση των πουλιών στο ελληνικό δημοτικό τραγούδι.

Όσον αφορά στη χρήση ηχογραφημένου birdsong σε συνθέσεις, ο Ιταλός συνθέτης Ottorino Respighi στο «*The Pines of Rome*» (1923-1924) υπήρξε ο πρώτος που συνέθεσε μουσική κατ' αυτόν τον τρόπο. Το 1972, ο Φινλανδός συνθέτης Einojuhani Rautavaara έγραψε το ορχηστρικό κομμάτι «*Cantus Arcticus*» στο οποίο έκανε εκτεταμένη χρήση birdsong από πουλιά της Αρκτικής. Στις δεκαετίες του 1960 και του 1970, ροκ μπάντες όπως οι Άγγλοι Pink Floyd άρχισαν να χρησιμοποιούν ηχητικά εφέ, συμπεριλαμβανομένου και birdsong στη μουσική τους. Σύγχρονοι συνθέτες όπως ο Mâche και οι R. Murray Schafer, Michel Gonneville, Rozalie Hirs και Stephen Preston έχουν εξίσου χρησιμοποιήσει ηχογραφήσεις πουλιών στα έργα τους.

Υπάρχουν ακόμη παραδείγματα όπου το birdsong χρησιμοποιείται ως η μουσική καθαυτή. Στις 19 Μαΐου 1924, το BBC μετέδωσε μία ραδιοφωνική εκπομπή στην οποία η Beatrice Harrison έπαιξε τσέλο στον κήπο της στο Oxted παράλληλα με το birdsong αηδονιών που προσελκύνονταν από τη μουσική του τσέλου της. Επίσης, ο μουσικός της τζαζ David Rothenberg έκανε το 2002 ένα αυτοσχέδιο ντουέτο με μία Τσίχλα στο Εθνικό Πτηνολογικό Πάρκο του Πίτσμπουργκ.

2.6 Μουσικές δομές στη φωνητική επικοινωνία των πτηνών

Στο «τραγούδι» των πουλιών έχουν παρατηρηθεί πάρα πολλά χαρακτηριστικά στοιχεία που υπάρχουν και στην ανθρώπινη μουσική, όπως αυξομειώσεις στο ρυθμό ή στην ένταση, μελωδικο-ρυθμικά μοτίβα και παραλλαγές, τρίλιες, glissandi κ.ά. Επίσης έχει αποδειχτεί ότι μέσω του διαχωρισμού της Σύριγγας, τα πουλιά μπορούν να δημιουργούν αρμονικούς (Σπυρίδης, 1996), ενώ ενδιαφέρον παρουσιάζουν και οι συνηχήσεις που δημιουργούνται στις χορωδίες της αυγής.



Εικόνα 2: Ήχοι χαράς και θλίψης ενός κλωσσόπουλου (Schaffer, 1977)

Ενδιαφέρουσες αναλογίες μεταξύ ανθρώπινης μουσικής και birdsong, υπάρχουν και ως προς την έκφραση συναισθήματος αφού – σε κάποια είδη - η χαρά εκφράζεται με ανιόντα διαστήματα και η θλίψη με κατιόντα (βλ. Εικόνα 2).

Ο πρώτος τρόπος ηχητικής καταγραφής των φωνημάτων των πουλιών που χρησιμοποιήθηκε είναι η μέθοδος της ονοματοποιίας, η οποία συναντάται αρχικά στο «*Chanson Chants des Oiseaux*» του Clément Janequin το 1528. Ο όρος «ονοματοποιία» χρησιμοποιείται για το χαρακτηρισμό λέξεων, των οποίων το φωνητικό σώμα είναι απομίμηση του πραγματικού ήχου του πράγματος στο οποίο αναφέρεται η λέξη. Σύμφωνα με τον Schaffer, κατ' αυτόν τον τρόπο καταγραφής «τα πρώτα χρόνια οι ορνιθολόγοι κατασκεύαζαν γοητευτικές λέξεις σε μια γλώσσα που δεν υπάρχει για να περιγράψουν τους ήχους των πτηνών» (Schaffer, 1977). Παρά τις ελλείψεις που παρουσιάζει αυτή η μέθοδος, ως κάποιο βαθμό μπορεί να περιγράψει μελωδικά και ρυθμικά στοιχεία. Ακολουθούν κάποια παραδείγματα με τέτοιου είδους καταγραφές, από το βιβλίο του Schaffer (Εικόνα 3).

Hawfinch	<i>Deak . . . waree-ree-ree Tchee . . . tchee . . . tur-wee-wee</i>
Greenfinch	<i>wah-wah-wah-wah-chow-chow-chow-chow-tu-we-we</i>
Crossbill	<i>jibb . . . chip-chip-chip-gee-gee-gee-gee</i>
Great Titmouse	<i>ze-too, ze-too, p'tsee-ée, tsoo-ée, tsoo-ée ching-see, ching-see, deeder-deeder-deeder, biple-be-wit-se-diddle</i>
Pied Flycatcher	<i>Tchéetle, tchéetle, tchéetle diddle-diddle-dée; tzit-tzit-tzit, trui, trui, trui</i>
Mistlethrush	<i>tre-wir-ri-o-ee; tre-wir-ri-o-ee-o; tre-we-o-wee-o-wee-o-wit</i>
Corncrake	<i>crex-crex, krek-krek, rerp-rerp</i>
Common Snipe	<i>tik-tik-tik-tuk-tik-tuk-tik-tuk-chip-it; chick-chuck; yuk-yuk</i>

Εικόνα 3: Ονοματοποιία (Schaffer, 1977)

Ένας άλλος τρόπος ηχητικής καταγραφής των μελωδικών μοτίβων του «τραγουδιού» των πουλιών είναι με τη χρήση μουσικής σημειογραφίας. Από τους συνθέτες του 20ου αιώνα αυτήν την πρακτική ακολούθησε κατά κόρον και καθιέρωσε ο Olivier Messiaen, ωστόσο η πρώτη μεταγραφή κατ' αυτόν τον τρόπο τελείται στο «*Musurgia Universalis*» του – Γερμανού Ιησουίτη ιερέα Athanasius Kircher το 1650 (βλ. Εικόνα 4). Αποτελεί ένα από τα σημαντικότερα έργα της μουσικολογίας και άσκησε μεγάλη επιρροή στην ανάπτυξη της δυτικής μουσικής - ιδιαίτερα στον Bach και στον Beethoven. Επίσης στο βιβλίο αυτό, εκτός των άλλων,

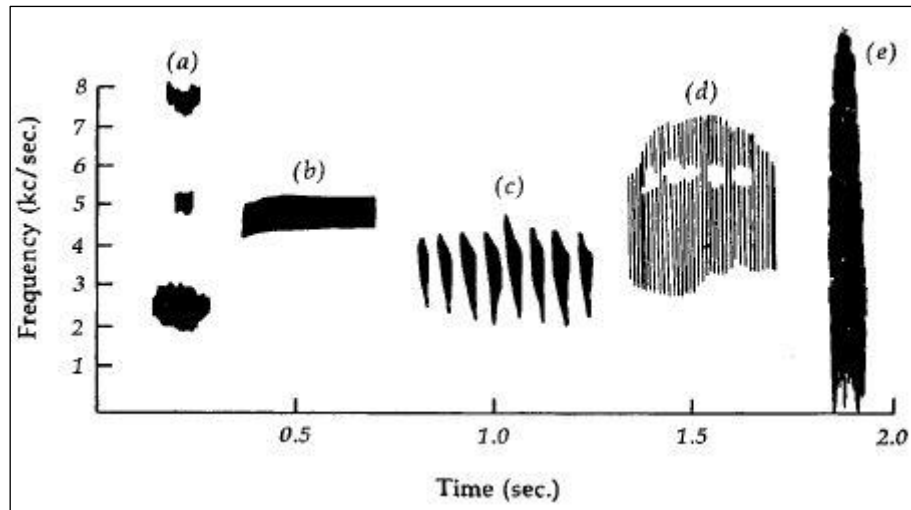
περιγράφεται ένα σύστημα αλγοριθμικής σύνθεσης με ονομασία «Arca Musarithmica», με σκοπό την αυτόματη παραγωγή αντιστικτικών συνθέσεων (Bumgardner, 2009). Στην ευρέως διαδεδομένη εικόνα του Musurgia Universalis με τα πουλιά, απεικονίζεται το «τραγούδι» του κόκορα, της όρνιθας που κλωσάει τα αυγά της, του κούκου, του ορτυκιού και του παπαγάλου. Ο παπαγάλος φέρεται να λέει «Χαίρε» στα ελληνικά, ενώ το «τραγούδι» του κούκου σημειώνεται ως συνήθως σαν ένα κατιόν διάστημα τρίτης.



Εικόνα 4: Musurgia Universalis (1650)

Η χρήση μουσικής σημειογραφίας για την καταγραφή των ήχων των πουλιών αποτέλεσε μεγάλη καινοτομία και διευκόλυνε πολύ τους μουσικολόγους και τους ορνιθολόγους, ωστόσο έχει αρκετούς περιορισμούς. Τα πουλιά κελαηδούν σε πολύ υψηλές συχνότητες, σε γρήγορα tempo, με διαρκείς αυξομειώσεις του τονικού ύψους και των δυναμικών, χρησιμοποιώντας παράλληλα πολύπλοκα ρυθμικά και μελωδικά στοιχεία που είναι δύσκολο να αποδοθούν με τη χρήση μουσικής σημειογραφίας. Σημείο σταθμός στην πλήρη καταγραφή τους αποτέλεσε η εξέλιξη της τεχνολογίας, και πιο συγκεκριμένα η χρήση του φασματογραφήματος. Κατέστη δηλαδή εφικτή η ηχογράφησή τους και εν συνεχεία η επεξεργασία και η λεπτομερής ανάλυση των ήχων τους, σε επίπεδο μικροδομής πλέον. Τέτοιου είδους στοιχεία της μικροδομής τους παρατηρούνται στο φασματογράφημα της Εικόνας 5 (Schaffer, 1977), όπου απεικονίζονται κατά σειρά: το φώνημα ενός αηδονιού με τους αρμονικούς του, το σφύριγμα ενός είδους σπουργιτιού

(white-throated sparrow), η τρίλια μιας βαλτοποταμίδας, ένας μη αρμονικός ήχος ενός είδους σπουργιτιού (clay-colored sparrow) και το κρώξιμο κατά την πτήση ενός είδους παπαγάλου (budgerigar).

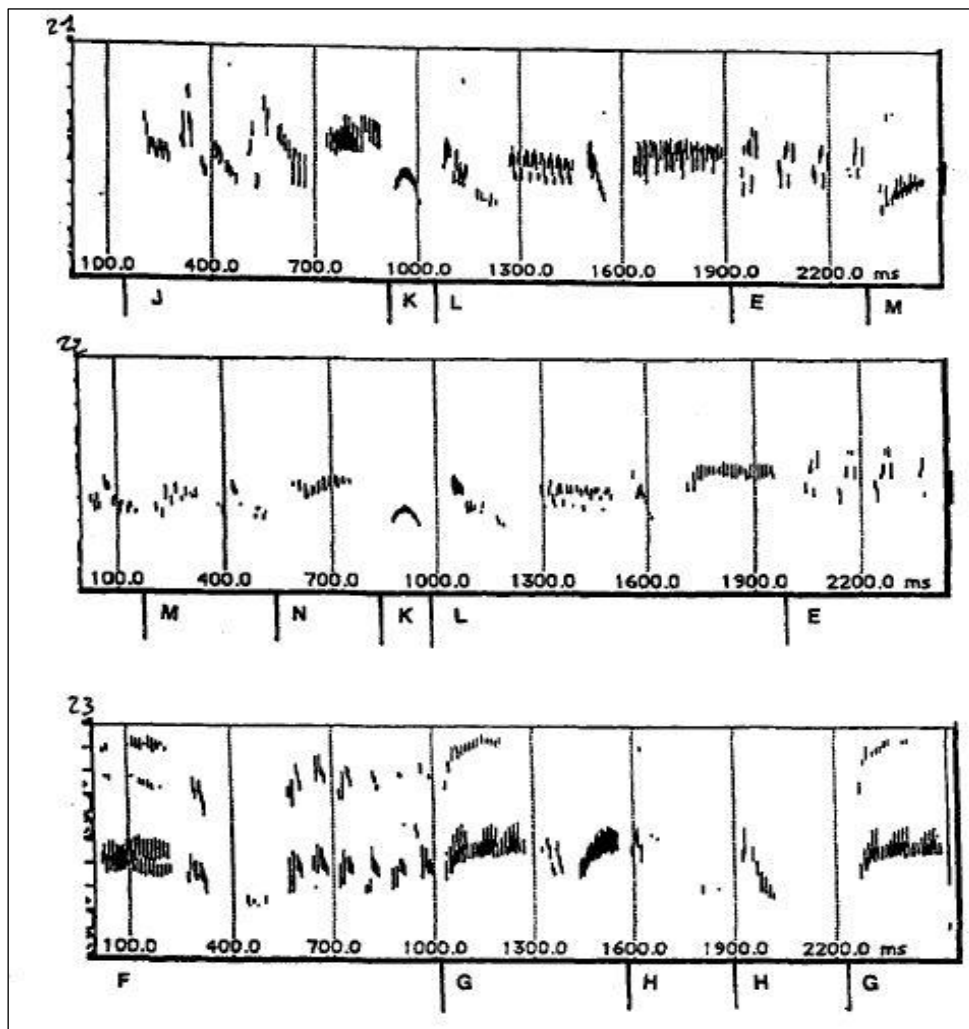


Εικόνα 5: Φασματογράφημα (Schaffer, 1977)

Προκειμένου να αναλυθεί και να μελετηθεί η δομή (syntax) του birdsong με ακρίβεια, ο Mâche ορίζει λεπτομερώς όλα τα δομικά στοιχεία που το απαρτίζουν (Mâche, 1997). Πιο συγκεκριμένα, το φώνημα (phonatom) ορίζεται ως η πιο απλή και αδιαίρετη μονάδα, το μικρότερο συνεχές ηχητικό στοιχείο. Η νότα (note) περιλαμβάνει πολλές μορφές. Μπορεί να οριστεί ως ένα ή περισσότερα (έως και 10) φωνήματα, τα οποία είναι είτε τοποθετημένα πολύ κοντά το ένα στο άλλο στον άξονα του χρόνου είτε είναι παρεμφερή ως προς το ηχόχρωμα. Το staccato είναι μία νότα που αποτελείται από επαναλαμβανόμενα φωνήματα, τα οποία δεν είναι τοποθετημένα πολύ κοντά το ένα στο άλλο στον άξονα του χρόνου, συνεπώς δεν ακούγονται ενιαία. Η τρίλια (trill) είναι ένα staccato που παίζει με δύο διαφορετικά, ασυνεχή φωνήματα. Το vibrato είναι μία τρίλια (trill) με όλα τα εσωτερικά στοιχεία συνδεδεμένα. Η κρατημένη νότα (sustained note) αποτελείται από ένα μεγάλο φώνημα. Το σύνολο (group) είναι μια διαδοχή διαφορετικών φωνημάτων που συνδέονται μεταξύ τους μόνο δια της χρονικής εγγύτητας. Ο κύκλος (cycle) είναι ένα σύνολο που έχει επαναληφθεί τουλάχιστον δύο φορές. Το μοτίβο (motive) είναι ένας συσχετισμός διαφόρων νοτών. Εάν ίδιες νότες επαναλαμβάνονται, τότε αυτό μπορεί να ονομαστεί «iteration». Το σχήμα (figure) είναι ένας συσχετισμός από διάφορα μοτίβα. Αν το ίδιο μοτίβο επαναλαμβάνεται, τότε αυτό μπορεί να ονομαστεί «ostinato». Το θέμα (theme) είναι ένας συσχετισμός από πολλά σχήματα. Εάν το ίδιο σχήμα επαναλαμβάνεται, τότε αυτό μπορεί να ονομαστεί «στροφή» («strophe»). Η ακολουθία (sequence) είναι ένα συνονθύλευμα από νότες, μοτίβα, σχήματα κ.ά. Τέλος, το τραγούδι (song) είναι οτιδήποτε καταγράφηκε από

τους ήχους του πουλιού. Διαχωρίζεται από άλλο τραγούδι με παύσεις που κυμαίνονται από μερικά δευτερόλεπτα έως μεγαλύτερες διάρκειες.

Δύο βασικές αρχές που διέπουν τη μελέτη της δομής είναι αφενός ο προσδιορισμός του ηχογραφημένου υλικού που θα αναλυθεί και αφετέρου ο ορισμός μια συντομογραφίας για κάθε διαφορετικό ήχο ή ομάδα ήχων που επαναλαμβάνεται. Αυτή η πρακτική είναι απαραίτητη καθώς το «τραγούδι» των πουλιών συχνά είναι μεγάλο σε χρονική διάρκεια, με πολλούς επιμέρους ήχους που το απαρτίζουν και πολλές επαναλήψεις. Ο χωρισμός σε «στροφές» διευκολύνει ιδιαίτερα την εξέταση της μικροδομής και της μακροδομής, αφού το τελικό αποτέλεσμα που προκύπτει από αυτού του είδους την ανάλυση είναι σχήματα από συντομογραφίες τα οποία τελικώς συγκρίνονται μεταξύ τους. Στην Εικόνα 6 παρουσιάζεται αυτή η πρακτική.



Εικόνα 6: Ο χωρισμός σε «στροφές» στη μελέτη της δομής (Mâche, 1997)

3. Αυτόματη αναγνώριση του birdsong

Ως αναγνώριση ενός ηχητικού σήματος, αναφέρεται ο προσδιορισμός του είδους της ηχητικής πηγής, της συχνότητας, καθώς και ο χρονικός προσδιορισμός του κάθε επιμέρους ήχου που το απαρτίζουν. Επειδή όμως μια πλήρης ανάλυση όλων των παραπάνω στοιχείων είναι δύσκολη, ο στόχος της αναγνώρισης επαναπροσδιορίζεται ως αναγνώριση ενός συγκεκριμένου μέρους του σήματος. Συγκεκριμένα στην αναγνώριση του birdsong, στον προσδιορισμό δηλαδή του είδους του πουλιού βάσει του ηχογραφημένου καλέσματός του, προκύπτουν διάφορες τεχνικές δυσκολίες, αφού η διαδικασία συλλογής δεδομένων από τα πουλιά είναι ιδιαίτερα απαιτητική και χρονοβόρα υπόθεση. Θόρυβοι, πουλιά που έχουν την δυνατότητα να μιμούνται ήχους άλλων πουλιών, διαφοροποιήσεις ανάμεσα σε υποείδη ή ανά περιοχή είναι συνήθη προβλήματα που δεν μπορούν να αποφευχθούν. Επίσης σημαντική προϋπόθεση είναι η ύπαρξη ειδικού εξοπλισμού εντοπισμού και καταγραφής ήχου, καθώς και εξειδικευμένου ανθρώπινου δυναμικού το οποίο θα είναι σε θέση να αναγνωρίζει τα διάφορα είδη πουλιών.

Ένα μέρος αυτών των προβλημάτων προσπαθεί να επιλύσει η Αυτόματη Αναγνώριση Birdsong, η χρήση υπολογιστικών συστημάτων δηλαδή, τα οποία μπορούν να παρέχουν γρήγορα, ακριβή αποτελέσματα στην αναγνώριση πουλιών. Οι έρευνες σε αυτό το πεδίο σχετίζονται άμεσα με την Αυτόματη Αναγνώριση Ομιλίας αφού η ανθρώπινη ομιλία και το τραγούδι των πουλιών έχουν παρεμφερή, ιεραρχική δομή⁸. Ωστόσο άνθρωποι και πουλιά διαφέρουν πολύ ως προς το συχνοτικό τους εύρος. Παράλληλα, όπως συμβαίνει και στην ανθρώπινη ομιλία, ένα πουλί δεν θα παραγάγει ποτέ το ίδιο ακριβώς κελάηδισμα δεύτερη φορά, γεγονός που καθιστά δύσκολη την αναγνώρισή του από τα υπολογιστικά συστήματα. Το σχετικά νέο αυτό πεδίο λοιπόν ενώ μας παρέχει πολλές δυνατότητες, ταυτόχρονα δημιουργεί και νέα ζητήματα προς επίλυση.

3.1 Αναγνώριση Προτύπων

Μία τυπική διαδικασία αναγνώρισης προτύπων από ένα σύστημα Αυτόματης Αναγνώρισης Birdsong περιλαμβάνει συγκεκριμένα στάδια. Αρχικά, τα ακατέργαστα δεδομένα ψηφιοποιούνται και προεπεξεργάζονται για να απομακρυνθεί ο ανεπιθύμητος θόρυβος. Ιδανική συχνότητα δειγματοληψίας είναι τα 44.100Hz ή τα 48.000Hz. Εάν τα αρχεία ήχου αποτελούνται από πολλά κανάλια, μετατρέπονται σε ένα μονό κανάλι. Στη συνέχεια, αφού τα δεδομένα χωριστούν σε πλαίσια (frames), γίνεται η εξαγωγή των περιοδικών και των συχνοτικών χαρακτηριστικών τους. Τα δεδομένα χωρίζονται σε δεδομένα εκπαίδευσης (training data) και

⁸ Σημ. χρησιμοποιούν δηλαδή συλλαβές για να σχηματίσουν λέξεις, και λέξεις για να σχηματίσουν προτάσεις.

δεδομένα δοκιμής (testing data), και ο ταξινομητής εκπαιδεύεται στα δεδομένα εκπαίδευσης. Στο επόμενο στάδιο τα δεδομένα δοκιμής εισάγονται στον ταξινομητή για να προβεί σε αναγνώριση του είδους του πουλιού. Τέλος, όλα τα αποτελέσματα που θα προκύψουν από την αναγνώριση των δεδομένων δοκιμής συνυπολογίζονται για να εκτιμηθεί το ποσοστό ακρίβειας του ταξινομητή. Είναι επίσης εφικτό ο ταξινομητής να επανεκπαιδευτεί με νέα δεδομένα προκειμένου να βελτιωθεί αυτό το ποσοστό.

3.1.1 Προεπεξεργασία – Αφαίρεση Θορύβου

Ένα πολύ σύνηθες πρόβλημα που υποστηρίζεται από την πλειοψηφία των ερευνητών ότι επηρεάζει σημαντικά την ακρίβεια της αναγνώρισης, είναι η ύπαρξη θορύβου στις ηχογραφήσεις των πουλιών. Αέρας, βροχή, ήχοι άλλων πουλιών και ζώων, και θόρυβοι που προκαλούνται από την ανθρώπινη παρουσία καλύπτουν ένα μεγάλο εύρος συχνοτήτων στις καταγραφές και πρέπει είτε να αφαιρεθούν - στο βαθμό που είναι εφικτό, είτε να μειωθούν. Η επικρατούσα άποψη είναι ότι πρέπει να χρησιμοποιηθούν όσο το δυνατόν «καθαρότερα» από θόρυβο δεδομένα για την εκπαίδευση του ταξινομητή. Για να επιτευχθεί αυτό συνήθως χρησιμοποιούνται διαφόρων ειδών φίλτρα (high-pass, low-pass, band-pass κ.ά.), αφού πρώτα έχει ληφθεί το προφίλ θορύβου της καταγραφής. Βέβαια αυτή η πρακτική έχει περιορισμούς και στην πράξη προκύπτουν δυσκολίες στην εφαρμογή της. Εάν παραδείγματος χάρη το συχνοτικό εύρος του συγκεκριμένου πουλιού συμπίπτει με το συχνοτικό εύρος του θορύβου, τότε το φίλτρο θα «κόψει» απαραίτητες συχνότητες του birdsong. Το ζήτημα αυτό παραμένει ακόμη και σήμερα το σημαντικότερο εμπόδιο που αντιμετωπίζουν οι ερευνητές, ενώ παράλληλα, προτείνονται πολλές διαφορετικές προσεγγίσεις προς επίλυσή του.

3.1.2 Κατάτμηση (Segmentation)

Αλληλένδετα ζητήματα είναι αυτά του θορύβου και της κατάτμησης. Μία ηχογράφιση μπορεί να περιέχει εκτός από το κάλεσμα του πουλιού που μας ενδιαφέρει, και περιττή πληροφορία όπως μεγάλες παύσεις και κενά, με θόρυβο ή σιωπή. Με τον όρο «κατάτμηση» εννοούμε τη διαίρεση ή χωρισμό του ήχου σε τμήματα, επικεντρώνοντας στο είδος της πληροφορίας που μας ενδιαφέρει να ανακτήσουμε. Παραδείγματος χάρη, στο birdsong, μία πρακτική που ακολουθείται είναι ο χωρισμός σε «συλλαβές». Είναι μία διαδικασία που – σε μεγάλο βαθμό, αλλά όχι αποκλειστικά – γίνεται χειροκίνητα⁹. η αυτοματοποιημένη κατάτμηση συνήθως αποφεύγεται αφού είτε δεν είναι ξεκάθαρη η διάκριση των συλλαβών, είτε υπάρχει η πιθανότητα να συμπεριληφθούν τμήματα με θόρυβο. Σημαντικά εργαλεία, βοηθητικά ως προς τη

⁹ Σημ. αυτή η επιλογή ακολουθήθηκε στο στάδιο της προεπεξεργασίας της εφαρμογής birdMelos (βλ. κεφ. 4).

διάκριση των συλλαβών, είναι η κυματομορφή και το φασματογράφημα. Μάλιστα οι de Oliveira et al. (2015) προτείνουν μία μέθοδο ανίχνευσης της ακουστικής δραστηριότητας των πουλιών, με βάση το μορφολογικό φιλτράρισμα του φασματογραφήματος, όταν αυτό επεξεργάζεται ως εικόνα. Παρεμφερείς τεχνικές έχουν χρησιμοποιήσει και άλλοι ερευνητές (π.χ. Potamitis 2014). Όσον αφορά στην αυτοματοποιημένη κατάτμηση, υπάρχουν και εκεί κάποιες χρησιμοποιούμενες τεχνικές, με τα πλεονεκτήματα και τα μειονεκτήματά τους, ενώ στην βιβλιογραφία συναντώνται ακόμη και κάποιες ημι-αυτόματες τεχνικές κατάτμησης.

3.1.3 Τεχνικές Ανάλυσης – Εξαγωγή Χαρακτηριστικών (Feature Extraction)

Η ανάλυση ηχητικών σημάτων (audio signal analysis) είναι μία διαδικασία με σκοπό την εξαγωγή γνώσης αναφορικά με το περιεχόμενο και την φύση των σημάτων αυτών. Οι τεχνικές ανάλυσης εφαρμόζονται αμέσως μετά το στάδιο της προεπεξεργασίας προκειμένου να εξαχθούν τα ηχητικά χαρακτηριστικά (audio features) τα οποία περιγράφουν το αντίστοιχο ηχητικό σήμα, με σκοπό την κατάτμηση, την ταξινόμηση (classification), την αυτόματη ανάκτηση (retrieval) κ.ά. Τα προς επεξεργασία και ανάλυση ηχητικά σήματα είναι στην συντριπτική τους πλειοψηφία σε ψηφιακή μορφή, συνεπώς, στο στάδιο αυτό χρησιμοποιούνται μέθοδοι από την ψηφιακή επεξεργασία σήματος. Συγκεκριμένα χρησιμοποιείται η βραχυπρόθεσμη επεξεργασία (short-term processing), η οποία βασίζεται στην διαίρεση ή παραθύρωση (windowing), δηλαδή την τμηματική εφαρμογή παραθύρων τα οποία κάθε φορά μετακινούνται στον χρόνο και έτσι απομονώνουν ένα ξεχωριστό «πλαίσιο» (frame) του σήματος.¹⁰

Η επιτυχημένη εξαγωγή των κατάλληλων χαρακτηριστικών είναι ο σημαντικότερος παράγοντας στην διαδικασία αναγνώρισης μοτίβων και η ακρίβεια του ταξινομητή που θα εκπαιδευτεί, εν πολλοίς, εξαρτάται από αυτήν. Στην Αυτόματη Αναγνώριση του birdsong έχει διαχρονικά χρησιμοποιηθεί μία μεγάλη ποικιλία τέτοιων χαρακτηριστικών, από πολύ απλά έως σύνθετα, στατιστικά χαρακτηριστικά κ.ά. Αφορούν είτε στην ένταση (ενέργεια) ολόκληρης της καταγραφής ή κάποιων παραθυροποιημένων τμημάτων της, είτε στη συχνότητα (spectrum, cepstrum).

Η Φασματική Ανάλυση είναι η βασικότερη μέθοδος ανάλυσης του ηχητικού σήματος. Το συχνотικό φάσμα είναι η μέτρηση της ενέργειας του ηχητικού σήματος σαν μια συνάρτηση της συχνότητας. Βρίσκει εφαρμογή στην ανάλυση του birdsong γιατί παρέχει μεγάλες δυνατότητες στην παρουσίαση της μικροδομής των φυσικών, φωνητικών και συνθετικών ήχων. Στην ανάλυση και επεξεργασία των ήχων των πουλιών πολύ διαδεδομένη είναι επίσης η

¹⁰ Πρόκειται για μία πολύπλοκη και απαιτητική διαδικασία, αφού οι απόψεις δίστανται σχετικά με το ιδανικό μέγεθος των frames. Στην εφαρμογή birdMelos (βλ. κεφ. 4) χρησιμοποιήθηκαν frames των 50 ms.

φασματογραφική (Sonogram) ανάλυση, η οποία απεικονίζει μία φασματική ανάλυση του αρχείου ήχου στον άξονα χρόνου-συχνότητας, όπου το πλάτος των αρμονικών καθορίζεται από την έντονη γραμμοσκίαση τους. Βέβαια, αναλόγως την περίπτωση, μπορούν να χρησιμοποιηθούν και άλλοι τρόποι φασματικής ανάλυσης. Σε αυτό το σημείο θα πρέπει να αναφερθεί και η ανάλυση κατά Fourier με τη χρήση του αλγορίθμου Fast Fourier Transform (FFT), η οποία είναι μία τεχνική μέτρησης που εφαρμόζεται στους ηλεκτρονικούς υπολογιστές, και η οποία παρέχει τη δυνατότητα να μεταβούμε από το δυναμικό πεδίο ενός ήχου της κυματομορφής του στο φασματικό πεδίο. Ο αλγόριθμος FFT παράγει σε σύντομο χρονικό διάστημα ανάλυση Fourier και χρησιμοποιείται σε ειδικά λογισμικά ακουστικής ανάλυσης και επεξεργασίας ακουστικών ψηφιακών σημάτων.

Τέλος, σημαντική μέθοδο ανάλυσης αποτελεί και η Γραμμική Προγνωστική Κωδικοποίηση¹¹. Πρόκειται για ένα εργαλείο που χρησιμοποιείται κυρίως στην επεξεργασία ηχητικού σήματος και στην επεξεργασία ομιλίας, για την απεικόνιση της φασματικής περιβάλλουσας ενός ψηφιακού σήματος ομιλίας σε συμπιεσμένη μορφή, χρησιμοποιώντας τις πληροφορίες ενός γραμμικού προγνωστικού μοντέλου (Deng & O'Shaughnessy, 2003). Είναι μία από τις πιο ισχυρές τεχνικές ανάλυσης ομιλίας και μία από τις πιο χρήσιμες μεθόδους για την κωδικοποίηση καλής ποιότητας ομιλίας με χαμηλό ρυθμό μετάδοσης bit και παρέχει εξαιρετικά ακριβείς εκτιμήσεις παραμέτρων ομιλίας. Τα χαρακτηριστικά που εξάγονται μέσω αυτής της διαδικασίας, ονομάζονται Linear Prediction Cepstrum Coefficients (συντομογρ.: LPCC) και βρίσκουν ιδιαίτερη εφαρμογή τόσο στην Αυτόματη Αναγνώριση Ομιλίας όσο και στην Αυτόματη Αναγνώριση του Birdsong (π.χ. Lee et al., 2006).

Πέραν από αυτές τις βασικές τεχνικές ανάλυσης και εξαγωγής χαρακτηριστικών υπάρχουν κι άλλες επιλογές που μπορούν να χρησιμοποιηθούν. Η πλέον διαδεδομένη είναι η ανάλυση cepstrum και η εξαγωγή MFCC χαρακτηριστικών¹², που χρησιμοποιείται στην πλειοψηφία των διάφορων ερευνών στο πεδίο της Αυτόματης Αναγνώρισης Ομιλίας και της Αυτόματης Αναγνώρισης του Birdsong. Άλλοι τύποι ανάλυσης είναι ανάλυση wavelets¹³, Short-Time Fourier Transform¹⁴ (STFT) κ.ά.

¹¹ Βλ. Linear Predictive Coding, συντομογρ.: LPC.

¹² Σημ. τα MFCC εξηγούνται στο κεφ. 4.2.

¹³ Η ανάλυση με μετασχηματισμό κυματιδίων χρησιμοποιεί παράθυρα μεταβλητού μεγέθους και απεικόνιση χρόνου – κλίμακας. Συγκεκριμένα ο ευρέως χρησιμοποιούμενος Διακριτός Μετασχηματισμός Κυματιδίων (Discrete Wavelet Transform, συντομογρ.: DWT) είναι κάθε μετασχηματισμός κυματιδίων, τα κυματίδια του οποίου είναι σε διακριτά δείγματα. Το βασικό του πλεονέκτημα σε σχέση με τους μετασχηματισμούς Fourier είναι η ανάλυση στο χρόνο, αφού διατηρεί και την πληροφορία της συχνότητας και την πληροφορία της θέσης

¹⁴ Κατά το μετασχηματισμό Short-Time Fourier το σήμα αναλύεται σε ένα χρονικό παράθυρο κάθε φορά και η ακρίβεια της πληροφορίας εξαρτάται από το μέγεθος αυτού του παραθύρου. Χρησιμοποιεί απεικόνιση χρόνου – συχνότητας.

3.1.4 Τεχνικές Ταξινόμησης

Απαραίτητη προϋπόθεση για να γίνει η αναγνώριση είναι η ύπαρξη ενός ταξινομητή (classifier), τον οποίο θα έχουμε εκπαιδεύσει στο να αναγνωρίζει αντίστοιχα πρότυπα. Η «μηχανική μάθηση», υποπεδίο της επιστήμης των υπολογιστών, ορίζεται από τον Άρθουρ Σάμουελ (1959) ως *«Πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί»*. Διερευνά τη μελέτη και την κατασκευή αλγορίθμων που μπορούν να μαθαίνουν¹⁵ από τα δεδομένα (Kohavi & Provost, 1998) και να κάνουν προβλέψεις με βάση αυτά. Μία συγκεκριμένη προσέγγιση της μηχανικής μάθησης που βρίσκει εφαρμογή στην αναγνώριση ομιλίας είναι η Βαθιά Μάθηση (Deep Learning), η οποία παρέχει τη δυνατότητα στους υπολογιστές να μαθαίνουν από σύνθετα δεδομένα. Παρόλο που η μηχανική μάθηση είναι ένα ευρύ πεδίο που μας παρέχει πολλές επιλογές και δυνατότητες, είναι δύσκολο να γίνει μία ασφαλής, ποιοτική σύγκριση μεταξύ των αλγορίθμων ταξινόμησης, εξαιτίας των διαφορετικών μεθόδων που χρησιμοποιούν αλλά και της διαφορετικής ποιότητας δεδομένων που εισάγονται κάθε φορά.

Τα Τεχνητά Νευρωνικά Δίκτυα¹⁶ είναι υπολογιστικά συστήματα που εκτελούν ορισμένες από τις χαρακτηριστικές λειτουργίες των βιολογικών νευρωνικών δικτύων. Αποτελούνται από τεχνητούς νευρώνες, που συνιστούν την υπολογιστική μηχανή τους, οι οποίοι είναι οργανωμένοι σε επάλληλα επίπεδα (layers). Η συνήθης δομή ενός ΤΝΔ είναι η εξής: αποτελείται από ένα επίπεδο εισόδου - όπου εισάγεται το εκπαιδευτικό διάνυσμα εισόδου, ένα ή περισσότερα κρυφά επίπεδα - όπου γίνεται η επεξεργασία των πληροφοριών και ένα επίπεδο εξόδου, το οποίο διαθέτει υπολογιστική ικανότητα και εξάγει τα αποτελέσματα. Γενικώς, χρησιμοποιούνται είτε για τη μοντελοποίηση σύνθετων σχέσεων μεταξύ δεδομένων εισόδου και εξόδου, είτε για την ανακάλυψη προτύπων στα δεδομένα, ή για τον εντοπισμό στατιστικής δομής σε μία άγνωστη κοινή κατανομή πιθανότητας μεταξύ των παρατηρούμενων μεταβλητών. Στην Αναγνώριση Φωνής, διαδεδομένη μέθοδος εκπαίδευσης ενός ΤΝΔ είναι η εκπαίδευση μέσω Οπισθοδιάδοσης του Σφάλματος (Error Back Propagation).

Σε μία πρώιμη προσπάθεια χρησιμοποίησης ΤΝΔ με τη μέθοδο της οπισθοδιάδοσης του σφάλματος ο ταξινομητής αναγνώρισε σωστά περίπου το 80 - 85% των δειγμάτων (McIlraith &

¹⁵ Υπάρχουν τρεις βασικές κατηγορίες εκμάθησης: η επιτηρούμενη (supervised learning), η μη επιτηρούμενη (unsupervised learning) και η ενισχυτική μάθηση. Στην επιτηρούμενη μάθηση το υπολογιστικό πρόγραμμα δέχεται τις παραδειγματικές εισόδους καθώς και τα επιθυμητά αποτελέσματα από έναν επιβλέποντα, και ο στόχος είναι να μάθει έναν γενικό κανόνα προκειμένου να αντιστοιχίσει τις εισόδους με τα αποτελέσματα. Στην περίπτωση της μη επιτηρούμενης μάθησης χωρίς να παρέχεται κάποια εμπειρία στον αλγόριθμο μάθησης, το υπολογιστικό σύστημα πρέπει να βρει την δομή των δεδομένων εισόδου. Τέλος, στην ενισχυτική μάθηση, ένα πρόγραμμα υπολογιστή αλληλεπιδρά με ένα δυναμικό περιβάλλον στο οποίο πρέπει να επιτευχθεί ένας συγκεκριμένος στόχος, χωρίς κάποιος επιβλέπων να του λέει ρητά εάν έχει φτάσει κοντά στο στόχο του.

¹⁶ Βλ. Artificial Neural Networks, συντομογρ.: ANN.

Card, 1995), ενώ στην πρώτη μεγάλη έρευνα σχετικά με την χρήση ΤΝΔ για μεγάλο σύνολο δεδομένων επιτεύχθηκε συνολική ακρίβεια 86,8% (Cai et al., 2007). Καινοτομία αυτής της μελέτης είναι ότι οι ερευνητές δημιούργησαν έναν δικό τους αλγόριθμο μείωσης θορύβου, και χρησιμοποίησαν γραμμικούς συντελεστές cepstral και συντελεστές Mel-Frequency Cepstral Coefficients (MFCC) για την εξαγωγή χαρακτηριστικών. Κατέληξαν επίσης στο συμπέρασμα ότι η παρεμβολή άλλων πουλιών και ζώων στις ηχογραφήσεις αποτελεί σημαντικό εμπόδιο στη διαδικασία της αναγνώρισης.

Στα ΤΝΔ η μάθηση μπορεί να είναι είτε επιτηρούμενη είτε μη επιτηρούμενη. Ένας αυτορρυθμιζόμενος χάρτης ή αυτορρυθμιζόμενος χάρτης χαρακτηριστικών (συντομογρ.: SOM ή SOFM), είναι ένας τύπος ΤΝΔ που εκπαιδεύεται χρησιμοποιώντας μη επιτηρούμενη μάθηση και διαφέρει από τους άλλους τύπους ΤΝΔ καθώς εφαρμόζει ανταγωνιστική μάθηση αντί της μάθησης μέσω διόρθωσης σφαλμάτων, όπως παραδείγματος χάρι ο αλγόριθμος error back-propagation. Σε μία πρόσφατη έρευνα (Stowell & Plumbley, 2014), η μη επιτηρούμενη μάθηση προτείνεται ως επιτυχημένος τρόπος εκπαίδευσης ως προς την ακρίβεια του ταξινομητή. Διευκρινίζουν όμως οι ερευνητές ότι απαραίτητη προϋπόθεση είναι η ύπαρξη μεγάλου όγκου δεδομένων. Πάντως σε άλλες έρευνες που έχουν διεξαχθεί αναφορικά με την αξιοπιστία της μη επιτηρούμενης μάθησης (Selin et al., 2007), έχουν βρεθεί αντίθετα αποτελέσματα.

Μία άλλη μέθοδος που ανήκει στην κατηγορία της μηχανικής μάθησης είναι οι Μηχανές Διανυσμάτων Υποστήριξης¹⁷, οι οποίες αποτελούν ένα σύνολο μεθόδων επιτηρούμενης μάθησης που χρησιμοποιούνται για την ταξινόμηση και την παλινδρόμηση. Σε αυτή τη μέθοδο, δίνεται ένα σύνολο παραδειγμάτων εκπαίδευσης και κάθε φορά δηλώνεται σε ποια από τις δύο κατηγορίες ανήκει το παράδειγμα. Η ΜΔΥ κατασκευάζει ένα μοντέλο που προβλέπει εάν το νέο παράδειγμα εμπίπτει στην μία κατηγορία ή την άλλη. Αντίθετα με τον αλγόριθμο back-propagation των ΤΝΔ, ο αλγόριθμος support vector είναι γενικότερης φύσης και με ευρύτερη εφαρμοσιμότητα. Επίσης έχει αποδειχτεί ότι οι ΜΔΥ λειτουργούν πολύ καλά με τους συντελεστές MFCC (Fagerlund, 2007). Άλλες σημαντικές έρευνες αναφορικά με τη χρήση ΜΔΥ στην αναγνώριση των καλεσμάτων των πουλιών είναι των Briggs et al. (2012), των Dufour et al. (2013) και τέλος, η έρευνα των Andreassen et al. (2014).

Τα Μεικτά Γκαουσιάν Μοντέλα¹⁸ είναι μια τεχνική που χρησιμοποιεί κατανομή Γκάους για να προσπαθήσει να μοντελοποιήσει διαφορετικές τάξεις. Μια σημαντική έρευνα για τη συγκεκριμένη μέθοδο κατηγοριοποίησης είναι αυτή των Papadopoulos et al. (2015). Ένας άλλος σημαντικός αλγόριθμος κατηγοριοποίησης στη μηχανική μάθηση είναι οι Αλυσίδες Markov, στοχαστικά μοντέλα τα οποία ήδη από την δεκαετία του 70' χρησιμοποιούνται για αναγνώριση

¹⁷ Βλ. Support Vector Machines, συντομογρ.: SVM.

¹⁸ Βλ. Gaussian Mixture Models, συντομογρ.: GMM.

και επεξεργασία εικόνας και ήχου. Τα Κρυφά Μαρκοβιανά Μοντέλα¹⁹ είναι ένα στατιστικό εργαλείο και βρίσκουν ευρεία εφαρμογή στην επεξεργασία σήματος και ιδιαίτερα στην επεξεργασία ομιλίας. Στην αναγνώριση του birdsong παρέχουν στατιστικές προβλέψεις ως προς την διαδοχή των συλλαβών του καλέσματος (Kogan & Margoliash, 1998). Χρήζει ιδιαίτερης αναφοράς το γεγονός ότι σε μία έρευνα (Jančonič & Köküer, 2016) δοκιμάστηκαν ταξινομητές KMM για αναγνώριση πολλαπλών ειδών, κάτι που ακόμη θεωρείται πρόκληση για τους περισσότερους ερευνητές.

Πέραν αυτών των διαδεδομένων τεχνικών, σε μικρότερο βαθμό έχουν δοκιμαστεί και άλλες μέθοδοι. Επιγραμματικά, κάποιες από αυτές είναι: η μέθοδος K-Κοντινότερων Γειτόνων²⁰, διάφοροι τύποι ΤΝΔ όπως Πολυεπίπεδοι Αισθητήρες²¹, Δένδρα Αποφάσεων και Τυχαία Δάση²². Τέλος, υπάρχουν και άλλες μέθοδοι αναγνώρισης που συνδυάζουν τεχνικές ταξινόμησης για να δημιουργήσουν υβριδικά συστήματα. Αυτές οι υβριδικές τεχνικές παρουσιάζουν καλύτερα αποτελέσματα και συνολικά υψηλότερη ακρίβεια σε σύγκριση με τις συμβατικές προσεγγίσεις. Στην αυτοματοποιημένη αναγνώριση ομιλίας η έρευνα γύρω από τέτοια συστήματα είναι ευρέως διαδεδομένη, ωστόσο, μικρότερη εφαρμογή βρίσκει – προς το παρόν - στην αναγνώριση του birdsong.

Κλείνοντας, θα πρέπει να αναφερθεί ότι τα τελευταία χρόνια έχουν πραγματοποιηθεί κάποιες ενδιαφέρουσες έρευνες με σκοπό τη σύγκριση διαφορετικών μοντέλων ταξινόμησης για το ίδιο σύνολο δεδομένων. Σε μία συγκριτική μελέτη (Ross, 2006) χρησιμοποιήθηκαν μοντέλα ΤΝΔ, μοντέλα ΜΔΥ και μοντέλα Εκτίμησης Πυκνότητας Πυρήνα²³. Τα αποτελέσματα έδειξαν ότι το ΤΝΔ πέτυχε το μεγαλύτερο ποσοστό ακρίβειας. Σε μία άλλη έρευνα της ίδιας περιόδου (Somervuo et al., 2006) συγκρίθηκαν μοντέλα KMM, Δυναμικής Χρονικής Στρέβλωσης²⁴ και Μεικτό Μοντέλο Γκάους. Μεγαλύτερη ακρίβεια έναντι των υπολοίπων πέτυχε ο αλγόριθμος ΔΧΣ. Σε μία πιο πρόσφατη έρευνα (Lopes et al. 2011a) δοκιμάστηκαν ο Αφελής ταξινομητής του Bayes²⁵, ένα Δένδρο Αποφάσεων, ένα ΤΝΔ τύπου MLP, και μία ΜΔΥ. Το ΤΝΔ και η ΜΔΥ παρουσίασαν καλύτερη ακρίβεια έναντι των υπολοίπων τεχνικών. Είναι γεγονός ότι ακόμα και

¹⁹ Βλ. Hidden Markov Models, συντομογρ.: HMM.

²⁰ Βλ. K-Nearest Neighbors, συντομογρ.: KNN.

²¹ Βλ. Multilayer Perceptron, συντομογρ.: MLP.

²² Βλ. Decision Trees και Random Forests. Τα Decision Trees εξηγούνται στο κεφ. 4.3.

²³ Όσον αφορά στη μέθοδο Εκτίμησης Πυκνότητας Πυρήνα (Kernel Density Estimation, συντομογρ.: KDE), τα πρότυπα πυκνοτήτων στοχεύουν στο να εντοπίσουν την χωρική συγκέντρωση και ένταση των προτύπων μετακίνησης μέσω του σχηματισμού ενός συνόλου σημείων σε συνεχείς τιμές πυκνότητας μιας ευρύτερης περιοχής. Η μέθοδος πυρήνα (Kernel Density) είναι τεχνική γενίκευσης που μέσω παρεμβολής αποδίδει τιμές σε τμήματα της περιοχής μελέτης.

²⁴ Ο αλγόριθμος Δυναμικής Χρονικής Στρέβλωσης (Dynamic Time Warping, συντομογρ.: DTW), είναι μία τεχνική μέτρησης της ομοιότητας μεταξύ δύο χρονικών ακολουθιών οποιος εφαρμόστηκε αρχικά στην Αυτόματη Αναγνώριση Ομιλίας. Επίσης, βρίσκει εφαρμογή στην Αυτόματη Αναγνώριση Ομιλητή ή ακόμη και στην αναγνώριση βιομετρικών χαρακτηριστικών (υπογραφής) στο διαδίκτυο.

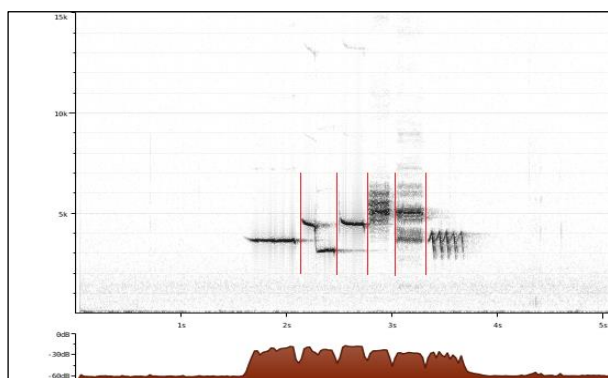
²⁵ Βλ. Naive Bayes classifier.

σε τέτοιου είδους έρευνες, που εξετάζουν διαφορετικούς αλγόριθμους ταξινόμησης πάνω στα ίδια δεδομένα, δεν μπορεί να εξαχθεί ένα ασφαλές συμπέρασμα. Κάποιοι αλγόριθμοι λειτουργούν καλά για συγκεκριμένες εργασίες ενώ ακόμη, οι επιδόσεις τους μπορεί να διαφέρουν αναλόγως των εκάστοτε χαρακτηριστικών που τους έχουν δοθεί.

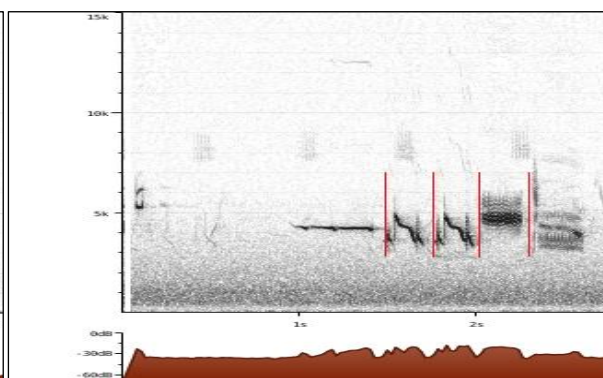
3.2 Συνήθη προβλήματα που προκύπτουν κατά την αναγνώριση

Υπάρχουν διάφοροι παράγοντες που μπορούν να επηρεάσουν το κελάηδισμα: η ηλικία (Williams, 2004) και το φύλο του πουλιού, ή ακόμα εάν το πουλί βρίσκεται σε αστικό ή υπαίθριο περιβάλλον - αφού τα πουλιά της πόλης κελαηδούν σε υψηλότερες συχνότητες λόγω του θορύβου (Luther & Baptista, 2009). Τα προβλήματα αυτά διαπιστώθηκαν και στην πράξη, κατά τη δημιουργία της εφαρμογής birdMelos (βλ. κεφ. 4), στο στάδιο της επιλογής των αρχείων ήχου που θα χρησιμοποιούνταν. Χαρακτηριστικό παράδειγμα είναι η χρησιμοποίηση της καταγραφής με κωδικό καταχώρησης XC188747 (για περισσότερες πληροφορίες βλ. Παράρτημα 1, Πίνακας 1). Πρόκειται για το κάλεσμα ενός νεαρού πουλιού, το οποίο διαπιστώθηκε ότι επιδείκνυε τη συνολική ακρίβεια της αναγνώρισης (βλ. κεφ. 4.4).

Σε παγκόσμια κλίμακα παρουσιάζονται διαφοροποιήσεις στο κελάηδισμα, ακόμα και σε ίδια είδη ή υποείδη πουλιών. Πρόκειται για το ζήτημα των «τοπικών διαλέκτων», όπου ίδια πουλιά παράγουν διαφορετικούς ήχους ανάλογα με τον τόπο τους. Προκειμένου αυτό να γίνει αντιληπτό, με πιο εύγλωττο τρόπο, θα χρησιμοποιηθούν δύο επεξεργασμένα παραδείγματα ήχων πουλιών που ανακτήθηκαν από την πλατφόρμα xeno-canto. Στην εικόνα 1 απεικονίζεται το κελάηδισμα από ένα Ορεινό Λευκόφρυδο Σπουργίτι στον Καναδά (XC250986, Ian Cruickshank) και στην εικόνα 2 απεικονίζεται το κελάηδισμα από ένα Ορεινό Λευκόφρυδο Σπουργίτι στις Η.Π.Α (XC325008, Jim Holmes).



Εικόνα 1: www.xeno-canto.org/250986



Εικόνα 2: www.xeno-canto.org/325008

Είναι προφανές ότι τα δύο παραδείγματα διαφέρουν τόσο ως προς τις «συλλαβές» τους όσο και ως προς το τονικό τους ύψος, ενώ επιμέρους διαφοροποιήσεις παρατηρούνται και στη

μικροδομή του ήχου. Συγκεκριμένα, το παράδειγμα από τον Καναδά αποτελείται από 6 συλλαβές και το τονικό του ύψος είναι κάτω από τα 4000 Hz, ενώ το παράδειγμα από τις Η.Π.Α. αποτελείται από 4 + 1 (που επαναλαμβάνεται) συλλαβές και το τονικό του ύψος είναι πάνω από τα 4000 Hz. Σχετικά με τη μικροδομή τους, υπάρχουν διαφορές στη διάρκεια των συλλαβών και «διαστηματικές» (με μουσικούς όρους), δηλαδή στη συχνοτική απόσταση που χωρίζει τις συλλαβές.

3.3 Τα λογισμικά

Ιδιαίτερο ενδιαφέρουν παρουσιάζουν νέες τεχνικές ανάλυσης για την εξαγωγή χαρακτηριστικών, διαφορετικοί αλγόριθμοι για την ταξινόμηση, καινούργιοι συνδυασμοί τεχνικών για βελτιωμένα αποτελέσματα, που εξελίσσονται διαρκώς τα τελευταία χρόνια. Με βάση τέτοιες έρευνες τα τελευταία χρόνια ορισμένες εταιρείες έχουν αναπτύξει λογισμικά ακουστικής ανάλυσης. Κάποια από αυτά είναι γενικότερης χρήσης, ενώ κάποια άλλα επικεντρώνονται αποκλειστικά στην ακουστική ανάλυση ζώων ή συγκεκριμένα στην αναγνώριση του birdsong.

3.3.1 Λογισμικά ακουστικής ανάλυσης και αναγνώρισης

Ενώ η σύγκριση διαφορετικών μεθόδων ανάλυσης και αλγορίθμων ταξινόμησης έχει ήδη πραγματοποιηθεί κατά το παρελθόν, σε μία πρόσφατη έρευνα (Priyadarshani et al., 2018) γίνεται παρουσίαση και αξιολόγηση του συνόλου των εφαρμογών αναγνώρισης. Σε αυτό το σημείο παρατίθενται τα πλέον διαδεδομένα από αυτά τα λογισμικά:

Το ARBIMON είναι ένα web-based δίκτυο για την αποθήκευση, την κοινή χρήση και την ανάλυση ακουστικής πληροφορίας (Aide et al. 2013). Χρησιμοποιεί χαρακτηριστικά του ήχου όπως: εύρος συχνοτήτων, διάρκεια, μέγιστη ένταση και εύρος ζώνης (bandwidth). Για την αναγνώριση χρησιμοποιεί KMM, ενώ οι δημιουργοί του δεν αναφέρουν τίποτα για την ακρίβεια αυτής.

Το Avisoft-SASLab Pro δημιουργήθηκε από την Avisoft Bioacoustics το 1990, και είναι ένα εμπορικό λογισμικό γενικής χρήσης για την ανάλυση ήχου (Specht, 1993) το οποίο μπορεί να προβαίνει σε μέτρηση παραμέτρων του ήχου. Αφαιρεί τον θόρυβο κάτω από ένα κατώφλι που έχει θέσει ο χρήστης και εξάγει τα εξής χαρακτηριστικά του ήχου: peak frequency, peak amplitude, συχνότητα, εύρος ζώνης και αριθμός αρμονικών.

Το Raven Pro είναι μία εφαρμογή η οποία αναπτύχθηκε από το Cornell Lab of Ornithology (Charif et al., 2010). Παρέχει εργαλεία (φασματογραφήματα) για την ανάλυση και την απεικόνιση του ήχου. Χρησιμοποιεί band-pass φίλτρα, εξάγει συχνοτικά και περιοδικά

χαρακτηριστικά, και για την κατάτμηση (segmentation), χρησιμοποιεί κατώφλι έντασης και συχνότητας.

Το Song Scope της Wildlife Acoustics αρχικά υπήρξε εμπορικό προϊόν – το πρώτο από τα λογισμικά αναγνώρισης που κυκλοφόρησε - και πλέον διατίθεται δωρεάν στην ιστοσελίδα της εταιρείας (Wildlife Acoustics, 2011). Χρησιμοποιεί φίλτρα για την μείωση του θορύβου, αλγόριθμο ανίχνευσης για τον διαχωρισμό των διάφορων καλεσμάτων σε κάθε ηχογράφιση, χαρακτηριστικά MFCC για την ανάλυση (Agranat 2009, Duan et al. 2011) και KMM για την κατηγοριοποίηση. Σε μία έρευνα (Duan et al., 2013) συγκρίθηκε το Song Scope με το λογισμικό Raven Pro ως προς την ακρίβεια τους. Το Song Scope παρουσίασε ακρίβεια 37% έναντι 43% του Raven Pro. Αποδείχτηκε ότι τα δεδομένα με μεγάλο ποσοστό θορύβου μειώνουν αισθητά την ακρίβεια της αναγνώρισής του, αφού ο ταξινομητής KMM αντιμετωπίζει τα τμήματα (segments) με θόρυβο ως συλλαβές του birdsong. Πρόσφατα, η Acoustics Wildlife κυκλοφόρησε το Kaleidoscope, μία εφαρμογή που ειδικεύεται στην αναγνώριση καλεσμάτων νυχτερίδων αλλά και άλλων ήχων, το οποίο θεωρείται ότι υπερέχει του Song Scope.

Το Sound Analysis Pro είναι ένα δωρεάν λογισμικό ανοιχτού κώδικα το οποίο δεν ενδείκνυται για καταγραφές πεδίου, αφού στοχεύει σε ηχογραφήσεις που δεν εμπεριέχουν θόρυβο. Αντίθετα, προτείνεται ως εφαρμογή για να διερευνηθεί η μάθηση και η μίμηση του τραγουδιού στα πουλιά. Εξάγει απλά χαρακτηριστικά του ήχου όπως: Wiener entropy, spectral continuity, τονικό ύψος, frequency modulation, ενώ όσον αφορά την κατάτμηση μπορεί να δεχτεί ολόκληρη την καταγραφή χωρίς να προβεί σε διάκριση συλλαβών (Tchernichovski et al., 2000, Tchernichovski 2012). Η αναγνώριση στηρίζεται σε εκτίμηση της Ευκλείδειας απόστασης.

Το SoundID είναι ένα εμπορικό σύστημα αναγνώρισης ήχου, το οποίο βρίσκει εφαρμογή στο πεδίο της βιοακουστικής. Χρησιμοποιεί band-pass φίλτρο, κατάτμηση με χρήση κατωφλίου ενέργειας, χαρακτηριστικά LPCC και σύγκριση γεωμετρικής απόστασης για την αναγνώριση (Jinnai et al., 2012). Στην ιστοσελίδα του προϊόντος (www.soundid.net) αναφέρεται ότι το συγκεκριμένο λογισμικό παρέχει 95% ακρίβεια, ωστόσο σε κάποιες δοκιμές της εφαρμογής (Priyadarshani et al., 2018) δεν επετεύχθη αυτό το ποσοστό.

Πέραν αυτών των διαδεδομένων εφαρμογών υπάρχουν κι άλλες, εμπορικές και μη, με τα πλεονεκτήματα και τα μειονεκτήματά τους (βλ. Priyadarshani et al., 2018). Κάποιες από αυτές, οι οποίες κυκλοφορούν ακόμα, είναι: BatSound, Ishmael, Luscinia, monitoR (R package), Praat, SonoBat, SpectraPLUS.

3.3.2 Εφαρμογές αναγνώρισης του birdsong αποκλειστικά για έξυπνα κινητά τηλέφωνα (smartphones).

Η ραγδαία εξέλιξη και εξάπλωση των νέων τεχνολογιών, αλλά και η αύξηση του ενδιαφέροντος για ζητήματα ακουστικής οικολογίας, έχει οδηγήσει τα τελευταία χρόνια στην δημιουργία τέτοιου είδους λογισμικών για smartphones. Αυτές οι εφαρμογές είναι πιο απλές στην χρήση, αλλά δεν παρέχουν εξειδικευμένες λειτουργίες ακουστικής ανάλυσης. Ωστόσο η μικρή ποικιλία ειδών, οι δυσκολίες που προκύπτουν από την ύπαρξη ανεπιθύμητου θορύβου και τα χαμηλά ποσοστά ακρίβειας, είναι στοιχεία που παρουσιάζονται και εδώ. Πρόκειται για ένα εντελώς νέο πεδίο που έχει δημιουργηθεί, και σε αυτό το πρώιμο στάδιο ύπαρξής του, η αξιολόγησή του προς το παρόν περιορίζεται στους ερασιτέχνες χρήστες της τεχνολογίας του. Τα πιο διαδεδομένα από αυτά τα λογισμικά είναι τρία: το Bird Song Id, το Song Sleuth και το ChirpOMatic.

Η εφαρμογή Bird Song Id της εταιρείας Isoperla δημιουργήθηκε το 2013 και είναι συμβατή τόσο με iOS όσο και με Android πλατφόρμες. Περιλαμβάνει πουλιά του Ηνωμένου Βασιλείου. Η διαδικασία της αναγνώρισης γίνεται εντός της συσκευής, συνεπώς δεν απαιτείται σύνδεση στο διαδίκτυο, όμως προϋποθέτει δεδομένα καλής ποιότητας για να προβεί σε επιτυχημένη εκτίμηση. Οι δημιουργοί της ισχυρίζονται ότι η ακρίβεια αναγνώρισης είναι της τάξης του 85%.

Η εφαρμογή Song Sleuth δημιουργήθηκε το 2017 από την Wildlife Acoustics σε συνεργασία με τον ορνιθολόγο David Sibley. Είναι συμβατή τόσο με iOS όσο και Android πλατφόρμες. Περιλαμβάνει 189 είδη πουλιών (κυρίως της Βορείου Αμερικής), 3 είδη βατράχων και 3 είδη σκίουρων. Κατά την χρήση, πρέπει να ορίσεις την τοποθεσία σου ώστε η εφαρμογή να αποκλείσει είδη πουλιών που δεν ανήκουν στη συγκεκριμένη περιοχή. Η ηχογράφηση που κάνει παρουσιάζεται σε μορφή φασματογραφήματος. Παρέχεται, επίσης, η δυνατότητα φιλτραρίσματος και επεξεργασίας του ήχου. Τέλος, η λειτουργία της δεν απαιτεί σύνδεση στο διαδίκτυο, αφού όλα τα δεδομένα που απαιτούνται για την ανάλυση υπάρχουν μέσα στην εφαρμογή.

Τέλος, η εφαρμογή ChirpOMatic δημιουργήθηκε το 2017 και προς το παρόν είναι συμβατή αποκλειστικά με iOS πλατφόρμες. Για να προβεί σε αναγνώριση παίρνει ένα δείγμα 12 δευτερολέπτων από την ηχογράφηση, ενώ παράλληλα παρέχει δυνατότητα αποθήκευσης του αρχείου.

3.4 Βάσεις δεδομένων - βιβλιοθήκες ήχων

Η ανάγκη εύρεσης μίας μεγάλης ποικιλίας καταγραφών ενός είδους, αναπόφευκτα θα οδηγήσει σε διαδικτυακές συλλογές από αρχεία ήχων, είτε ερασιτεχνών χρηστών είτε έμπειρων ορνιθολόγων. Αυτές οι ψηφιακές βιβλιοθήκες ήχων πουλιών παρέχουν άμεση και εύκολη πρόσβαση σε μία μεγάλη βάση δεδομένων με είδη από όλον τον κόσμο. Εκτός από τους ήχους καθαυτούς, δίδονται πληροφορίες για το είδος του πουλιού, για την περιοχή και την ώρα που πραγματοποιήθηκε η καταγραφή, για τον εξοπλισμό που χρησιμοποιήθηκε ή ακόμη και το φασματογράφημα της καταγραφής. Καθιστούν έτσι εφικτή την σύγκριση ανάμεσα σε πολλά διαφορετικά είδη και υποείδη, σε διαφορετικές περιοχές του κόσμου. Ωστόσο η χρήση αρχείων ήχου από το διαδίκτυο έχει και μειονεκτήματα, αφού, αφενός η ποιότητα του ήχου πολλές φορές είναι κακή, αφετέρου τυγχάνει κάποιες φορές το πουλί της καταγραφής να μην είναι το αναφερόμενο (δηλαδή να έχει αναγνωριστεί λάθος). Οι πιο διαδεδομένες πλατφόρμες με ήχους πουλιών είναι οι εξής:

Το Avibase (βλ. <https://avibase.bsc-eoc.org>) αποτελεί ένα έργο που βρίσκεται σε εξέλιξη από το 1992 και είναι ένα εκτενές σύστημα πληροφοριών για όλα τα πουλιά του κόσμου, το οποίο περιέχει πάνω από 24 εκατομμύρια αρχεία για περίπου 10.000 είδη και 22.000 υποείδη πτηνών. Τον ιστότοπο τον διαχειρίζεται ο Denis Lepage και φιλοξενείται από την Bird Studies Canada, την καναδική συνιδιοκτήτρια της Birdlife International.

Η βιβλιοθήκη Macaulay του Cornell Lab of Ornithology (βλ. www.macaulaylibrary.org) είναι ένα διαδικτυακό επιστημονικό αρχείο ήχου, βίντεο και φωτογραφίας σχετικά με τη φυσική ιστορία. Εκτός από πτηνά, η βάση δεδομένων περιλαμβάνει αμφίβια, ψάρια και θηλαστικά, και η συλλογή διατηρεί καταγραφές της συμπεριφοράς και της φυσικής ιστορίας κάθε είδους. Παρέχεται πρόσβαση σε αρχεία ήχου και βίντεο της βιβλιοθήκης, για έρευνα, εκπαιδευτική και εμπορική χρήση.

Το Xeno-canto (βλ. www.xeno-canto.org) είναι ένας ιστότοπος για την κοινή χρήση ηχογραφημένων ήχων άγριων πουλιών από όλον τον κόσμο. Ξεκίνησε το 2005 από τους Bob Planqué και Willem-Pier Vellinga και διατηρείται από μια ομάδα διαχειριστών με τη βοήθεια της κοινότητας του xeno-canto. Διοικείται από το ίδρυμα Xeno-canto (ή επίσημα Stichting Xeno-canto voor natuurgeluiden) από την Ολλανδία.

Το Mangoverde World Bird Guide (βλ. www.mangoverde.com/birdsound) είναι ένα εθελοντικό πρόγραμμα για τη δημιουργία ενός ελεύθερα προσβάσιμου οδηγού πολυμέσων για όλα τα είδη πουλιών παγκοσμίως. Ο κύριος στόχος αυτού του ιστότοπου είναι να εγείρει την περιβαλλοντική συνειδητοποίηση του κόσμου. Αυτή τη στιγμή περιλαμβάνει 9917 είδη πτηνών, 2990 ήχους που καλύπτουν 1556 είδη, 23109 φωτογραφίες που καλύπτουν 5641 είδη και 494 βίντεο που καλύπτουν 247 είδη.

4. Η δημιουργία της εφαρμογής birdMelos

Ζητούμενο της παρούσας εργασίας ήταν η δημιουργία μίας πρωτότυπης εφαρμογής αναγνώρισης πουλιών η οποία θα εφαρμόζει τις αρχές και τις μεθόδους της μηχανικής μάθησης. Ο απώτερος στόχος είναι το πρόγραμμα αυτό να έχει τη δυνατότητα μελλοντικά να αναγνωρίζει είδη πουλιών της ελληνικής ορνιθοπανίδας. Προς το παρόν όμως αυτό δεν είναι εφικτό, αφού την προκειμένη χρονική στιγμή δεν υπάρχει κάποια ολοκληρωμένη βιβλιοθήκη ήχων με τα ελληνικά πουλιά, ενώ τα δείγματα ήχων που υπάρχουν στο διαδίκτυο από πουλιά που έχουν καταγραφεί στην Ελλάδα, είναι ελάχιστα και κακής ποιότητας.

Δεδομένων των συνθηκών, αποφασίστηκε να χρησιμοποιηθούν δεδομένα από το διαδίκτυο – χρησιμοποιώντας καταγραφές πουλιών ανεξαρτήτου περιοχής, από όλον τον κόσμο - προκειμένου αφενός να δοκιμαστεί η απόδοση των διαφόρων αλγορίθμων ταξινόμησης ως προς την ακρίβεια αναγνώρισης, αφετέρου να εξεταστούν τα διάφορα χαρακτηριστικά (features) του ήχου ως προς την καταλληλότητά τους. Μάλιστα η χρησιμοποίηση του συνόλου αυτών των συγκεκριμένων χαρακτηριστικών αποτελεί και καινοτομία της συγκεκριμένης εργασίας. Αναπόφευκτα, ζητούμενο ήταν και το κατά πόσο τα δεδομένα του διαδικτύου θα μπορούσαν να ανταποκριθούν επαρκώς σε αυτήν την απαιτητική διαδικασία.

Επιλέχθηκαν τρία διαφορετικά πουλιά, τα οποία εμφανίζονται και στον ελλαδικό χώρο: ο γκιώνης («*otus scops*»), ο κούκος («*cuculus canorus*») και η δεκαοχτούρα («*streptopelia decaocto*»). Τα καλέσματα αυτών των πουλιών αποτελούνται από μία συλλαβή (ένας βαθύς συριστικός ήχος), δύο συλλαβές («κου-κου»), και τρεις συλλαβές («κου-κουου-κου») αντίστοιχα. Το κριτήριο για την επιλογή των συγκεκριμένων πουλιών έγκειται στο γεγονός ότι τα καλέσματά τους έχουν σαφείς, διακριτές δομές.

Σημαντικό κομμάτι της εργασίας υπήρξε η εμβάθυνση στις διάφορες τεχνικές μηχανικής μάθησης και προγραμματισμού, μέσα από την οποία θα μπορούσαν να προκύψουν συμπεράσματα και τεχνογνωσία. Όλος ο κώδικας για την εξαγωγή χαρακτηριστικών και την εκπαίδευση του ταξινομητή γράφηκε στην πλατφόρμα του MATLAB, με χρήση της βιβλιοθήκης MIRtoolbox 1.7.1 (Lartillot & Toivainen, 2007) και της εφαρμογής Classification Learner App. Τέλος, δημιουργήθηκε ένα εύχρηστο GUI interface για το χρήστη μέσω της εφαρμογής MATLAB App Designer. Το birdMelos είναι σε θέση να εξάγει χαρακτηριστικά από κάθε νέο δείγμα – stereo ή mono - και να προβαίνει σε αναγνώριση του είδους, ενώ παράλληλα διαθέτει δικά του φίλτρα αφαίρεσης θορύβου προς βελτίωση της ποιότητας της ακρίβειας του.

4.1 Προεπεξεργασία

Ανακτήθηκαν από την διαδικτυακή πλατφόρμα www.xeno-canto.org συνολικά 213 αρχεία ήχου, 71 δηλαδή για κάθε ένα από τα τρία πουλιά. Το κριτήριο επιλογής αυτών των δεδομένων ήταν αφενός η καλύτερη δυνατή ποιότητα του δείγματος και αφετέρου η συνολική διάρκειά του να μην ξεπερνάει τα 40 sec, για πρακτικούς λόγους. Από τα 71 δείγματα που αντιστοιχούσαν σε κάθε πουλί, τα 50 χρησιμοποιήθηκαν ως δεδομένα εκπαίδευσης του συστήματος και τα 21 χρησιμοποιήθηκαν ως δεδομένα δοκιμής, σύμφωνα με την πάγια πρακτική το 30% των συνολικών δεδομένων να είναι δεδομένα δοκιμής.

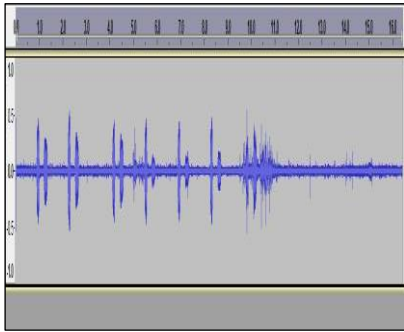
Τα δεδομένα εκπαίδευσης του Γκιώνη αποτελούνταν κατά 30% από calls (15/50), κατά 68% από songs (34/50) και κατά 2% ήταν μικτά (1/50) (βλ. Παράρτημα 1, Εικόνα 1). Τα δεδομένα εκπαίδευσης του Κούκου αποτελούνταν κατά 12% από calls (6/50), κατά 86% από songs (43/50) και κατά 2% ήταν μικτά (1/50) (βλ. Παράρτημα 1, Εικόνα 2). Τα δεδομένα εκπαίδευσης της Δεκαοχτούρας αποτελούνταν κατά 14% από calls (7/50), κατά 78% από songs (39/50) και κατά 8% ήταν μικτά (4/50) (βλ. Παράρτημα 1, Εικόνα 3).

Τα δεδομένα δοκιμής του Γκιώνη αποτελούνταν κατά 19% από calls (4/21), κατά 71% από songs (15/21) και κατά 10% ήταν μικτά (2/21) (βλ. Παράρτημα 2, Εικόνα 1). Τα δεδομένα δοκιμής του Κούκου αποτελούνταν κατά 9% από calls (2/21), κατά 67% από songs (14/21) και κατά 24% ήταν μικτά (5/21) (βλ. Παράρτημα 2, Εικόνα 2). Τα δεδομένα δοκιμής της Δεκαοχτούρας αποτελούνταν κατά 9% από calls (2/21), κατά 81% από songs (17/21) και κατά 10% ήταν μικτά (2/21) (βλ. Παράρτημα 2, Εικόνα 3).

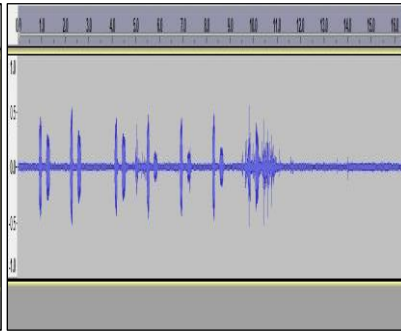
Όσον αφορά στην προ-επεξεργασία των δεδομένων εκπαίδευσης χρησιμοποιήθηκε το λογισμικό Audacity και ακολουθήθηκαν τα εξής βήματα:

- 1) Τα αρχεία μετατράπηκαν από .mp3 σε .wav.
- 2) Τα κανάλια σε κάθε αρχείο από stereo μετατράπηκαν σε mono (βλ. Εικόνα 1).
- 3) Αφού το συχνοτικό εύρος κανενός από τα τρία πουλιά δεν ξεπερνούσε τα 2000 Hz, εφαρμόστηκε ένα low-pass φίλτρο προκειμένου να αφαιρεθούν συχνότητες άλλων πουλιών που ξεπερνούσαν αυτήν τη συχνότητα (βλ. Εικόνα 2).
- 4) Χρησιμοποιήθηκε ένα ακόμα φίλτρο για τις χαμηλές συχνότητες, κάτω από τα 400 Hz, προκειμένου να αφαιρεθεί ο ανεπιθύμητος θόρυβος (βλ. Εικόνα 3).
- 5) Τέλος, αφαιρέθηκαν «κενά» σημεία στην αρχή και στο τέλος των δειγμάτων.

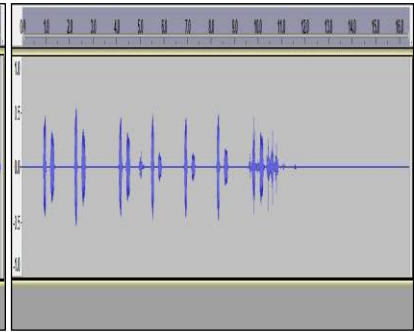
Το αποτέλεσμα αυτής της διαδικασίας απεικονίζεται στις εικόνες 1, 2 και 3, σε επίπεδο κυματομορφής. Το συγκεκριμένο παράδειγμα που χρησιμοποιήθηκε είναι από έναν κούκο, όπου πλέον στην εικόνα 3 οι δύο συλλαβές του είναι «καθαρισμένες» από περιττό περιεχόμενο.



Εικόνα 1

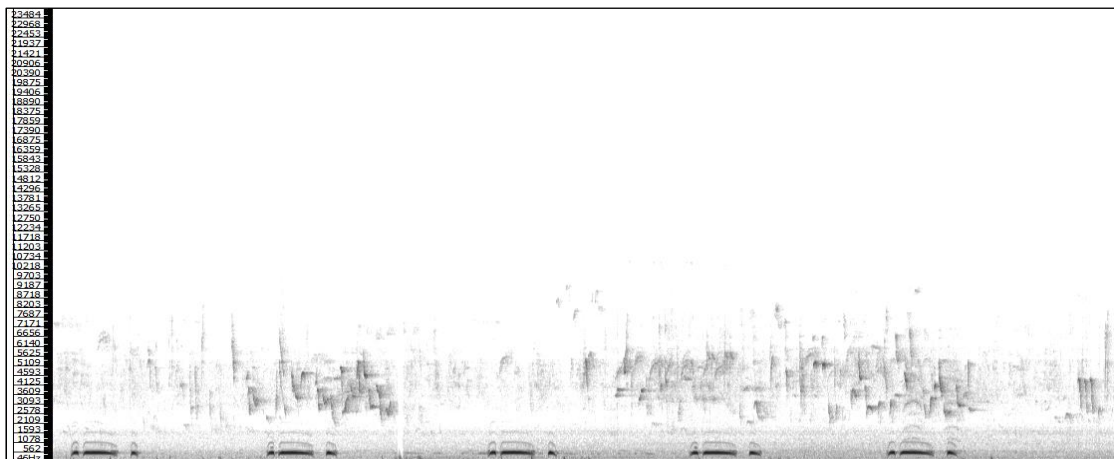


Εικόνα 2



Εικόνα 3

Εφαρμόζοντας την ίδια διαδικασία στο αρχείο ήχου μιας δεκαοχτούρας, προκύπτουν ενδιαφέρουσες παρατηρήσεις σε επίπεδο φάσματος αυτή τη φορά. Στην εικόνα 4 απεικονίζεται το αρχικό φασματογράφημα του αρχείου στο λογισμικό Sonic Visualiser. Στην εικόνα 5 απεικονίζεται το επεξεργασμένο δείγμα:



Εικόνα 4: Φασματογράφημα πριν τη χρήση φίλτρων



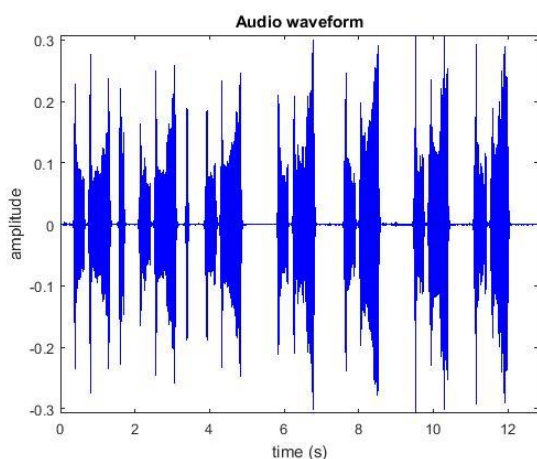
Εικόνα 5: Φασματογράφημα μετά τη χρήση φίλτρων

4.2 Εξαγωγή χαρακτηριστικών

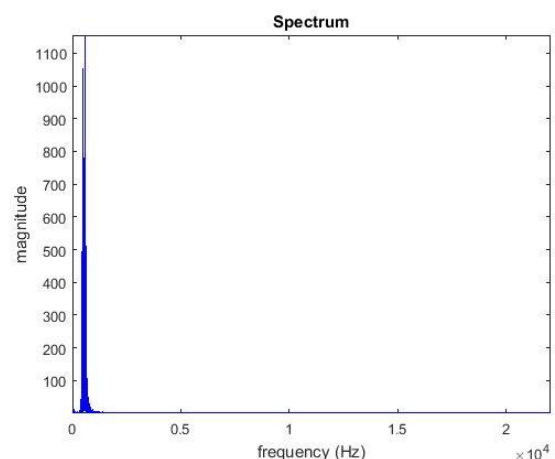
Η ανάλυση του ήχου και η εξαγωγή των συχνοτικών και περιοδικών χαρακτηριστικών του (Feature Extraction) είναι μία σύνθετη και συνήθως χρονοβόρα διαδικασία. Είναι δεδομένο ότι υπάρχει πληθώρα τρόπων και μεθόδων για να πραγματοποιηθεί η παραμετροποίηση του σήματος, η στρατηγική όμως που τελικώς θα επιλέξει κάποιος να ακολουθήσει έγκειται στα μέσα, στην ποιότητα των δεδομένων που διαθέτει, καθώς και στους στόχους που έχει θέσει. Η επιλογή των χαρακτηριστικών παραδείγματος χάρη, μπορεί να γίνει είτε με κριτήριο την προσδοκώμενη ακρίβεια της αναγνώρισης είτε με κριτήριο την προσδοκώμενη ταχύτητα.

Στην εφαρμογή birdMelos τα εξαγόμενα χαρακτηριστικά επιλέχθηκαν με κριτήριο να μην εξαρτώνται από χρονικές πληροφορίες (temporal information) αφού κάτι τέτοιο μπορεί να επηρέαζε την ακρίβεια της αναγνώρισης, δεδομένου ότι στο σύστημα θα εισάγονταν αρχεία διαφορετικής χρονικής διάρκειας. Με βάση την ονομασία της συνάρτησης τους στη βιβλιοθήκη MIRtoolbox τα χαρακτηριστικά αυτά είναι τα εξής: mirrms, mirlowenergy, mirzerocross, mirrolloff, mirbrightness, mircentroid, mirspread, mirskewness, mirflatness, mirentropy, mirmfcc, mirregularity. Επιπρόσθετα, αρχικά χρησιμοποιήθηκαν τρία ακόμα χαρακτηριστικά: mirattacktime, mirattackleap και mirattckslope. Ωστόσο κατά το στάδιο της εκπαίδευσης του ταξινομητή αυτά κρίθηκαν ως ανεπαρκή και εν τέλει απορρίφθηκαν.

Σε αυτό το σημείο θα χρησιμοποιηθεί ως παράδειγμα ένα αρχείο ήχου από τα δεδομένα εκπαίδευσης, το «ecdo.wav», προκειμένου να αναδειχθεί η λειτουργία και η χρησιμότητα των χαρακτηριστικών αυτών. Αρχικώς εισάγεται το αρχείο αυτό στο MATLAB σε μορφή .wav και εν συνεχεία «ανοίγεται» με τη συνάρτηση miraudio της βιβλιοθήκης MIRtoolbox. Στην εικόνα 6 απεικονίζεται η κυματομορφή του, ενώ σε δεύτερο χρόνο, μέσω της συνάρτησης mirspectrum – η οποία κάνει χρήση του αλγορίθμου Fast Fourier Transform - πραγματοποιείται η φασματική απεικόνισή του (βλ. Εικόνα 7).



Εικόνα 6



Εικόνα 7

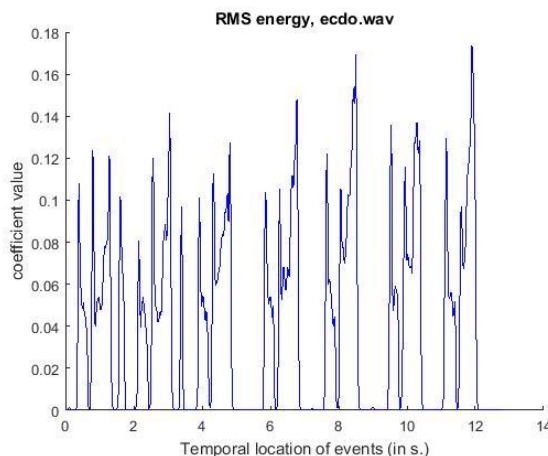
Μία πολύ συνηθισμένη διάκριση των εξαγόμενων παραμέτρων, είναι η διάκριση σε χρονικά χαρακτηριστικά (temporal features), σε φασματικά χαρακτηριστικά (spectral features) και σε σαφματικά χαρακτηριστικά (cepstral features). Το MIRtoolbox ακολουθεί διαφορετική ταξινόμηση: Δυναμικές (Dynamics), Ρυθμός (Rhythm), Ηχόχρωμα (Timbre), Τονικό ύψος (Pitch) και Τονικότητα (Tonality). Παρέχει επίσης μια ποικιλία συναρτήσεων βιβλιοθήκης οι οποίες καλύπτουν στον μέγιστο βαθμό το πεδίο της Music Information Retrieval.

Πέραν των πρώτων δύο χαρακτηριστικών που χρησιμοποιήθηκαν, Root Mean Square Energy και Low Energy Rate, τα οποία σχετίζονται με τις δυναμικές, όλα τα υπόλοιπα αφορούν στο ηχόχρωμα. Αρχικά λοιπόν χρησιμοποιήθηκαν δεκαπέντε χαρακτηριστικά, ενώ στην συνέχεια αφαιρέθηκαν τρία. Πρόκειται για τη στήλη με τα είκοσι επτά Features στους πίνακες του Παραρτήματος 3 στην οποία συμπεριλαμβάνονται και τα τρία που αφαιρέθηκαν. Είναι χρήσιμο να διευκρινιστεί ότι η συνάρτηση `mirmfcc` κάνει εξαγωγή δεκατριών διαφορετικών τιμών και όχι μίας. Αναλυτικά, και με τη σειρά που είναι ταξινομημένα στη βάση δεδομένων, αυτά είναι:

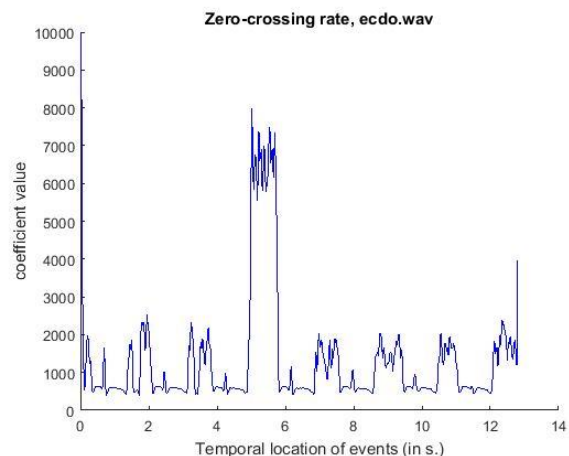
- Feature 1 - Ρίζα του Μέσου Τετραγώνου της Ενέργειας (Root Mean Square Energy):

Η συνάρτηση `mirrms` (βλ. Εικόνα 8) υπολογίζει την ενέργεια του σήματος χρησιμοποιώντας την RMS Energy, δηλαδή τη μέση τετραγωνική ρίζα του τετραγώνου του πλάτους. Ο τύπος με τον οποίο υπολογίζεται είναι:

$$x_{\text{rms}} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}}$$



Εικόνα 8



Εικόνα 9

- Feature 2 – Ποσοστό Χαμηλής Ενέργειας (Low Energy Rate):

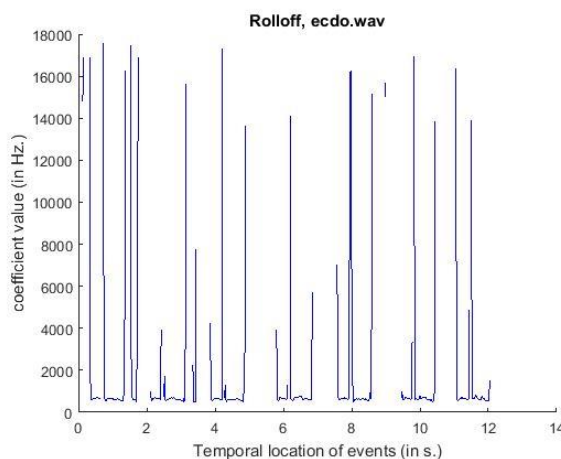
Το function `mirlowerenergy` μας παρέχει πληροφορίες για την κατανομή της ενέργειας στον άξονα του χρόνου. Συγκεκριμένα υπολογίζει το Low Energy Rate, το ποσοστό δηλαδή των πλαισίων - γνωστών και ως «frames» - που παρουσιάζουν ενέργεια χαμηλότερη του μέσου όρου της ενέργειας του σήματος. Η προεπιλεγμένη χρονική διάρκεια των frames είναι 50 ms.

- Feature 3 – Ρυθμός Διέλευσης από το Μηδέν (Zero-Crossing Rate):

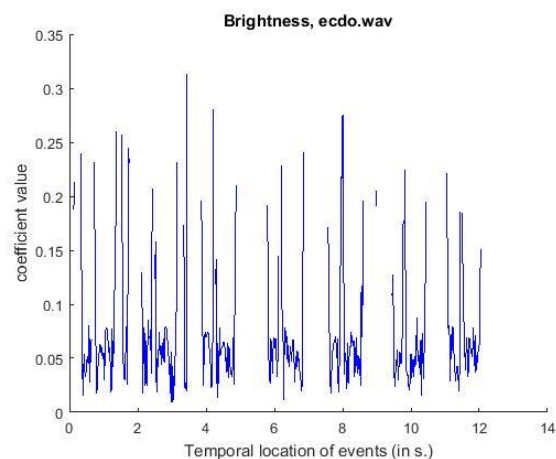
Η συνάρτηση `mirzerocross` υπολογίζει το πόσο θορυβώδες είναι ένα σήμα, μετρώντας τις φορές που το σήμα αυτό αλλάζει πρόσημο διασχίζοντας τον άξονα X (βλ. Εικόνα 9).

$$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} 1_{\mathbb{R}_{<0}}(s_t s_{t-1})$$

Οι περιοδικοί ήχοι συνήθως έχουν χαμηλότερη τιμή αυτού του χαρακτηριστικού ενώ οι θορυβώδεις ήχοι, μεγαλύτερη. Συνεπώς βρίσκει εφαρμογή τόσο στην εκτίμηση του συνολικού θορύβου που εμπεριέχει ένα σήμα, όσο και στην εκτίμηση της θεμελιώδους συχνότητας. Παράλληλα, το Zero-Crossing Rate, χρησιμοποιείται ιδιαίτερα σε περιπτώσεις διάκρισης ομιλίας / μουσικής (Panagiotakis & Tziritas, 2005) αλλά και σε άλλους τομείς που σχετίζονται με ανάλυση του ήχου και κατηγοριοποίηση.



Εικόνα 10



Εικόνα 11

- Feature 4 - Φασματική Απόκλιση (Rolloff):

Η συνάρτηση `mirrolloff` μετρά τις υψηλές συχνότητες του σήματος (High - Frequency Energy) βρίσκοντας μία συχνότητα κάτω από την οποία υπάρχει ένα συγκεκριμένο ποσοστό της συνολικής ενέργειας (βλ. Εικόνα 10). Το προεπιλεγμένο ποσοστό της συνάρτησης είναι το 85%

το οποίο έχει χρησιμοποιηθεί επιτυχώς κατά το παρελθόν (Tzanetakis & Cook, 2002), ωστόσο σε άλλες έρευνες έχει προταθεί και το 95% (Pohle et al., 2005).

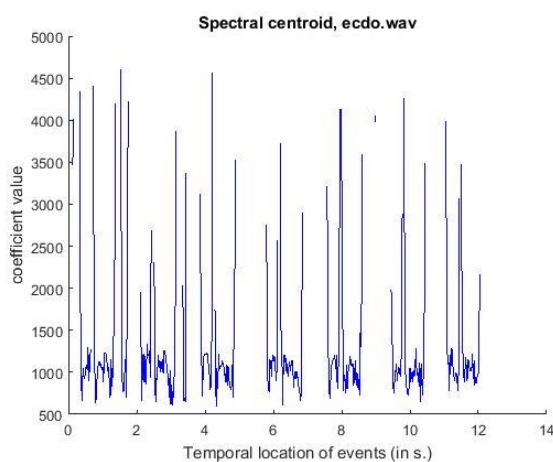
- Feature 5 - Φασματική Κατανομή των Συχνοτήτων (Brightness):

Η συνάρτηση *mirbrightness* αποτελεί εναλλακτική επιλογή για τον υπολογισμό της High - Frequency Energy, θέτοντας μία συχνότητα και υπολογίζοντας το ποσοστό ενέργειας πάνω από αυτήν (Juslin, 2000). Μετρά δηλαδή το κατά πόσο επικρατούν οι ψηλές ή οι χαμηλές συχνότητες στο σήμα. Ορίζεται ως η συνολική ενέργεια τους σήματος που βρίσκεται πάνω από ένα Κατώφλι. Το αποτέλεσμα του υπολογισμού είναι ένας αριθμός ανάμεσα στο 0 και στο 1 (βλ. Εικόνα 11).

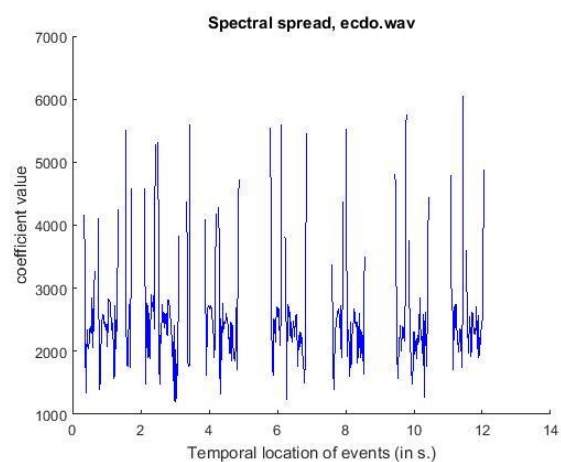
- Feature 6 - Φασματικό Κεντροειδές (Spectral Centroid):

Στατιστική περιγραφή της φασματικής κατανομής. Η συνάρτηση *mircentroid* υπολογίζει το «κέντρο βάρους» του φάσματος (βλ. Εικόνα 12). Αποτελεί δηλαδή ένα μέτρο της φασματικής θέσης, όπου οι υψηλότερες τιμές του χαρακτηριστικού αυτού αντιστοιχούν σε εντονότερους ήχους. Έχει υψηλές διακυμάνσεις για σήματα ομιλίας, και για τον λόγο αυτό έχει χρησιμοποιηθεί σε μεθόδους διαχωρισμού ομιλίας - μουσικής (Giannakopoulos, 2009).

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



Εικόνα 12



Εικόνα 13

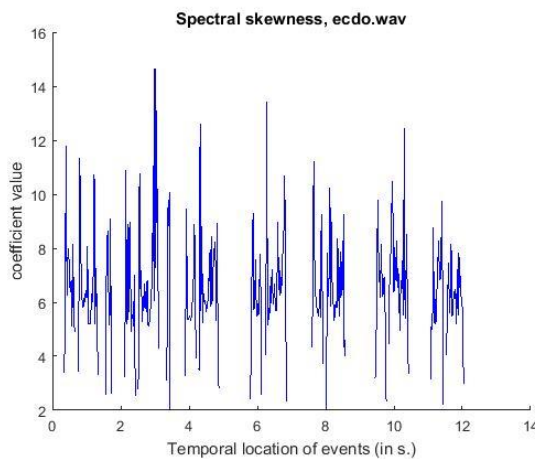
- Feature 7 - Τυπική Απόκλιση (Spectral Spread):

Στατιστική περιγραφή της φασματικής κατανομής. Η συνάρτηση `mirspread` υπολογίζει την «τυπική απόκλιση» (standard deviation) του φάσματος από το φασματικό κέντρο βάρους (βλ. Εικόνα 13). Είναι ένα μέτρο που χρησιμοποιείται στην στατιστική για να υπολογιστεί το ποσό της μεταβολής ή της διασποράς ενός συνόλου τιμών δεδομένων. Η τυπική απόκλιση μιας τυχαίας μεταβλητής είναι η τετραγωνική ρίζα της διακύμανσης (διασποράς) της.

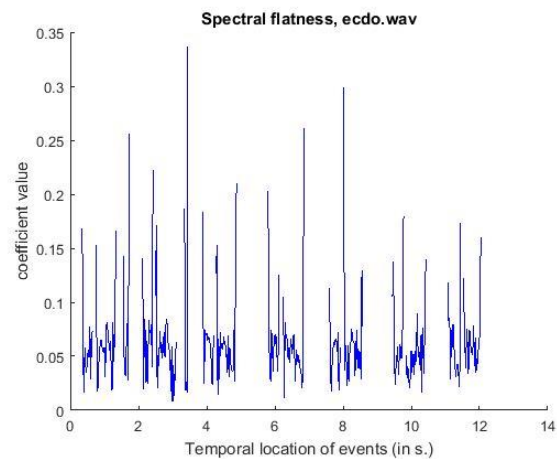
- Feature 8 – Φασματική Ασυμμετρία (Spectral Skewness):

Στατιστική περιγραφή της φασματικής κατανομής. Η συνάρτηση `mirskewness` αποτελεί ένα μέτρο υπολογισμού της συμμετρίας μίας κατανομής ενός συνόλου δεδομένων (βλ. Εικόνα 14). Μπορεί να λάβει είτε θετικό πρόσημο – άρα θα έχει κάποιες τιμές πάνω από το μέσο και η κατανομή θα εμφανίζει «ουρά» προς τα δεξιά, είτε αρνητικό πρόσημο – άρα θα έχει κάποιες τιμές κάτω από το μέσο και η κατανομή θα εμφανίζει «ουρά» προς τα αριστερά. Ο συντελεστής (coefficient) της ασυμμετρίας ισούται με την αναλογία της συμμετρίας προς την τυπική απόκλιση, υψωμένες στην τρίτη.

$$\frac{\mu_3}{\sigma^3}$$



Εικόνα 14



Εικόνα 15

- Feature 9 – Φασματική Επιπεδότητα (Spectral Flatness):

Στατιστική περιγραφή της φασματικής κατανομής. Η συνάρτηση `mirflatness` υπολογίζει το ποσοστό της ομαλότητας στην κατανομή των δεδομένων (βλ. Εικόνα 15).

$$\text{Flatness} = \frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\frac{\sum_{n=0}^{N-1} x(n)}{N}} = \frac{\exp\left(\frac{1}{N} \sum_{n=0}^{N-1} \ln x(n)\right)}{\frac{1}{N} \sum_{n=0}^{N-1} x(n)}$$

Ο υπολογισμός του χαρακτηριστικού αυτού γίνεται με διαίρεση του γεωμετρικού μέσου του φάσματος ισχύος με τον αριθμητικό μέσο, όπου $x(n)$ είναι το πλάτος της n -οστής φασματικής συνιστώσας. Οι υψηλές τιμές αυτού του χαρακτηριστικού μεταφράζονται σε ομοιόμορφη κατανομή της φασματικής ενέργειας σε όλες τις φασματικές ζώνες του φάσματος ισχύος.

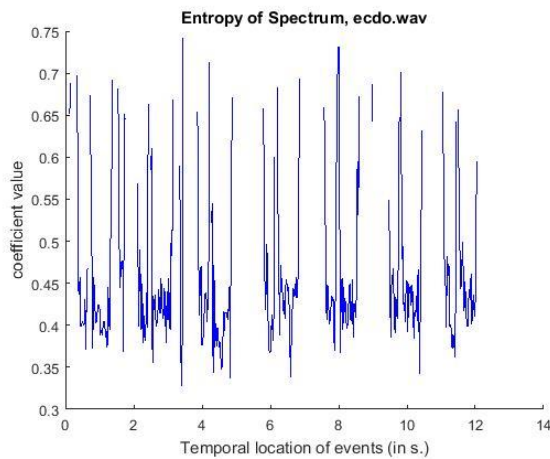
- Feature 10 – Φασματική Εντροπία (Entropy of Spectrum):

Στατιστική περιγραφή της φασματικής κατανομής. Το function `mirentropy` επιστρέφει την σχετική Εντροπία Shannon (βλ. Εικόνα 16). Η Εντροπία Shannon (1948) στη θεωρία πληροφορίας είναι ένα «μέτρο αβεβαιότητας» το οποίο διακατέχει ένα σύστημα, και βασίζεται στην ακόλουθη εξίσωση:

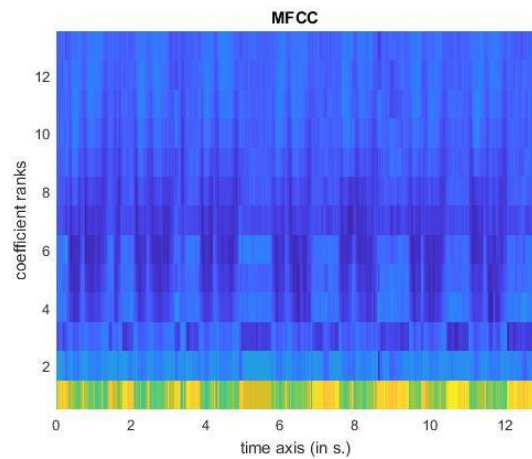
$$H(X) := - \sum_{i=1}^n p(x_i) \log_b p(x_i)$$

όπου b είναι η βάση του λογάριθμου. Προκειμένου η μέτρηση της εντροπίας να είναι ανεξάρτητη από το μήκος της ακολουθίας, χρησιμοποιείται η σχετική εντροπία (Lartillot, 2018) η οποία υπολογίζεται ως εξής:

$$H(p) = -\text{sum}(p.*\log(p)) / \log(\text{length}(p)) ;$$



Εικόνα 16

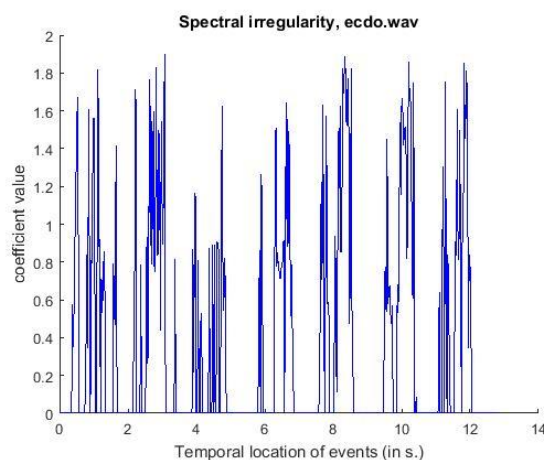


Εικόνα 17

- Features 11 έως 23 - Σαφματικοί Συντελεστές της Κλίμακας Mel (MFCC):

Με τη συνάρτηση `mirmfcc` εξάγονται οι συντελεστές Mel Frequency Cepstral Coefficients ή MFCC (βλ. Εικόνα 17), τα πλέον διαδεδομένα και πολυχρησιμοποιημένα χαρακτηριστικά στην αναγνώριση ανθρώπινης ομιλίας (Makhoul & Schwartz, 1995). Τις τελευταίες δεκαετίες εφαρμόζονται ευρέως στην αναγνώριση των φωνημάτων των πουλιών (Stattner et al., 2013), αλλά και άλλων ζώων (Clemins & Johnson, 2003). Προσφέρουν μια περιγραφή της φασματικής μορφής του ήχου και η διαδικασία εξαγωγής τους μοιάζει αρκετά με την διαδικασία υπολογισμού του `cepstrum` («σάφμα»), όπου ένας μετασχηματισμός Fourier εφαρμόζεται πάνω σε έναν μετασχηματισμό Fourier («Αντίστροφος Fourier»).

Συγκεκριμένα στην περίπτωση των MFCC οι συχνοτικές μπάντες τοποθετούνται λογαριθμικά στην κλίμακα Mel, η οποία προσεγγίζει την απόκριση του ανθρώπινου ακουστικού συστήματος περισσότερο από ότι οι γραμμικά διαχωρισμένες συχνοτικές μπάντες. Εφαρμόζονται τριγωνικά φίλτρα των οποίων οι κεντρικές συχνότητες ισαπέχουν στις χαμηλότερες συχνότητες και τα οποία είναι λογαριθμικά τοποθετημένα στις υψηλότερες συχνότητες. Αντίθετα με το `cepstrum`, εδώ αντί ενός δεύτερου μετασχηματισμού Fourier εφαρμόζεται ένας Διακριτός Μετασχηματισμός Συνημιτόνου²⁶, μετασχηματισμός παρεμφερής με τον Διακριτό Μετασχηματισμό Fourier ο οποίος όμως χρησιμοποιεί μόνο πραγματικούς αριθμούς. Ο ΔΜΣ έχει μια αυξημένη ικανότητα «ενεργειακής συμπίκνωσης» αφού η περισσότερη πληροφορία του σήματος συνήθως τείνει να συγκεντρώνεται σε κάποια λίγα στοιχεία χαμηλής συχνότητας του ΔΜΣ (Lartillot, 2018). Ο προεπιλεγμένος αριθμός στο MIRtoolbox είναι 13 εξαγωγή στοιχεία - χαρακτηριστικά MFCC, συνεπώς η συνάρτηση `mirmfcc` μας δίνει 13 τιμές οι οποίες καλύπτουν τις θέσεις 11 έως 23 των Features των πινάκων του Παραρτήματος 3.



Εικόνα 18

²⁶ Βλ. Discrete Cosine Transform, συντομογρ.: DCT.

- Feature 24 - Φασματική Διακύμανση (Spectral Irregularity):

Η συνάρτηση *mirregularity* προβαίνει σε υπολογισμό της διακύμανσης του spectrum (βλ. Εικόνα 18), της διαφοροποίησης δηλαδή ανάμεσα σε διαδοχικές κορυφές (peaks). Υπάρχουν δύο τύποι μέτρησης της Φασματική Διακύμανσης: Ο τύπος του Krimphoff (Krimphoff et al., 1994), όπου η Διακύμανση προκύπτει από το άθροισμα των πλατών μείον τη μέση τιμή του προηγούμενου, του ίδιου, και του επόμενου πλάτους

$$\sum_{k=2}^{N-1} \left| a_k - \frac{a_{k-1} + a_k + a_{k+1}}{3} \right|$$

και ο τύπος του Jensen (1999), όπου η Διακύμανση προκύπτει από το άθροισμα των τετραγώνων των διαφορών μεταξύ των πλατών δύο διαδοχικών κορυφών.

$$(\sum_{k=1}^N (a_k - a_{k+1})^2) / \sum_{k=1}^N a_k^2$$

- Features 25 έως 27 - Παραλειπόμενες παράμετροι:

Όπως προαναφέρθηκε, αρχικά, χρησιμοποιήθηκαν τρία επιπλέον χαρακτηριστικά τα οποία τελικώς απορρίφθηκαν. Η λειτουργία των *mirattacktime*, *mirattackleap* και *mirattackslope* αφορά στην φάση της ατάκας της ADSR, δηλαδή της δυναμικής περιβάλλουσας ενός ήχου. Ειδικότερα: η συνάρτηση *mirattacktime* προβαίνει σε μέτρηση της χρονικής διάρκειας της ατάκας ενός ήχου, η συνάρτηση *mirattackleap* προβαίνει σε υπολογισμό της διαφοράς της έντασης (πλάτους) κατά το ξεκίνημα και κατά το τελείωμα της ατάκας, ενώ η συνάρτηση *mirattackslope* εκτιμά την κλίση της ατάκας.

Ενώ θεωρητικά θα ήταν ωφέλιμο να χρησιμοποιηθούν αυτά τα χαρακτηριστικά, προκειμένου να συγκριθεί η μικροδομή των επιμέρους ήχων, πρακτικά υπήρξαν δυσκολίες στη χρησιμοποίησή τους. Συγκεκριμένα, δεν κατέστη εφικτή η εισαγωγή τους στον ταξινομητή γιατί σε μερικά αρχεία ήχου τα χαρακτηριστικά αυτά δεν είχαν λάβει τιμή. Αυτό καταγράφεται με εύγλωττο τρόπο στους πίνακες του Παραρτήματος 3, όπου σε κάποια αρχεία ήχου στις θέσεις Feature 25, Feature 26 και Feature 27 υπάρχει η ένδειξη «NaN». Πιθανώς η εξήγηση για αυτό αφορά στην κακή ποιότητα του αρχείου προέλευσης, σε συνδυασμό ίσως με την – υποχρεωτική – χρήση φίλτρου για την αφαίρεση του ανεπιθύμητου θορύβου. Αναπόφευκτα, η εκπαίδευση του ταξινομητή έγινε με τις υπόλοιπες είκοσι τέσσερις τιμές.

4.2.1 Ο κώδικας

Για την εκπαίδευση του ταξινομητή χρειάστηκε να γίνει μια βάση δεδομένων με όλα τα features και από τα τρία είδη πουλιών. Για να διευκολυνθεί η εξαγωγή τους και η εισαγωγή τους στην database, αλλά και επειδή σε δεύτερο χρόνο, όταν ο ταξινομητής θα έπρεπε να αναγνωρίσει άγνωστα δείγματα, έπρεπε να δημιουργηθεί μία συνάρτηση στο MATLAB η οποία θα είχε τη δυνατότητα να κάνει την εξαγωγή όλων των χαρακτηριστικών μαζί. Σε μερικές απλές γραμμές κώδικα γράφτηκε η συνάρτηση `extractBasicFeats(a)`, όπου το `input` είναι ένα `miraudio` object – ένα αρχείο ήχου δηλαδή που έχει «διαβαστεί» με την συνάρτηση `miraudio` του `MIRtoolbox`, και το `output` είναι ένα διάνυσμα με τις τιμές από όλα τα χαρακτηριστικά που συλλέχτηκαν. Όσον αφορά στην εκπαίδευση του ταξινομητή, η εξαγωγή έγινε σε ολόκληρους τους φακέλους με τους ήχους των πουλιών²⁷, σε 50 αρχεία ανά πουλί δηλαδή. Όλα τα διανύσματα συγκεντρώθηκαν σε ένα αρχείο `.xls` χωρισμένα σε τρεις κλάσεις: «Γκιώνης», «Κούκος», «Δεκαοχτούρα». Παρατίθεται ο κώδικας της συνάρτησης:

```
function f = extractBasicFeats(a)

    f = [];

    f = [f; mirgetdata(mirrms(a))];

    f = [f; mirgetdata(mirlowenergy(a))];

    f = [f; mirgetdata(mirzerocross(a))];

    f = [f; mirgetdata(mirrolloff(a))];

    f = [f; mirgetdata(mirbrightness(a))];

    f = [f; mirgetdata(mircentroid(a))];

    f = [f; mirgetdata(mirspread(a))];

    f = [f; mirgetdata(mirskewness(a))];

    f = [f; mirgetdata(mirflatness(a))];

    f = [f; mirgetdata(mirentropy(a))];

    f = [f; mirgetdata(mirmfcc(a))];

    f = [f; mirgetdata(mirregularity(a))];

    f = f.';

end
```

²⁷ Σημ. αντίθετα, στην αναγνώριση άγνωστων δεδομένων προς το παρόν δεν υπάρχει η δυνατότητα επεξεργασίας πολλών αρχείων ταυτόχρονα.

4.3 Η επιλογή του ταξινομητή

Στο τελικό στάδιο της διαδικασίας εκπαιδεύτηκε ένας ταξινομητής, η επιλογή του οποίου έγινε με κριτήριο την μέγιστη δυνατή ακρίβεια. Συγκεκριμένα, με χρήση της εφαρμογής MATLAB Classification Learner, εισήχθη στο σύστημα η βάση δεδομένων με τα 150 training data (βλ. Παράρτημα 5, Εικόνα 1). Αρχικά ορίστηκαν 27 predictors και 1 response (otus scops / cuculus canorus / streptopelia decaocto), τελικώς όμως, όπως προαναφέρθηκε, 3 από τα χαρακτηριστικά αφαιρέθηκαν γιατί δεν πληρούσαν τις προϋποθέσεις και η εκπαίδευση έγινε με 24 predictors (βλ. Παράρτημα 5, Εικόνα 2). Έγινε διασταυρούμενη επικύρωση με 5 folds – η προεπιλεγμένη επιλογή - και εκπαιδεύτηκαν χωρίς χρήση PCA – λόγω των δεδομένων – όλοι οι αλγόριθμοι ταξινόμησης. Καλύτερα ποσοστά πέτυχαν τα Δένδρα Αποφάσεων (93.3%) και συγκεκριμένα τα «Ensembles of Bagged Trees» (βλ. Παράρτημα 5, Εικόνα 3). Τέλος, μετά την πρώτη εκπαίδευση του μοντέλου δεν κρίθηκε σκόπιμο να παραληφθεί κάποιο άλλο από τα features που χρησιμοποιήθηκαν.

Πριν την εκπαίδευση των ταξινομητών που δοκιμάστηκαν, επιλέχθηκε διασταυρούμενη επικύρωση 5 τμημάτων (5 fold cross-validation)²⁸, που ήταν και η προεπιλεγμένη μέθοδος του Classification Learner. Αυτή η πρακτική ανήκει στην κατηγορία της πολλαπλής επικύρωσης (k-fold cross-validation) όπου τα δεδομένα χωρίζονται τυχαία σε k ισομεγέθη υποσύνολα. Ένα από τα υποσύνολα χρησιμοποιείται ως σύνολο επικύρωσης (testing data) και τα υπόλοιπα συνενώνονται και χρησιμοποιούνται ως σύνολο εκπαίδευσης (training data). Η διαδικασία επαναλαμβάνεται k φορές, κάθε φορά χρησιμοποιώντας ένα διαφορετικό σύνολο ως σύνολο επικύρωσης και τα υπόλοιπα ως σύνολο εκπαίδευσης. Στο τέλος έχουν χρησιμοποιηθεί όλα τα σύνολα από μία φορά ως σύνολο εκπαίδευσης, και βάσει των αποτελεσμάτων υπολογίζεται η μέση επίδοση του μοντέλου.

Στο στάδιο της εκπαίδευσης των ταξινομητών, προκειμένου να αναδειχθεί ο πιο ακριβής, επιλέχθηκε να μην γίνει PCA (Principal Component Analysis)²⁹. Η PCA μπορεί να διευκολύνει ως προς τον υπολογιστικό χρόνο που απαιτείται σε διάφορες εργασίες στη μηχανική μάθηση,

²⁸ Η διασταυρούμενη επικύρωση (cross-validation) αποτελεί ένα σύνολο παρόμοιων τεχνικών επικύρωσης μοντέλου για την αξιολόγηση του τρόπου γενίκευσης των αποτελεσμάτων μιας στατιστικής ανάλυσης σε ένα ανεξάρτητο σύνολο δεδομένων. Ο στόχος της διασταυρούμενης επικύρωσης είναι να δοκιμάσει την ικανότητα του μοντέλου να προβλέπει νέα δεδομένα που δεν χρησιμοποιήθηκαν για την εκτίμησή του, προκειμένου να επισημάνει προβλήματα όπως η Υπερμοντελοποίηση (Overfitting) – όταν δηλαδή η παραγωγή μίας ανάλυσης είναι προσαρμοσμένη σε ένα συγκεκριμένο σύνολο δεδομένων με αποτέλεσμα να μην μπορεί να προσαρμοστεί σε νέα δεδομένα ή να κάνει νέες προβλέψεις αξιόπιστα - και να δώσει μια εικόνα για το πώς το μοντέλο θα γενικευθεί σε ένα άγνωστο σύνολο δεδομένων.

²⁹ Η Ανάλυση Κύριων Συνιστωσών (συντομογρ.: PCA) είναι μία γραμμική μέθοδος συμπίεσης δεδομένων η οποία υπολογίζει τους άξονες στους οποίους παρατηρείται η μέγιστη διασπορά δεδομένων. Αποτελεί βασική τεχνική στην μείωση διαστάσεων (Dimensionality Reduction, συντομογρ.: DR), μια μεθοδολογία η οποία έχει ως σκοπό να προβάλει ένα σύνολο πολυδιάστατων δεδομένων σε μία δομή χαμηλής διάστασης.

ωστόσο δεν ενδείκνυται για δεδομένα που περιέχουν πολύ θόρυβο ή για περιπτώσεις κατηγοριοποίησης που η διακύμανση εντός των κλάσεων είναι μεγάλη συγκριτικά με αυτήν ανάμεσα στις κλάσεις. Λαμβάνοντας υπόψη την ποιότητα των δεδομένων, το ότι επιλέχθηκαν να συμπεριληφθούν calls και songs μαζί στην ίδια κλάση, και σε συνδυασμό με τις πληροφορίες (βλ. Παράρτημα 5, Εικόνα 3) από το Διάγραμμα Διασποράς (Scatter Plot) κρίθηκε ωφέλιμη η μη χρησιμοποίηση της.

Αφού πραγματοποιήθηκε εκπαίδευση όλων των δυνατών αλγορίθμων ταξινόμησης, τελικώς επιλέχθηκαν να χρησιμοποιηθούν τα Ensembles of Bagged Trees που πέτυχαν και το μεγαλύτερο ποσοστό ακρίβειας (93.3%). Τα Δένδρα Αποφάσεων³⁰ αποτελούν ίσως τον δημοφιλέστερο αλγόριθμο μηχανικής μάθησης με επιτήρηση, η περίπτωση δηλαδή που όλα τα δεδομένα είναι κατηγοριοποιημένα εξ αρχής και οι αλγόριθμοι εκπαιδεύονται στο να δίνουν αποτέλεσμα πρόβλεψης μέσω των δεδομένων που τους εισάγονται. Όσον αφορά συγκεκριμένα στα Ensembles of Bagged Trees, πρόκειται για εφαρμογή πολλαπλών Δένδρων Αποφάσεων με την τεχνική Bagging (Bootstrap Aggregation). Με την τεχνική αυτή δημιουργούνται διάφορα υποσύνολα δεδομένων από ένα δείγμα εκπαίδευσης που επιλέγεται τυχαία με αντικατάσταση. Κάθε συλλογή δεδομένων από τα υποσύνολα αυτά χρησιμοποιείται για την εκπαίδευση των Δένδρων Αποφάσεων, με αποτέλεσμα να καταλήγουμε σε ένα σύνολο διαφορετικών μοντέλων. Στο τέλος χρησιμοποιείται ο μέσος όρος των προβλέψεων από τα διάφορα δέντρα, κάτι που ως πρακτική παρέχει μεγαλύτερη ακρίβεια από ένα μόνο Δένδρο Αποφάσεων.

Στον Πίνακα Σύγχυσης³¹ για το μοντέλο εκπαίδευσης του birdMelos (Εικόνες 19 και 20) σημειώθηκαν τα εξής αποτελέσματα: Ο κούκος αναγνωρίστηκε σωστά 48 φορές, αναγνωρίστηκε εσφαλμένα ως δεκαοχτούρα 1 φορά και αναγνωρίστηκε εσφαλμένα 1 φορά ως γκιώνης. Η δεκαοχτούρα αναγνωρίστηκε σωστά 42 φορές, αναγνωρίστηκε εσφαλμένα ως κούκος 7 φορές και αναγνωρίστηκε εσφαλμένα ως γκιώνης 1 φορά. Ο γκιώνης αναγνωρίστηκε σωστά 50 φορές. Από το άθροισμα των ποσοστών σωστής αναγνώρισης προέκυψε η ακρίβεια του ταξινομητή 93.3%.

³⁰ Η ονομασία τους προέρχεται από τη δενδροειδή δομή τους: κάθε εσωτερικός κόμβος του δένδρου προσδιορίζει τον έλεγχο των χαρακτηριστικών και κάθε κλαδί που συνδέει τους εσωτερικούς με τους απογόνους κόμβους αντιστοιχεί σε μία πιθανή τιμή για το χαρακτηριστικό. Στους εξωτερικούς κόμβους αντιστοιχούν οι τιμές εξόδου, οι κλάσεις των δεδομένων.

³¹ Στη μηχανική μάθηση, και δη στη στατιστική ταξινόμηση, ένας Πίνακας Σύγχυσης ή Πίνακας Σφάλματος (Stehman, 1997) είναι μια ειδική διάταξη πίνακα που επιτρέπει την απεικόνιση της απόδοσης ενός αλγορίθμου. Εφαρμόζεται συνήθως στην μάθηση με επίβλεψη. Κάθε σειρά του πίνακα αντιπροσωπεύει μία πραγματική κλάση ενώ κάθε στήλη αντιπροσωπεύει μία προβλεπόμενη. Το όνομα του σχετίζεται με το ότι μας διευκολύνει να δούμε σε τι ποσοστό το σύστημα συγχέει δύο διαφορετικές κλάσεις.

Model 1

True class	cuckoo	48	1	1
	decaocto	7	42	1
	otus			50
		cuckoo	decaocto	otus
		Predicted class		

Εικόνα 19: Confusion matrix – number of observations

Model 1

True class	cuckoo	96%	2%	2%		96%	4%
	decaocto	14%	84%	2%		84%	16%
	otus			100%		100%	
		cuckoo	decaocto	otus		True Positive Rate	False Negative Rate
		Predicted class					

Εικόνα 20: Confusion matrix – true positive rates / false negative rates

4.4 Δοκιμές – αποτελέσματα

Τα δεδομένα δοκιμής αναγνωρίστηκαν από τον ταξινομητή με ποσοστό ακρίβειας 80% για το σύνολό τους (βλ. Παράρτημα 4). Συγκεκριμένα στα δεδομένα του γκιώνη πραγματοποιήθηκαν 13 σωστές αναγνωρίσεις (13/21, 62%), στα δεδομένα του κούκου όλα τα δείγματα αναγνωρίστηκαν σωστά (21/21, 100%), τέλος, στα δεδομένα της δεκαοχτούρας πραγματοποιήθηκαν 17 σωστές αναγνωρίσεις (17/21, 81%). Από τις 8 εσφαλμένες αναγνωρίσεις στην περίπτωση του γκιώνη οι 6 έδειξαν «κούκο» και οι 2 έδειξαν «δεκαοχτούρα», ενώ στην περίπτωση της δεκαοχτούρας και οι 4 εσφαλμένες αναγνωρίσεις έδειξαν «κούκο». Βάσει των δεδομένων που χρησιμοποιήθηκαν η συνολική ακρίβεια του συστήματος κρίνεται ως αρκετά ικανοποιητική.

Μία προσπάθεια ερμηνείας των αποτελεσμάτων αφορά στην ποιότητα των δειγμάτων που χρησιμοποιήθηκαν τόσο για την εκπαίδευση του ταξινομητή, όσο και για τις δοκιμές που ακολούθησαν: τα αρχεία που αποκτήθηκαν από το διαδίκτυο δεν είχαν ηχογραφηθεί από τον ίδιο εξοπλισμό, ούτε υπό τις ίδιες συνθήκες. Ανεπιθύμητοι θόρυβοι και κυρίως η ύπαρξη άλλων πουλιών και ζώων αναπόφευκτα επηρέασαν την αναγνώριση. Είναι γεγονός ότι τα προβλήματα που απορρέουν από την κακή ποιότητα των δεδομένων συναντώνται στις πρώτες κιόλας μελέτες που έγιναν πάνω στο πεδίο της Αυτόματης Αναγνώρισης του Birdsong. Το ζήτημα του θορύβου κυριαρχεί στην πλειοψηφία των ερευνών (Baker & Logue , 2003), ενώ παράλληλα σε πολλές από αυτές (Cai et al., 2007) γίνεται λόγος για το κατά πόσο η παρουσία άλλων πουλιών στις ηχογραφήσεις αποτελεί εμπόδιο στη διαδικασία της αναγνώρισης.

Στην περίπτωση του κούκου - όπου επιτεύχθηκε απόλυτη ακρίβεια – είναι ενδιαφέρον ότι ακόμα και οι λανθασμένες αναγνωρίσεις των άλλων πουλιών έδειξαν «κούκο» σε ποσοστό 83.3% (10/12). Μία πιθανή εξήγηση είναι η εξής: ό,τι δείγμα εισάγουμε στο σύστημα για αναγνώριση- είτε αυτό είναι άλλο πουλί, είτε οποιοσδήποτε άλλος ήχος – το σύστημα πάντα θα μας δώσει το πλησιέστερο αποτέλεσμα στα δεδομένα εκπαίδευσής του. Ειδικά όσον αφορά στην σχέση κούκου - δεκαοχτούρας, τα δύο πουλιά πέραν της μίας συλλαβής που τα χωρίζει, ομοιάζουν πολύ στο συχνοτικό πεδίο. Ο κούκος παράγει συχνότητες σε ένα μέσο εύρος από 700 έως 900 Hz, ενώ το συχνοτικό εύρος της δεκαοχτούρας είναι ελαφρώς χαμηλότερο. Αντίθετα ο γκιώνης παράγει πιο ψηλές συχνότητες, περίπου στα 1500 Hz. Λαμβάνοντας υπόψη τα παραπάνω, μπορεί να γίνει αντιληπτό γιατί και οι 4 λανθασμένες αναγνωρίσεις της δεκαοχτούρας έδειξαν «κούκο». Το γεγονός αυτό συνάδει πλήρως και με τα δεδομένα του Πίνακα Σύγχυσης κατά την εκπαίδευση του μοντέλου.

Τέλος, μία άλλη παράμετρος είναι η ενσωμάτωση κάποιων calls (5) στα δεδομένα εκπαίδευσης που παρέκλιναν αρκετά από τα τυπικά calls του είδους. Συγκεκριμένα είναι τα αρχεία: XC188747, XC383067 και XC382542 (γκιώνης), XC402141 (κούκος), XC118361 (δεκαοχτούρα). Ειδικά όσον αφορά στα καλέσματα του γκιώνη και το κάλεσμα της δεκαοχτούρας, είναι σχεδόν παρεμφερή. Σε κάποιες δοκιμές που έγιναν, επανεκπαιδεύτηκε ο ταξινομητής χωρίς να συμπεριληφθούν αυτά τα αρχεία και τα αποτελέσματα παρουσιάστηκαν σχετικά βελτιωμένα³².

³² Σημ. η συνολική ακρίβεια της αναγνώρισης προσέγγισε το 82%.

4.5 Η δημιουργία του interface

Ένα Γραφικό Περιβάλλον Χρήστη (Graphical User Interface, GUI) είναι ένα σύνολο εικονικών στοιχείων, τα οποία χρησιμοποιούνται για να διευκολύνουν και να επιταχύνουν την αλληλεπίδραση μεταξύ του χρήστη και της συσκευής. Παρέχει, μέσω εικόνων, ενδείξεις και μέσα προκειμένου ο χρήστης γρήγορα και με απλές ενέργειες να εκτελέσει συγκεκριμένες εργασίες. Με την ολοκλήρωση της εξαγωγής των παραμέτρων και της εκπαίδευσης του classifier, έπρεπε να δημιουργηθεί ένα τέτοιο, εύχρηστο interface μέσω του οποίου θα μπορούσε κάποιος να αναγνωρίζει νέα δείγματα χωρίς να χρειάζεται να γράφει κώδικα στο MATLAB. Σε αυτό το πρόβλημα έδωσε λύση το MATLAB App Designer, ένα περιβάλλον για την κατασκευή εφαρμογών στο MATLAB. Τα αρχεία που εξάγει το App Designer δεν είναι τύπου script αλλά αρχεία τύπου app, εφαρμογές δηλαδή. Αποκτά πλέον ο χρήστης ένα περιβάλλον GUI με το οποίο μπορεί εύκολα, με «point-and-click», να ελέγξει τις διάφορες λειτουργίες της εφαρμογής.

Μόλις ολοκληρώθηκε η εφαρμογή, δοκιμάστηκε ως προς την ταχύτητά της. Δεν διαπιστώθηκαν προβλήματα ως προς τις διάφορες λειτουργίες του GUI. Όσον αφορά την ταχύτητα της αναγνώρισης, ήταν ακριβώς η ίδια ταχύτητα με την οποία αναγνώριζε τα δείγματα το MATLAB γράφοντας κώδικα. Με βάση διάφορες δοκιμές που έγιναν, σε έναν σύγχρονο ηλεκτρονικό υπολογιστή η ταχύτητα αναγνώρισης του birdMelos καλύπτει ένα εύρος διάρκειας από 5 sec έως και 50 sec. Αρχεία «καθαρά» και με μικρή χρονική διάρκεια θα αναγνωριστούν ευκολότερα, ενώ για αρχεία με μεγαλύτερη διάρκεια των 40 sec και με πλούσιο συχνοτικό περιεχόμενο θα χρειαστεί περισσότερος χρόνος. Η πιο χρονοβόρα διαδικασία είναι η ανίχνευση «κορυφών» και οι διαφοροποιήσεις που παρουσιάζουν (Feature 24 - mirregularity).

4.5.1 Άνοιγμα της εφαρμογής

Η εφαρμογή κατά την εκκίνησή της (βλ. Παράρτημα 6, Εικόνα 1) εκχωρεί μηδενικές μεταβλητές για τα δεδομένα (data) και την αναγνώριση (yfit) στο workspace του MATLAB. Αυτό γίνεται για λειτουργικούς λόγους, στη συνέχεια αυτές οι μεταβλητές αντικαθίστανται από τις μεταβλητές που προκύπτουν από το αρχείο ήχου που έχει επιλεγεί και από το αποτέλεσμα της αναγνώρισης.

```
function startupFcn(app)
    yfit='0';
    assignin('base','yfit', yfit);
    data = struct('data', {}, 'fs', {});
    assignin('base','data',data);
end
```

4.5.2 search file Button

Το «search file» button δίνει στον χρήστη την δυνατότητα να εισαγάγει στο σύστημα το αρχείο ήχου προς αναγνώριση. Η εφαρμογή δέχεται αποκλειστικά αρχεία τύπου .wav. Εφόσον επιλέξει αρχείο, τα δεδομένα του εισάγονται στο workspace:

```
data = importdata(file);
(assignin('base','data',data));
```

Δημιουργείται ένα νέο UI Button με γκρι φόντο στο τετράγωνο πλαίσιο στο κέντρο της εφαρμογής:

```
BlankButton = uibutton(app.birdMelosVol11LabMAT2019UIFigure);
BlankButton.Text = '';
BlankButton.HorizontalAlignment = 'center';
BlankButton.VerticalAlignment = 'top';
BlankButton.FontName = 'Arial Black';
BlankButton.FontSize = 24;
BlankButton.FontColor = [1.00,0.00,0.00];
BlankButton.BackgroundColor = [0.94,0.94,0.94];
BlankButton.Position = [123.5 67 351 230];
```

Μέσα στο οποίο αναγράφεται το όνομα του αρχείου:

```
lbl = uilabel(app.birdMelosVol11LabMAT2019UIFigure);
lbl.Text = file;
lbl.HorizontalAlignment = 'center';
lbl.VerticalAlignment = 'center';
lbl.Enable = 'on';
lbl.Visible = 'on';
lbl.FontName = 'Arial Black';
lbl.FontSize = 12;
lbl.FontColor = [0.00,0.00,0.00];
lbl.BackgroundColor = [0.94,0.94,0.94];
lbl.Position = [149.5 185 300 15];
lbl.HandleVisibility = 'on';
```

(βλ. Παράρτημα 6, Εικόνα 2)

4.5.3 play / stop Buttons

Με το πάτημα του κουμπιού «play», η εφαρμογή αναζητά στο workspace τα δεδομένα που θα «τρέξει». Εάν ο χρήστης δεν έχει επιλέξει αρχείο ήχου, τότε ενημερώνεται με μήνυμα από την εφαρμογή. Εφόσον έχει επιλεγεί αρχείο τότε η εφαρμογή μέσω της συνάρτησης `audioplayer(Y,Fs)` δημιουργεί ένα object το οποίο αναπαράγει τον ήχο του αρχείου:

```
function playButtonPushed(app, event)
    data = evalin('base', 'data');
    if data.data == 0
        f = msgbox('No file selected', 'Warning', 'warn');
        beep;
    else pl=audioplayer(data.data,data.fs);
        assignin('base','pl', pl);
        play(pl);
    end
```

Κατά τον ίδιο ακριβώς τρόπο λειτουργεί και το κουμπί «stop».

4.5.4 recognise Button

Όταν πατηθεί από τον χρήστη το κουμπί «recognize», το σύστημα ξεκινά να κάνει όλες τις διαδικασίες που απαιτούνται για την εξαγωγή των χαρακτηριστικών και προβαίνει σε αναγνώριση του νέου δείγματος. Αρχικά «φορτώνει» τα δεδομένα του νέου δείγματος που έχουν εισαχθεί στο σύστημα όταν ο χρήστης επέλεξε αρχείο ήχου με το κουμπί «select file». Στην περίπτωση που ο χρήστης δοκιμάσει να προβεί σε αναγνώριση χωρίς πρώτα να έχει επιλέξει αρχείο, πληροφορείται μέσω ενός μηνύματος που αναδύεται στην οθόνη ότι πρέπει πρώτα να επιλέξει αρχείο ήχου. Στην συνέχεια ελέγχεται από την εφαρμογή – με συνθήκες `if` – εάν το αρχείο που έχει επιλεγεί είναι stereo ή mono. Εάν τελικώς το αρχείο είναι stereo μετατρέπεται σε mono:

```
if size(data.data,2) > 1
    Mono = (data.data(:,1) + data.data(:,2)) / 2;
```

Σε αυτό το σημείο αναδύεται ένα παράθυρο (progress bar) που ενημερώνει τον χρήστη ότι το σύστημα επεξεργάζεται τα δεδομένα (βλ. Παράρτημα 6, Εικόνα 2). Τα δεδομένα μετατρέπονται σε `miraudio` object προκειμένου να εξαχθούν τα χαρακτηριστικά τους με το

MIRtoolbox και στη συνέχεια φιλτράρονται για να αφαιρεθεί ο θόρυβος. Συγκεκριμένα περνάνε οι συχνότητες από 400 Hz έως 2000 Hz:

```
a = miraudio(Mono, data.fs);
mirfilterbank(a, 'Manual', [400 2000]);
```

Μέσω του function `extractBasicFeats(a)` εξάγονται τα χαρακτηριστικά τους και εκχωρούνται σε έναν πίνακα στο workspace. Στη συνέχεια το σύστημα φορτώνει τον εκπαιδευμένο ταξινομητή (`trainedModel.mat`) και ξεκινάει η διαδικασία της αναγνώρισης. Σε αυτό το σημείο το αποτέλεσμα της αναγνώρισης εκχωρείται ως μεταβλητή στο workspace προκειμένου να ενημερωθεί το κουμπί «info» και κλείνει η progress bar. Με συνθήκες if, αναλόγως το ποιά κλάση θα υποδείξει ο ταξινομητής, δημιουργείται ένα uibutton στο τετράγωνο πλαίσιο της εφαρμογής, με την εικόνα και το όνομα του πουλιού, ενώ παράλληλα ακούγεται και ένας ήχος προς ενημέρωση του χρήστη ότι η διαδικασία ολοκληρώθηκε (βλ. Παράρτημα 6, Εικόνα 3).

4.5.5 info Button

Το κουμπί «info» παρέχει στον χρήστη πληροφορίες σχετικά με το είδος του πουλιού που αναγνωρίστηκε από την εφαρμογή (βλ. Παράρτημα 6, Εικόνα 3). Αρχικά ανατρέχει στο workspace για να βρει τη μεταβλητή `yfit` που περιέχει το αποτέλεσμα της αναγνώρισης του ταξινομητή. Εάν δεν υπάρχει καταχωρημένη μεταβλητή, ενημερώνει τον χρήστη ότι δεν έχει πραγματοποιηθεί αναγνώριση. Αναλόγως το αποτέλεσμα ανοίγει ένα ξεχωριστό παράθυρο – UI Figure με τις πληροφορίες του πουλιού:

```
function Button_4Pushed(app, event)
    yfit = evalin('base', 'yfit');
    if yfit == '0'
        f = msgbox('No bird has been classified yet', 'Warning', 'warn');
        beep;
    elseif yfit == 'otus'
        Otusinfo;
    elseif yfit == 'cuckoo'
        cuckooinfo;
    elseif yfit == 'decaocto'
        streptopeliainfo;
    end
```

4.6 Σύνοψη και συμπεράσματα

Με το πέρας της διαδικασίας αυτής αναδύονται σημαντικά συμπεράσματα και προβληματισμοί για μελλοντική έρευνα. Αναμφίβολα, για τα δεδομένα αυτά, ο συγκεκριμένος αλγόριθμος λειτούργησε πολύ καλά. Το 80% ακρίβειας είναι ένα αρκετά ικανοποιητικό ποσοστό, και σίγουρα επιδέχεται βελτίωσης με επανεκπαίδευση του ταξινομητή.

Αναφορικά με τα δεδομένα εκπαίδευσης που χρησιμοποιήθηκαν μπορεί να λεχθεί συμπερασματικά ότι αδιαμφισβήτητα η κακή ποιότητα τους δυσκόλεψε τη διαδικασία επεξεργασίας τους και επιδείνωσε την ακρίβεια της αναγνώρισης. Ιδανικά, στη μηχανική μάθηση, ισχύει ότι πάντα τα πολλά δεδομένα υπερτερούν των λίγων. Ωστόσο, από τα ήδη λίγα διαθέσιμα δεδομένα του διαδικτύου χρειάστηκε να γίνει επιμέρους διαλογή με κριτήριο την ποιότητα και την χρονική τους διάρκεια.

Ένα άλλο ζήτημα που προκύπτει από το στάδιο της προεπεξεργασίας είναι η χρήση υψιπερατών (high-pass) και χαμηλοπερατών (low-pass) φίλτρων για την αφαίρεση του θορύβου. Είναι γεγονός ότι γνωρίζοντας εξ αρχής το συχνοτικό εύρος των πουλιών που χρησιμοποιήθηκαν, αφαιρέθηκαν όλες οι συχνότητες πάνω και κάτω από αυτό. Όμως σε συνθήκες πεδίου, και εφόσον η αναγνώριση δεν αφορά μόνο τρεις ομάδες πουλιών αλλά περισσότερων, κάτι τέτοιο δεν θα λειτουργήσει σωστά. Υπάρχει πολύ μεγάλη συζήτηση ως προς το θέμα αυτό, και πλέον προτείνονται και εναλλακτικές επιλογές για την αφαίρεση του θορύβου, όπως παραδείγματος χάρη η χρήση wavelets (Priyadarshani et al. 2016, Priyadarshani 2017) ή ακόμα και η εφαρμογή ανάλυσης εικόνας στα φασματογραφήματα (Potamitis, 2014).

Ως προς την χρήση των συγκεκριμένων χαρακτηριστικών που επιλέχθηκαν, από άποψη απόδοσης και βάσει αποτελέσματος, κρίνονται ικανοποιητικά. Ωστόσο όταν το ζητούμενο δεν είναι η μέγιστη δυνατή ακρίβεια της αναγνώρισης αλλά η ταχύτητα, τότε προκύπτουν δυσκολίες. Ιδιαίτερα δε όταν η διάρκεια του αρχείου είναι μεγάλη και η ποιότητα του ήχου κακή. Σίγουρα μία πιο εύχρηστη εφαρμογή θα απαιτούσε την παράλειψη κάποιων χαρακτηριστικών που επιβάρυναν την ταχύτητα του συστήματος. Αυτό δε είναι επιβεβλημένο, εφόσον μελλοντικά προστεθούν στην εφαρμογή και άλλα είδη πουλιών με πιο σύνθετα φωνήματα, τα οποία ζουν σε πιο θορυβώδη περιβάλλοντα.

Κλείνοντας, αντικείμενο υπό εξέταση αποτελεί επίσης το είδος των μετατροπών που πρέπει να γίνουν στον κώδικα ώστε μία εφαρμογή για ηλεκτρονικό υπολογιστή, δημιουργημένη στο MATLAB, να μπορεί να προβεί σε αναγνωρίσεις σε πραγματικό χρόνο (real time). Σίγουρα για την προσθήκη μίας τέτοιας λειτουργίας απαιτείται τελείως διαφορετική προσέγγιση, ωστόσο, όπως λέχθηκε και νωρίτερα, εξαρτάται τι στόχους έχει θέσει κάποιος αρχικά, όταν ξεκινά να δημιουργεί μία εφαρμογή.

5. Επίλογος

Συνοψίζοντας, είναι έκδηλο ότι υπάρχει ακόμη περιθώριο προς περαιτέρω έρευνα σχετικά με τους ήχους των πουλιών, ιδιαίτερα ως προς το τεχνολογικό κομμάτι. Ζητούμενο για τον μελλοντικό ερευνητή δεν μπορεί παρά να είναι η βελτίωση της υπάρχουσας τεχνολογίας ώστε τα συστήματα αναγνώρισης να αποκτήσουν μεγαλύτερη ακρίβεια, να γίνει δυνατή η ταυτόχρονη αναγνώριση πολλών ειδών στην ίδια ηχογράφηση και να καταστεί, πλέον, εφικτή η διεξοδική καταγραφή περισσότερων ειδών. Παράλληλα, με αναβάθμιση των αλγορίθμων αναγνώρισης και με επέκταση των βάσεων δεδομένων του διαδικτύου, θα αντιμετωπιστούν ζητήματα όπως το αυτό των διαλέκτων των πουλιών. Ειδικότερα όσον αφορά στα ελληνικά πουλιά, η δημιουργία βάσης δεδομένων με καταγραφές από την ελληνική επικράτεια είναι απαραίτητη, αφού προς το παρόν μόνο από το διαδίκτυο μπορούν να ανακτηθούν κάποιες λίγες καταγραφές ελληνικών πουλιών, συνήθως χαμηλής ποιότητας. Στα πρώτα στάδια της δημιουργίας της εφαρμογής birdMelos έγιναν κάποιες προσπάθειες ανεύρεσης δεδομένων από ελληνικά πουλιά, είτε μέσω επαφών με περιβαλλοντολόγους είτε με την Ελληνική Ορνιθολογική Εταιρεία, αλλά ήταν άκαρπες.

Η καταγραφή της τοπικής ορνιθοπανίδας και η χρήση των νέων τεχνολογιών μπορούν, παράλληλα, να αποφέρουν οφέλη σε διάφορους τομείς. Παραδείγματος χάρη στην πολιτιστική οικονομία και στον πολιτιστικό τουρισμό: φορητά ή εγκατεστημένα διαδραστικά συστήματα σε εθνικούς δρυμούς ή αρχαιολογικούς χώρους μπορούν να παρέχουν πληροφορίες στον επισκέπτη σχετικά με το ηχοτοπίο που τον περιβάλλει. Θα πρέπει βέβαια μέσω μίας σύμπραξης της επιστημονικής κοινότητας με την κεντρική εξουσία, να οριστούν διεξοδικά οι μέθοδοι και οι πρακτικές ενός τέτοιου εγχειρήματος. Εφαρμογή θα βρουν επίσης στην επιστήμη και στην έρευνα, σε πεδία όπως η Οικολογία, η Περιβαλλοντολογία και η Βιολογία. Τέλος, στον τομέα της εκπαίδευσης, τα τελευταία χρόνια στην χώρα μας έχουν ήδη εφαρμοστεί εκπαιδευτικές προτάσεις που αφορούν ηχητικούς περιπάτους και ανάλυση του ηχοτοπίου του σχολικού περιβάλλοντος. Είναι γεγονός ότι η μελέτη των ήχων των πουλιών μπορεί να λειτουργήσει ευεργετικά για τους μαθητές, τόσο σε επίπεδο περιβαλλοντικής ευαισθητοποίησης όσο και εξοικείωσης με την τεχνολογία, αφού πλέον το μάθημα γίνεται διαδραστικό, δίνοντας τη δυνατότητα στα παιδιά να αξιοποιούν με διαφορετικό τρόπο τα καθημερινά τεχνολογικά μέσα που διαθέτουν (smartphone, υπολογιστής).

Με την ολοκλήρωση αυτής της εργασίας ευελπιστούμε να έχει προστεθεί ένα λιθαράκι προς αυτήν την κατεύθυνση, έχοντας καλύψει τις περισσότερες πτυχές του ζητήματος άνθρωπος – ήχοι πουλιών και, παράλληλα, έχοντας θέσει κάποιες βάσεις για μελλοντική έρευνα.

Βιβλιογραφία

1. Agranat, I. (2009). Automatically identifying animal species from their vocalizations. *Proceedings of the 5th International Conference on Bio-Acoustics*, Holywell Park.
2. Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G. & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ* 1: e103.
3. Andreassen, T., Surlykke, A. & Hallam, J. (2014). Semi-automatic long-term acoustic surveying: a case study with bats. *Ecol. Inform.* 21: 13–24.
4. Araya-Salas, M. (2012). Is birdsong music? Evaluating harmonic intervals in songs of a Neotropical songbird. *Animal Behaviour*, 84, 309–313.
5. Baker, M. C. & Logue, D. M. (2003). Population differentiation in a complex bird sound: a comparison of three bioacoustical analysis procedures. *Ethology* 109: 223–242.
6. Berwick, R. C., Okanoya, K., Beckers, G. J. L. & Bolhuis, J. J. (2011). *Songs to syntax: the linguistics of birdsong*. - CELL Press.
7. Brenowitz, E. A., Margoliash, D., Nordeen, K. W. (1997). An introduction to birdsong and the avian song system. *J Neurobiol.* 33:495–500.
8. Briggs, F., Lakshminarayanan, B., Neal, L., Fern, X. Z., Raich, R., Hadley, S. J., Hadley, A. S. and Betts, M. G. (2012). Acoustic classification of multiple simultaneous bird species: a multi-instance multi-label approach. *J. Acoust. Soc. Am.* 131: 4640–4650.
9. Bumgardner, J. (2009): Kircher's Mechanical Composer: A Software Implementation. - *Proceedings of Bridges 2009: Mathematics, Music, Art, Architecture, Culture*.
10. Cai, J., Ee, D., Pham, B., Roe, P., Zhang, J. (2007). Sensor Network for the monitoring of Ecosystem: Bird Species Recognition. *Proc. of the 3rd International Conference on Intelligent Sensors, Sensor Networks and Information (ISSNIP 2007)*, Melbourne, Australia.
11. Campbell, B., Lack, E. (2011). *A Dictionary of Birds*. A&C Black. pp. 369–372.
12. Charif, R., Strickman, L. & Waack, A. (2010). *Raven Pro 1.4 user's manual. Revision 11*. The Cornell Lab of Ornithology, Ithaca, NY.
13. Clemens, P. J. & Johnson, M. T. (2003). Application of speech recognition to African elephant (*Loxodonta africana*) vocalizations. *IEEE Int. Conf. Acoust. Speech Signal Process.* 1: 484–487.
14. de Oliveira, A. G., Ventura, T. M., Ganchev, T. D., de Figueiredo, J. M., Jahn, O., Marques, M. I. & Schuchmann, K. L. (2015). Bird acoustic activity detection based on morphological filtering of the spectrogram. *Appl. Acoust.* 98: 34–42.

15. Deng, L. & O'Shaughnessy, D. (2003). *Speech processing: a dynamic and optimization-oriented approach*. Marcel Dekker. pp. 41–48.
16. Duan, S., Towsey, M., Zhang, J., Truskinger, A., Wimmer, J. & Roe, P. (2011). Acoustic component detection for automatic species recognition in environmental monitoring. *Proc. 7th Int. Conf. Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pp. 514–519.
17. Duan, S., Zhang, J., Roe, P., Wimmer, J., Dong, X., Truskinger, A. & Towsey, M. (2013). Timed probabilistic automaton: a bridge between Raven and Song Scope for automatic species recognition. *Proc. 25th Innovative Applications of Artificial Intelligence Conference*, pp. 1519–1524.
18. Dufour, O., Artieres, T., Glotin, H. and Giraudet, P. (2013). Clusterized mel filter cepstral coefficients and support vector machines for bird song identification. *Proc. 1st workshop on Machine Learning for Bioacoustics 951*: 89–93.
19. Fagerlund, S. (2007). Bird species recognition using support vector machines. *EURASIP J. Appl. Signal Process.*, 2007(1):64-64.
20. Georgaki, A. (2017). Spectres d'une écho-mythologie dans l'oeuvre de François-Bernard Mâche. «*F.B Mâche, le compositeur et le savant face à l'univers sonore*», ed. Hermann, Paris (Broché).
21. Giannakopoulos, T. (2009). *Study and application of acoustic information for the detection of harmful content, and fusion with visual information*. PhD Thesis, University of Athens.
22. Griffiths, P. (1985). *Olivier Messiaen and the Music of Time*. Ithaca, New York: Cornell University Press.
23. Jančovič, P. & Köküer, M. (2016). «*Recognition of Multiple Bird Species Based on Penalised Maximum Likelihood and HMM-Based Modelling of Individual Vocalisation Elements*». Interspeech 2016, 09, 2016.
24. Jensen, K. (1999). «*Timbre Models of Musical Sounds*». Rapport 99/7, University of Copenhagen.
25. Jinnai, M., Boucher, N., Fukumi, M. & Taylor, H. (2012). A new optimization method of the geometric distance in an automatic recognition system for bird vocalisations. *Proceedings of the Acoustics 2012 Nantes Conference*, pp. 2439–2445.
26. Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26(6), 1797-813.

27. Kogan, J. A. & Margoliash, D. (1998). Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. *J. Acoust. Soc. Am.* 103: 2185–2196.
28. Kohavi, R. & Provost, F. (1998). «Glossary of terms». *Machine Learning* 30: 271–274.
29. Kraft, D. (2013). *Birdsong in the Music of Olivier Messiaen*. London: Arosa Press.
30. Krimphoff, J., McAdams, S. & Winsberg, S. (1994). Caractérisation du timbre des sons complexes. II : Analyses acoustiques et quantification psychophysique. *Journal de Physique*, 4(C5), 625-628.
31. Lartillot, O. & Toivianen, P. (2007). A Matlab Toolbox for Musical Feature Extraction From Audio. *International Conference on Digital Audio Effects*, Bordeaux.
32. Lartillot, O. (2018). *MIRtoolbox 1.7.1 User's Manual*. University of Oslo, Norway, Department of Musicology.
33. Lee, C. H., Lee, Y. K. & Huang, R. Z. (2006). Automatic recognition of bird songs using cepstral coefficients. *J. Inform. Technol. Appl.* 1: 17–23.
34. Lopes, M. T., Gioppo, L. L., Higushi, T. T., Kaestner, C. A., Silla Jr, C. N. & Koerich, A. L. (2011a). Automatic bird species identification for large number of species. *Proc. IEEE Int. Symp. Multimedia*, pp. 117–122.
35. Luther, D. & Baptista, L. (2009). Urban noise and the cultural evolution of bird songs. 277. *Proceedings of the Royal Society B: Biological Sciences*.
36. Mache F. B, (1983). *Musique, mythe, nature, ou les Dauphins d'Arion*. Paris: Méridiens Klincksieck.
37. Mâche, F. B. (1997). Syntagms and paradigms in zoomusicology. *Contemporary Music Review*, 16:3, 55-78.
38. Makhoul, J. & Schwartz, R. (1995). State of the art in continuous speech recognition. *Proc. Natl Acad. Sci. USA* 92: 9956–9963.
39. Marler, P. & Peters, S. (1987). A sensitive period for song acquisition in the song sparrow: a case of age-limited learning. *Ethology* 76:89-100.
40. Marler, P. & Slabbekoorn, H. (2004). *Natures Music: The Science of Birdsong*. Elsevier, Academic Press. pp. 386.
41. Marler, P. (1997). Three models of song learning : evidence from behavior. *J. Neurobiol.* 33:501-516.
42. McIlraith, A. L. & Card, H. C. (1995). Birdsong recognition with DSP and neural networks. *IEEE WESCANEX'95 Proceedings, Winnipeg, Canada*, pp. 409–414.
43. Notebohm, F., Stokes, T. & Leonard, C. (1976). Central control of song in the canary. *J. Comp. Neurol.* 165:457-486.

44. Nottebohm, F. (1987). Birdsong. *Encyclopedia of Neuroscience, Vol. I*. G. Adelman, Ed. Birkhauser Boston, Boston, pp. 133-136.
45. Nottebohm, F., Alvarez-Buylla, A., Cynx, J., Ling, C. Y., Nottebohm, M., Suter, R., Tolles, A., & Williams, H. (1990). Song learning in birds: the relation between perception and production. *Phil. Trans. R. Soc. B* 329:115-124.
46. Panagiotakis, C. & Tziritas, G. (2005). A speech/music discriminator based on rms and zero-crossings. *IEEE Transactions on Multimedia*, 7(1):155–166.
47. Papadopoulos, T., Roberts, S. & Willis, K. (2015). *Detecting bird sound in unknown acoustic background using crowdsourced training data*. ArXiv e-prints arXiv: 1505.06443.
48. Pohle, T., Pampalk, E. & Widmer, G. (2005). Evaluation of Frequently Used Audio Features for Classification of Music Into Perceptual Categories. *Proceedings of the Fourth International Workshop on Content-Based Multimedia Indexing (CBMI'05)*, Riga, Latvia, June 21-23.
49. Potamitis, I. (2014). Automatic Classification of a Taxon-Rich Community Recorded in the Wild. *PLoS ONE* 9(5): e96936.
50. Priyadarshani N, Marsland S, Castro I 2018. Automated birdsong recognition in complex acoustic environments: a review. *Journal of Avian Biology* 49: jav-01447.
51. Priyadarshani, N. (2017). *Wavelet-based birdsong recognition for conservation*. PhD thesis, Massey University, Palmerston North, New Zealand.
52. Priyadarshani, N., Marsland, S., Castro, I., & Punchihewa, A. (2016). Birdsong denoising using wavelets. *PloS one*, 11(1):e0146790.
53. Ross, D. J. (2006). *Bird Call Recognition With Artificial Neural Networks, Support Vector Machines and Kernel Density Estimation*. University of Manitoba.
54. Schafer, R. M. (1977). *The Tuning of the World (The Soundscape)*. Michigan: Knopf; original: University of Michigan.
55. Selin, A., Turunen, J. & Tantt, J. T. (2007). Wavelets in recognition of bird sounds. *EURASIP J. Appl. Signal Process.* 2007: 141–141.
56. Somervuo, P., Härmä, A. & Fagerlund, S. (2006). Parametric representations of bird sounds for automatic species recognition. *IEEE Trans. Audio Speech Language Process.* 14: 2252–2263.
57. Specht, R. (1993). *AVISOF—sound analysis and synthesis laboratory Pro: a PC program for sonographic analysis*. AVISOFT, Berlin, Germany.

58. Stattner, E., Segretier, W., Collard, M., Hunel, P. & Vidot, N. (2013). Song-based classification techniques for endangered bird conservation. *ICML 2013 Workshop on Machine Learning for Bioacoustics*, Atlanta, GA, USA.
59. Stehman, S. V. (1997). "Selecting and interpreting measures of thematic classification accuracy". *Remote Sensing of Environment*. 62 (1): 77–89.
60. Stowell, D. & Plumbley, M. D. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ* 2: e488.
61. Tchernichovski, O. (2012). *Sound Analysis Pro 2011 user manual*. Sound Analysis Pro.
62. Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Anim. Behav.* 59: 1167–1176.
63. Thorpe, W. H. (1958). *The learning of song patterns by birds, with especial references to the song of the chaffinch, Fringilla coelebs*. *Ibis* 100:535-570.
64. Tzanetakis & Cook. Musical genre classification of audio signals. *IEEE Tr. Speech and Audio Processing*, 10(5),293-302, 2002.
65. Wallin, N. L. (1991). *Biomusicology: Neurophysiological, Neuropsychological and Evolutionary Perspectives on the Origins and Purposes of Music*. Stuyvesant, NY: Pendragon Press.
66. Wildlife Acoustics (2011). *Song Scope bioacoustics software version 4.0 documentation*. Wildlife Acoustics.
67. Williams, H. (2004). Birdsong and singing behavior. *Ann. N. Y. Acad. Sci.* 1016: 1–30.
68. Σπυρίδης, Χ. (1996). *Μουσική Ακουστική*. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης. pp. 351 – 365.

Διαδικτυογραφία – Ηλεκτρονικές πηγές

1. ARBIMON: www.sieve-analytics.com
2. Audacity: www.audacityteam.org
3. Avibase: <https://avibase.bsc-eoc.org>
4. Avisoft-SASLab Pro: www.avisoft.com/soundanalysis
5. BatSound: www.batsound.com
6. Bird Song Id: isoperla.co.uk
7. ChirpOMatic: www.chirpomatic.com
8. Ishmael: www.bioacoustics.us/ishmael
9. Kaleidoscope Pro: www.wildlifeacoustics.com/download/kaleidoscope-software
10. Luscinia: rflachlan.github.io/Luscinia

11. Macaulay Library: www.macaulaylibrary.org
12. Mangoverde World Bird Guide: www.mangoverde.com/birdsound
13. MATLAB App Designer documentation:
www.mathworks.com/products/matlab/app-designer
14. MATLAB Classification Learner documentantion:
www.mathworks.com/help/stats/classificationlearner-app
15. MIRtoolbox: www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox
16. monitoR (R package): <https://CRAN.R-project.org/package=monitoR>
17. Praat: www.fon.hum.uva.nl/praat
18. Raven Pro: www.ravensoundsoftware.com/software/raven-pro
19. Song Scope: www.wildlifeacoustics.com/download/song-scope-software
20. Song Sleuth: www.songsleuth.com
21. Sonic Visualiser: www.sonicvisualiser.org
22. SonoBat: www.sonobat.com
23. Sound Analysis Pro: www.soundanalysispro.com
24. SoundID: www.soundid.net
25. SpectraPLUS: www.spectraplus.com
26. XC250986: White-crowned Sparrow - *Zonotrichia leucophrys oriantha*. Recordist: Ian Cruickshank. License: Creative Commons Attribution-NonCommercial-ShareAlike 4.0. url: <https://www.xeno-canto.org/250986>
27. XC325008: White-crowned Sparrow - *Zonotrichia leucophrys oriantha*. Recordist: Jim Holmes. License: Creative Commons Attribution-NonCommercial-ShareAlike 4.0. url: <https://www.xeno-canto.org/325008>
28. Xeno-canto: www.xeno-canto.org
29. Κεφ. 1, Εικόνα 1: http://www.birdsongs.it/songs/luscinia_megarhynchos/spectro2.html
30. Κεφ. 1, Εικόνα 2: http://www.birdsongs.it/songs/luscinia_megarhynchos/spectro3.html

Παραρτήματα

Παράρτημα 1

Πίνακας 1: Δεδομένα εκπαίδευσης - Γκιώνης

Eurasian Scops Owl (Otus Scops)				
Κωδικός Καταχώρησης	Αρχική* Διάρκεια (sec)	Είδος	Τοποθεσία	Ηχογραφήθηκε από τον/την
XC383067	5	alarm call	Italy	Guido Ceccolini
XC385242	40	alarm call	Italy	Marco Dragonetti
XC19665	18	call	Kazakhstan	David Farrow
XC211203	33	call	Greece	Veljo Runnel
XC235526	21	call	Portugal	Helder Cardoso
XC304922	29	call	Portugal	Paulo Belo
XC308948	16	call	Greece	Giannis Gasteratos
XC315631	16	call	Italy	Marco Mastrorilli
XC319632	19	call	Greece	Giannis Gasteratos
XC329833	38	call	Italy	Fabrizio Grieco
XC362940	29	call	Israel	Daniel Melamed
XC374680	20	call	Italy	Roberto Mucchiutti
XC418819	5	call	Algeria	Karim Haddad
XC436092	29	call	Spain	Robin Fifield
XC376717	16	call, song	France	BrieucD
XC368566	19	female, male, song	Bulgaria	Samuel Odrzykoski
XC188747	31	juvenile	France	Latour
XC177732	17	male, song	Germany	Norbert Uhlhaas
XC177734	21	male, song	Germany	Norbert Uhlhaas
XC314037	39	male, song	France	julien Rochefort
XC102920	23	song	Italy	Francesco Sottile
XC105170	19	song	Germany	Nils Agster
XC113416	32	song	Switzerland	Jerome Fischer
XC142694	26	song	France	Fernand DEROUSSEN
XC174448	27	song	Italy	Corrado Zanini
XC176933	38	song	Greece	Jerome Fischer
XC187973	29	song	Turkey	Lars Lachmann
XC240938	36	song	France	Cedric MROCZKO
XC29068	13	song	Kazakhstan	Jan Hein van Steenis
XC309456	24	song	Portugal	Alexandre Moreira
XC309885	40	song	Tunisia	Cedric Mroczko
XC31756	15	song	Spain	Juan Malo de Molina
XC320969	20	song	Italy	Dimitri
XC343142	19	song	Turkey	Tero Linjama
XC359029	34	song	Spain	Jerome Fischer
XC362428	24	song	Italy	Sauro Giannerini
XC363908	27	song	France	Stanislas Wroza
XC366246	62	song	Italy	Francesco Sottile
XC369915	24	song	Portugal	Alexandre Moreira
XC372378	38	song	Spain	Niels Van Doninck

XC377454	33	song	Ukraine	Stuart Fisher
XC392928	29	song	Spain	Jonathan Farooqi
XC407296	28	song	Italy	Marco Preziosi
XC412217	21	song	Italy	maudoc
XC42019	35	song	Germany	Matthias Feuersenger
XC428980	35	song	France	Étienne Leroy
XC436734	15	song	France	Thomas ARMAND
XC54736	25	song	France	manuel Grosselet
XC55191	14	song	France	GROSSELET Olivier
XC78930	33	song	France	Jacques PREVOST

Πίνακας 2: Δεδομένα εκπαίδευσης - Κούκος

Common Cuckoo (Cuculus Canorus)				
Κωδικός Καταχώρησης	Αρχική* Διάρκεια (sec)	Είδος	Τοποθεσία	Ηχογραφήθηκε από τον/την
XC246319	31	aberrant song, song	Germany	Frank Holzapfel
XC143258	40	call	Norway	Elias A. Ryberg
XC411984	24	call	Nepal	Aladdin
XC402141	5	call, female	India	Mike Dooher
XC323954	30	call, male	Hungary	Csaba Moskat
XC323955	30	call, male	Hungary	Csaba Moskat
XC380809	30	call, male	Hungary	Csaba Moskat
XC141442	13	call, song in flight	Spain	Jordi Calvet
XC134671	18	male, song	United Kingdom	Ashley Fisher
XC234675	23	male, song	Belarus	dmitry yakubovich
XC289716	39	male, song	Bulgaria	Nikolay Sariev
XC289720	33	male, song	Bulgaria	Nikolay Sariev
XC307854	16	male, song	Spain	Carlos W.
XC363797	23	male, song	Poland	Antoni Knychala
XC363858	18	male, song	Spain	Carlos Rossi
XC415288	28	male, song	Poland	Romuald Mikusek
XC100628	20	song	Slovakia	Rudo Jureček
XC101485	25	song	Norway	Ingar Bringsvor
XC101506	13	song	France	Jacques PREVOST
XC103075	21	song	Italy	Francesco Sottile
XC120462	39	song	Kazakhstan	Patrick Franke
XC130651	9	song	Italy	maudoc
XC174676	18	song	Germany	Michele Peron
XC179281	21	song	Estonia	Uku Paal
XC181230	15	song	Germany	Buhl Johannes
XC193760	18	song	Netherlands	Peter Lindenburg
XC239345	19	song	Hungary	Zsanett Rozendaal-Pandur
XC244396	11	song	Spain	Jordi Calvet
XC246317	28	song	Germany	Frank Holzapfel
XC247953	16	song	Russian Federation	Albert Lastukhin
XC281234	32	song	Belgium	Peter Boesman

XC285816	22	song	Japan	Peter Boesman
XC315571	21	song	Netherlands	Ad Postma
XC318234	37	song	France	Jérémy Simar
XC320045	38	song	Georgia	Franck HIDVEGI
XC343138	24	song	Poland	Tero Linjama
XC361057	10	song	France	Olivier SWIFT
XC37447	38	song	United Kingdom	David Farrow
XC375221	27	song	Russian Federation	Benjamin Herold
XC412686	16	song	France	Stanislas Wroza
XC416702	20	song	Portugal	Peter Boesman
XC420664	9	song	Russian Federation	Lars Edenius
XC424141	35	song	Montenegro	Ptice Forum
XC437553	30	song	France	Jack Berteau
XC53936	20	song	Russian Federation	Sergey Slyusarev
XC56032	20	song	Russian Federation	Sergey Slyusarev
XC75595	19	song	Czech Republic	Tomas Belka
XC78137	31	song	Germany	Volker Arnold
XC80746	39	song	Germany	Volker Arnold
XC87424	29	song	Norway	Stein Ø. Nilsen

Πίνακας 3: Δεδομένα εκπαίδευσης - Δεκαοχτούρα

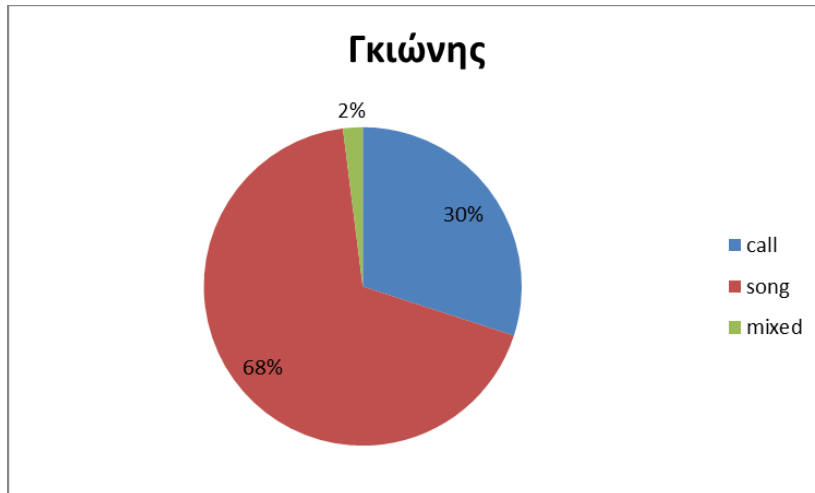
Eurasian Collared Dove (<i>Streptopelia decaocto</i>)				
Κωδικός Καταχώρησης	Αρχική* Διάρκεια (sec)	Είδος	Τοποθεσία	Ηχογραφήθηκε από τον/την
XC268063	11	call	United Kingdom	david m
XC324031	22	call	Hungary	Csaba Moskat
XC402031	40	call	United Kingdom	david m
XC240579	17	call, female	United Kingdom	david m
XC165577	22	call, male	United Kingdom	david m
XC118361	10	call, male	Italy	Marco Dragonetti
XC263584	28	call, male	Spain	José Carlos Sires
XC129508	27	flight call, song	Italy	Roberto Lerco
XC149328	26	flight call, song	India	Frank Lambert
XC283278	40	flight call, song	United Kingdom	david m
XC304302	33	flight call, song	United Kingdom	david m
XC184834	9	male, song	Uzbekistan	Albert Lastukhin
XC189692	17	male, song	United Kingdom	david m
XC364621	13	male, song	Germany	Matthias Hemprich
XC365473	10	male, song	France	Gerard OLIVIER
XC375708	14	male, song	United States	Jonathan Fosdick
XC380916	10	male, song	Morocco	José Carlos Sires
XC433624	12	male, song	Spain	Esperanza Poveda
XC101941	22	song	United Kingdom	Richard Dunn
XC112763	13	song	France	Jan-Kees Bossenbroek
XC11632	20	song	United States	Nathan Pieplow
XC141942	17	song	Greece	Stein Ø. Nilsen
XC149327	19	song	India	Frank Lambert

XC160085	17	song	Canada	Ian Cruickshank
XC165852	38	song	India	Frank Lambert
XC211916	12	song	Germany	Matthias Hemprich
XC236606	19	song	France	Cedric MROCZKO
XC243905	29	song	United Kingdom	david m
XC270553	36	song	United Kingdom	david m
XC292965	32	song	United Kingdom	david m
XC292975	35	song	United Kingdom	david m
XC292976	23	song	United Kingdom	david m
XC297326	16	song	United Kingdom	David Darrell-Lambert
XC297454	9	song	United States	Ross Gallardy
XC310660	20	song	United States	Sue Riffe
XC316367	18	song	Portugal	Cedric Mroczko
XC322661	20	song	Israel	Thomas Lüthi
XC343136	14	song	Poland	Tero Linjama
XC367406	14	song	Spain	Yoann Blanchon
XC370150	23	song	Montenegro	Uku Paal
XC37443	29	song	United Kingdom	David Farrow
XC377925	14	song	Denmark	Louis A. Hansen
XC377927	15	song	Denmark	Louis A. Hansen
XC381184	38	song	France	Olivier SWIFT
XC420573	35	song	Bosnia Herzegovina	Ptice Forum
XC424168	15	song	United Kingdom	Guy Kirwan
XC58766	15	song	France	Niels Krabbe
XC6539	12	song	United States	Darrell L. Peterson
XC77301	19	song	India	vir joshi
XC88219	10	song	Spain	Jordi Calvet

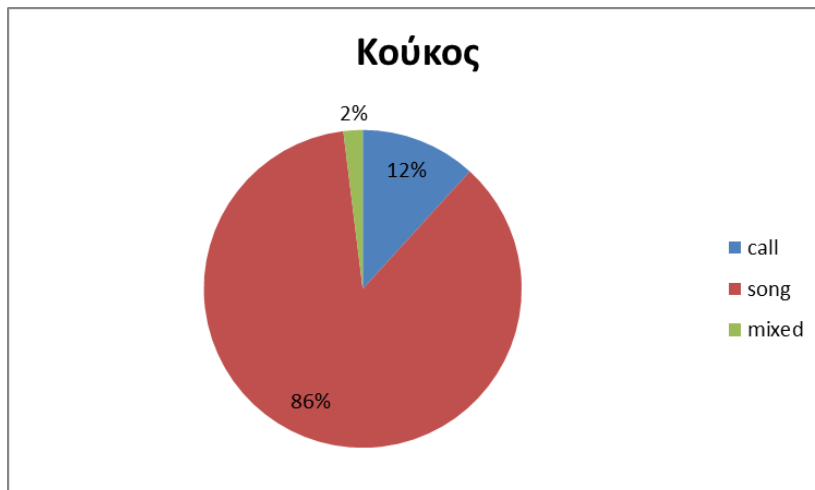
Διευκρίνιση: Όλα τα αρχεία ήχου συγκεντρώθηκαν από την διαδικτυακή πλατφόρμα xeno-canto.org, προστατεύονται από άδειες πνευματικών δικαιωμάτων Creative Commons (CC) και ανήκουν σε αυτούς που τα ηχογράφησαν και τα κοινοποίησαν στην συγκεκριμένη ιστοσελίδα. Οι νόμιμοι ιδιοκτήτες (ονομαστικώς) και οι κωδικοί καταχώρησης των αρχείων αναφέρονται στους Πίνακες του Παραρτήματος 1. Η ακριβής τοποθεσία (url) κάθε αρχείου στην ιστοσελίδα βρίσκεται προσθέτοντας τον αριθμό καταχώρησής του μετά από την διεύθυνση της αρχικής σελίδας www.xeno-canto.org/. Χρησιμοποιήθηκαν στην παρούσα έρευνα συμβαδίζοντας πλήρως με τους όρους χρήσης που έχει θέσει η ιστοσελίδα και οι δημιουργοί τους. Κανένα αρχείο ήχου ή παράγωγό του δεν διαμοιράστηκε/κοινοποιήθηκε ή θα διαμοιραστεί/κοινοποιηθεί στο μέλλον, σε οποιαδήποτε μορφή, για εμπορικό ή οποιονδήποτε άλλο σκοπό. Το αντικείμενο της έρευνας αυτής είναι αποκλειστικά η εξέταση των χαρακτηριστικών του ήχου και των τεχνικών της μηχανικής μάθησης στο birdsong, και μέσα σε αυτό το πλαίσιο δημιουργήθηκε η εφαρμογή birdMelos Vol.1 που χρησιμοποίησε τα συγκεκριμένα αρχεία.

*Κατά την προεπεξεργασία μειώθηκε η χρονική διάρκεια ορισμένων δεδομένων.

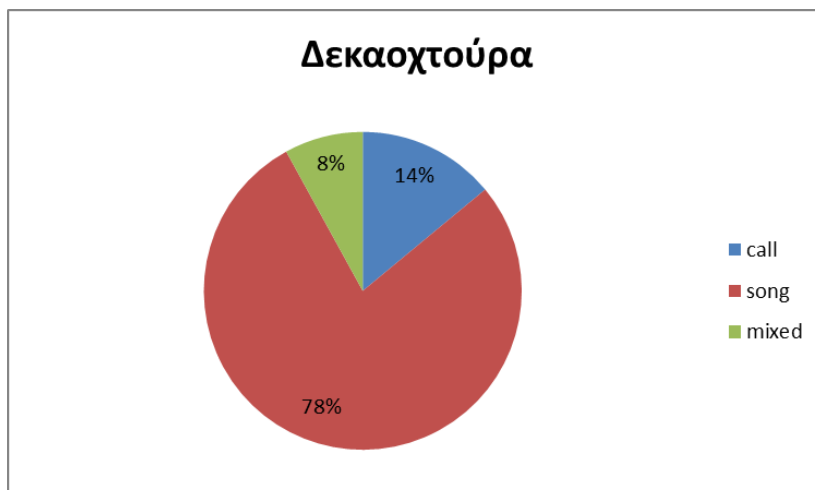
Εικόνα 1: Δεδομένα εκπαίδευσης – καλέσματα Γκιώνη



Εικόνα 2: Δεδομένα εκπαίδευσης – καλέσματα Κούκου



Εικόνα 3: Δεδομένα εκπαίδευσης – καλέσματα Δεκαοχτούρας



Παράρτημα 2

Πίνακας 1: Δεδομένα δοκιμής - Γκιώνης

Eurasian Scops Owl (Otus Scops)				
Κωδικός Καταχώρησης	Αρχική* Διάρκεια (sec)	Είδος	Τοποθεσία	Ηχογραφήθηκε από τον/την
XC318109	34	call	Kazakhstan	Jake Turin
XC319632	19	call	Greece	Giannis Gasteratos
XC345701	12	call	Serbia	alderash
XC207289	38	call, male	Romania	David Melichar
XC376717	16	call, song	France	BrieucD
XC411492	18	female, male, song	Bosnia Herzegovina	Aidan Place
XC136200	31	male, song	Italy	Gabriele Papale
XC183997	24	male, song	Kazakhstan	Albert Lastukhin
XC385246	27	male, song	Italy	Marco Dragonetti
XC240939	37	song	France	Cedric MROCZKO
XC252237	14	song	Austria	Miklos Heincz
XC315043	11	song	Spain	iSpiny
XC385832	14	song	France	Cedric Mroczko
XC388996	12	song	Macedonia**	Danuta Peplowska-Marczak
XC393848	34	song	Italy	Daniele Baroni
XC407296	28	song	Italy	Marco Preziosi
XC412452	28	song	Spain	Marc Anderson
XC416709	29	song	Finland	Matti Pajunen
XC420462	8	song	Greece	Stephan Risch
XC436736	26	song	France	Morvan Coentin
XC358166	27	song, territorial call	Bulgaria	Nikolay Sariev

Πίνακας 2: Δεδομένα δοκιμής - Κούκος

Common Cuckoo (Cuculus Canorus)				
Κωδικός Καταχώρησης	Αρχική* Διάρκεια (sec)	Είδος	Τοποθεσία	Ηχογραφήθηκε από τον/την
XC319117	24	call, song	Belgium	Alain Malengreau
XC323807	32	call, male	Hungary	Csaba Moskat
XC367571	10	call, song	Germany	Gerhard Huber
XC380974	29	call, male	Hungary	Csaba Moskat
XC394530	11	call, song	Netherlands	Frank Roos
XC374457	8	flight call, male, song	Spain	Esperanza Poveda
XC144003	9	male, song	France	julien Rochefort
XC185077	11	male, song	Uzbekistan	Albert Lastukhin
XC247641	12	male, song	Finland	Eetu Paljakka
XC307855	19	male, song	Spain	Carlos W.
XC176878	17	song	Germany	Sonnenburg
XC239540	35	song	Germany	Sonnenburg
XC239561	26	song	Italy	Roberto Lerco
XC242024	16	song	France	Yoann BLANCHON
XC281230	20	song	Belgium	Peter Boesman
XC342337	9	song	France	Timo Tschentscher
XC358976	9	song	Estonia	James Lidster
XC412685	9	song	France	Stanislas Wroza
XC414701	15	song	Spain	Stanislas Wroza
XC281236	15	song and call	Belgium	Peter Boesman
XC374165	25	song, stifled song	France	Stanislas Wroza

Πίνακας 3: Δεδομένα δοκιμής - Δεκαοχτούρα

Eurasian Collared Dove (Streptopelia decaocto)				
Κωδικός Καταχώρησης	Αρχική* Διάρκεια (sec)	Είδος	Τοποθεσία	Ηχογραφήθηκε από τον/την
XC391109	11	call	Italy	Vivek Puliyeri
XC331666	14	call, flight call	United Kingdom	david m
XC428769	19	call, song	Portugal	Jorge Leitão
XC374748	14	city sounds, song	France	Manuel Grosselet
XC292477	35	flight call, song	United Kingdom	david m
XC188717	17	male, song	United Kingdom	david m
XC254351	22	male, song	United Kingdom	david m
XC282490	17	male, song	Spain	José Carlos Sires
XC124694	12	song	Denmark	Louis A. Hansen
XC133855	8	song	United Kingdom	Mike Nelson
XC134431	14	song	France	Manuel Grosselet
XC184835	15	song	Uzbekistan	Albert Lastukhin
XC283277	14	song	United Kingdom	david m
XC363124	13	song	United Kingdom	GABRIEL LEITE

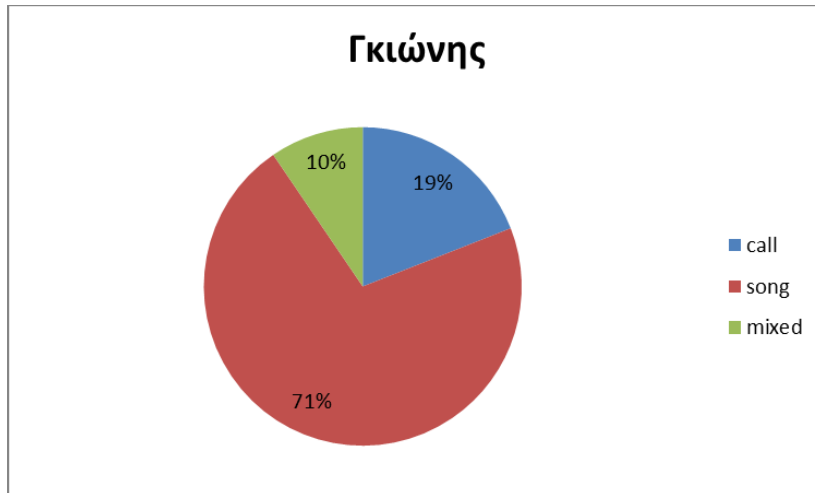
XC364870	11	song	Germany	Sonnenburg
XC365606	12	song	Denmark	Louis A. Hansen
XC370631	24	song	Italy	Dimitri
XC378443	13	song	France	Cedric Mroczko
XC383634	24	song	Morocco	Terje Kolaas
XC412791	15	song	Ireland	Harry Hussey
XC381882	15	song, wing beat	Mexico	Manuel Grosselet

Διευκρίνιση: Όλα τα αρχεία ήχου συγκεντρώθηκαν από την διαδικτυακή πλατφόρμα xeno-canto.org, προστατεύονται από άδειες πνευματικών δικαιωμάτων Creative Commons (CC) και ανήκουν σε αυτούς που τα ηχογράφησαν και τα κοινοποίησαν στην συγκεκριμένη ιστοσελίδα. Οι νόμιμοι ιδιοκτήτες (ονομαστικώς) και οι κωδικοί καταχώρησης των αρχείων αναφέρονται στους Πίνακες του Παραρτήματος 2. Η ακριβής τοποθεσία (url) κάθε αρχείου στην ιστοσελίδα βρίσκεται προσθέτοντας τον αριθμό καταχώρησής του μετά από την διεύθυνση της αρχικής σελίδας www.xeno-canto.org/. Χρησιμοποιήθηκαν στην παρούσα έρευνα συμβαδίζοντας πλήρως με τους όρους χρήσης που έχει θέσει η ιστοσελίδα και οι δημιουργοί τους. Κανένα αρχείο ήχου ή παράγωγό του δεν διαμοιράστηκε/κοινοποιήθηκε ή θα διαμοιραστεί/κοινοποιηθεί στο μέλλον, σε οποιαδήποτε μορφή, για εμπορικό ή οποιονδήποτε άλλο σκοπό. Το αντικείμενο της έρευνας αυτής είναι αποκλειστικά η εξέταση των χαρακτηριστικών του ήχου και των τεχνικών της μηχανικής μάθησης στο birdsong, και μέσα σε αυτό το πλαίσιο δημιουργήθηκε η εφαρμογή birdMelos Vol.1 που χρησιμοποίησε τα συγκεκριμένα αρχεία.

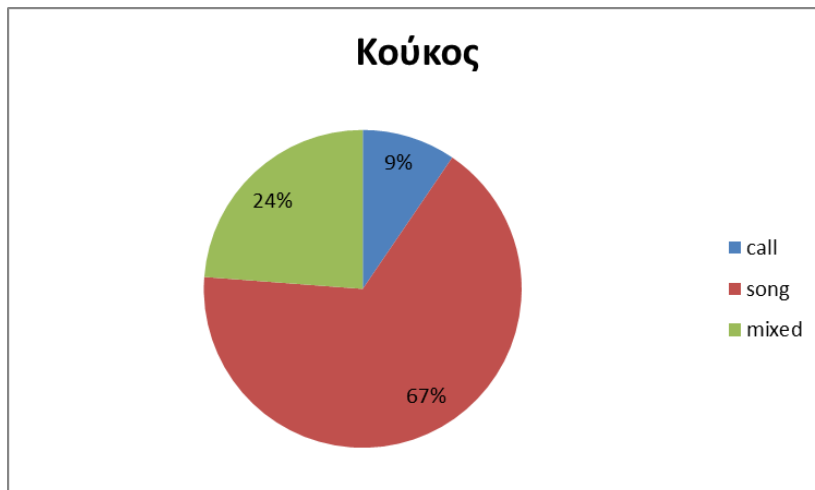
*Κατά την προεπεξεργασία μειώθηκε η χρονική διάρκεια ορισμένων δεδομένων.

**Ονομασία σύμφωνα με την καταχώρηση στην ιστοσελίδα www.xeno-canto.org.

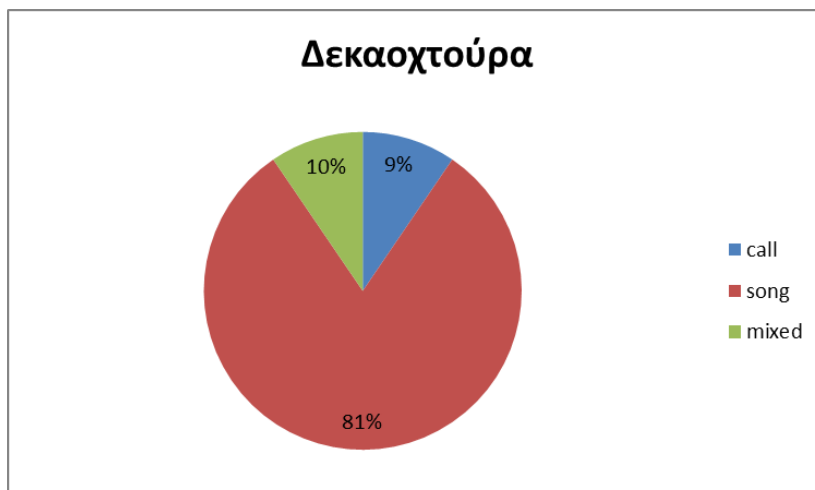
Εικόνα 1: Δεδομένα δοκιμής – καλέσματα Γκιώνη



Εικόνα 2: Δεδομένα δοκιμής – καλέσματα Κούκου



Εικόνα 3: Δεδομένα δοκιμής – καλέσματα Δεκαοχτούρας



Παράρτημα 3

Πίνακας 1: Εξαχθέντα χαρακτηριστικά δεδομένων εκπαίδευσης - Γκιώνης*

#	Bird 1	Bird 2	Bird 3	Bird 4	Bird 5	Bird 6	Bird 7	Bird 8	Bird 9	Bird 10
Feature 1	0,013749	0,086474	0,061837	0,022612	0,062706	0,027419	0,115976	0,006776	0,076603	0,062031
Feature 2	0,866667	0,899329	0,886241	0,874388	0,893781	0,90328	0,890385	0,922797	0,859672	0,85303
Feature 3	2948,247	10745,92	9804,304	12316,54	12884,27	12915,46	7562,135	3071,662	11486,27	11200,8
Feature 4	1469,01	1715,212	3386,811	1350,662	1719,292	1872,986	1405,678	1311,212	1479,609	5590,768
Feature 5	0,146674	0,325188	0,346928	0,084757	0,335395	0,233123	0,105199	0,112893	0,141101	0,332509
Feature 6	2222,854	1856,969	2660,319	1769,515	1611,462	2298,524	1868,427	2467,662	1526,394	2924,696
Feature 7	3680,125	1909,164	3673,042	2841,438	1532,432	3090,377	2940,712	4353,788	1965,123	3611,105
Feature 8	4,005586	5,606432	3,010992	5,916962	10,64961	4,116851	5,549881	3,442533	8,294658	2,50196
Feature 9	0,093824	0,049793	0,165558	0,018059	0,007444	0,088626	0,047321	0,093458	0,011193	0,168196
Feature 10	0,688195	0,746146	0,788688	0,666378	0,726439	0,743687	0,67536	0,700669	0,718334	0,784718
Feature 11	3,133099	0,773694	2,042933	2,230568	1,004273	1,44955	2,280649	3,561218	5,303982	0,584272
Feature 12	-2,34332	-2,91884	-2,23305	-5,82686	-8,76787	-2,81384	-3,21433	-1,13806	-3,58867	-1,72389
Feature 13	-0,81097	-0,50256	-0,81737	-1,40585	-0,38105	0,07952	-1,0743	-0,68051	-2,67338	-2,12505
Feature 14	1,239621	1,781894	0,525611	2,013346	1,093417	2,571247	1,281892	0,359187	-0,82103	-0,36865
Feature 15	1,700405	1,455116	0,431477	3,893453	2,765752	0,785266	2,048035	0,695924	2,605256	0,552215
Feature 16	-0,03866	-0,62943	0,14515	3,061689	1,392776	0,203195	0,062691	0,264363	-0,58851	0,659209
Feature 17	-1,16474	-1,54277	-0,61799	-1,32897	0,082579	-2,33325	-1,95301	-0,99488	-0,8991	-0,96329
Feature 18	-0,37652	-0,00272	-0,48706	0,583737	0,485852	-0,47146	-0,86772	-1,12773	0,074233	-0,58291
Feature 19	0,943263	1,01708	0,54522	1,014856	0,559392	0,612494	1,411372	0,462588	1,24521	0,630505
Feature 20	0,684136	0,117008	0,458092	-0,00264	0,323825	0,214727	0,696583	1,136882	0,371422	0,287667
Feature 21	-0,03008	-0,4909	-0,09934	0,938111	-0,68441	0,713207	-0,7305	-0,0197	-0,7474	0,229962
Feature 22	-0,16929	-0,06423	-0,17984	0,602142	1,389043	-0,59236	-0,0722	-0,84417	0,846962	-0,08564
Feature 23	-0,37377	-0,53372	-0,21622	0,305497	-0,02163	0,487886	-0,40333	-0,39949	0,602568	0,019824
Feature 24	1,886604	0,685408	1,797926	1,863218	1,750553	1,905385	0,950253	1,908979	1,77667	1,862697
Feature 25	NaN	0,031814	0,03839	NaN	NaN	0,072	0,026396	0,025193	NaN	0,07619
Feature 26	NaN	0,943475	0,899513	NaN	NaN	0,297287	0,74731	0,546058	NaN	0,401859
Feature 27	NaN	29,65593	23,43091	NaN	NaN	4,128986	28,31166	21,6752	NaN	5,274402
#	Bird 11	Bird 12	Bird 13	Bird 14	Bird 15	Bird 16	Bird 17	Bird 18	Bird 19	Bird 20
Feature 1	0,08635	0,029577	0,028979	0,029053	0,006552	0,013895	0,010388	0,027258	0,020033	0,007465
Feature 2	0,841855	0,910714	0,888532	0,895252	0,878929	0,862186	0,830782	0,832822	0,895812	0,896185
Feature 3	10632,55	13304,01	11574,44	12481,98	4243,458	11112,7	11584,61	11759,04	13070,39	12677,98
Feature 4	1704,446	1373,667	5141,914	2307,817	1757,858	1726,483	1303,432	1597,705	1588,41	1352,26
Feature 5	0,244527	0,097875	0,942709	0,504188	0,21505	0,261462	0,060985	0,19222	0,200726	0,093622
Feature 6	2016,87	2613,407	3850,791	2161,644	2867,925	1970,5	2249,359	1665,102	1689,338	2701,505
Feature 7	2514,327	4281,617	2691,371	2448,345	4844,436	2683,235	4054,398	2050,654	2341,436	4724,744
Feature 8	4,294575	3,065352	2,972443	5,004994	3,029962	5,891617	3,922859	8,2541	7,018715	3,187042
Feature 9	0,084411	0,031711	0,145094	0,08379	0,083091	0,047586	0,027174	0,008886	0,009693	0,050793
Feature 10	0,74334	0,67899	0,863631	0,783218	0,697385	0,753619	0,659929	0,70478	0,716035	0,659005
Feature 11	2,097269	7,479599	-1,93019	1,22433	1,321447	1,198816	2,832964	1,396602	1,001191	2,671548
Feature 12	-1,97219	0,21736	-2,0137	-3,02028	-5,26469	-4,48964	-2,58761	-9,62965	-6,99059	-2,11182
Feature 13	-1,00184	-1,53532	0,136575	-0,31382	-1,29534	-1,37398	-3,24795	-2,25373	-3,58446	-2,44217
Feature 14	0,821732	2,986199	-0,40011	1,15114	-0,3683	0,17103	1,828842	3,495596	2,843567	-0,77321
Feature 15	1,049853	2,212934	1,337796	0,709068	0,484085	0,048018	3,166445	0,66563	4,468096	1,492006
Feature 16	0,069039	-0,65585	-1,05357	-0,37085	0,113769	-0,9278	1,566406	2,206154	-0,64147	-1,04294
Feature 17	-1,44821	-4,06131	-0,34743	-0,58434	-1,68854	-1,13124	-2,30261	-0,96853	-1,72968	-1,83166
Feature 18	-0,84142	-1,8385	0,96286	0,494528	-0,03155	0,653808	-1,9371	-1,66743	1,037357	-0,73432
Feature 19	1,270997	1,285942	-0,07146	0,656505	0,859277	1,392917	0,397754	0,994954	-0,36225	0,720616
Feature 20	0,625473	1,79858	-0,68783	-0,70782	0,923471	0,93893	1,536942	0,089553	0,144453	2,655107
Feature 21	-0,57316	0,436378	0,38148	-0,51109	0,453223	0,967775	0,447511	-0,29746	1,164402	0,054415
Feature 22	-0,21923	-0,70012	-0,73958	-0,26466	-0,22759	0,505801	-0,93675	1,131896	0,946721	-0,76092
Feature 23	-0,10205	0,088522	-0,01065	-0,33372	-0,18704	0,691684	-0,73691	-0,13345	-1,06893	-0,24431
Feature 24	1,737349	0,969245	1,661584	1,605066	0,958558	1,783096	0	0,890962	0,772775	0,978328
Feature 25	0,039592	0,037188	0,069592	0,071383	0,031791	0,022812	NaN	0,019206	0,04322	NaN
Feature 26	0,958994	0,403395	0,566737	0,725717	0,545844	0,82193	NaN	0,691304	0,653287	NaN
Feature 27	24,22202	10,8474	8,143729	10,16649	17,16956	36,03092	NaN	35,99351	15,11539	NaN

#	Bird 21	Bird 22	Bird 23	Bird 24	Bird 25	Bird 26	Bird 27	Bird 28	Bird 29	Bird 30
Feature 1	0,067734	0,052293	0,109484	0,025054	0,019787	0,141345	0,104023	0,01089	0,001995	0,047617
Feature 2	0,886292	0,858896	0,893179	0,860666	0,906291	0,858065	0,898762	0,900609	0,906408	0,882917
Feature 3	9062,707	10634,97	12940,66	12397	1393,522	12463,41	12694,47	5156,262	1663,576	10670,59
Feature 4	1666,321	1644,935	2520,058	1652,169	1668,634	1320,419	1336,215	1653,157	6859,734	4854,875
Feature 5	0,253314	0,234677	0,199621	0,329366	0,262851	0,067311	0,053166	0,280757	0,379263	0,463736
Feature 6	1914,824	1763,379	2033,686	1735,198	1810,973	1438,706	1392,559	2014,233	3650,246	2936,831
Feature 7	2159,947	2071,761	2297,717	2243,979	2259,116	1396,736	1313,933	3185,278	5438,321	3354,433
Feature 8	5,690236	6,558287	4,654102	7,605461	6,669621	12,59285	11,36075	5,074571	2,213674	2,745164
Feature 9	0,054482	0,055274	0,040209	0,01381	0,049061	0,015544	0,017672	0,048603	0,284226	0,173366
Feature 10	0,74184	0,715898	0,719682	0,683066	0,743895	0,622345	0,701835	0,731919	0,865985	0,821477
Feature 11	-0,01897	2,543931	0,572614	0,996586	1,977889	2,780689	-0,78319	2,456191	2,78496	-0,41059
Feature 12	-4,88398	-2,33796	-2,68156	-7,99489	-2,84209	-3,77762	-6,20528	-3,78071	-1,35223	-3,16641
Feature 13	-1,2857	-1,27747	-1,24593	-1,72119	-0,0491	-1,70918	-3,33854	-0,56166	-0,6099	-1,49021
Feature 14	1,151844	0,972545	0,69951	1,195101	2,44892	1,010077	-0,69907	1,768037	0,118161	0,02181
Feature 15	1,847133	0,903659	2,17561	1,118618	2,114826	1,034916	1,479996	1,868128	0,495284	0,633622
Feature 16	0,370623	-1,11521	-0,54482	-0,35453	-1,10702	-0,0149	1,433091	-0,02485	0,003042	-0,70403
Feature 17	-0,18866	-0,65669	-1,35555	-2,06362	-1,58736	-1,02796	-0,69416	-1,82875	-0,18182	-0,77641
Feature 18	1,153066	0,585143	-0,46714	0,326403	0,235106	-0,69593	-1,56891	-0,31624	-0,32772	-0,41691
Feature 19	1,273683	1,500864	0,891382	0,518243	1,206007	0,599865	0,565288	1,146061	0,291246	0,256388
Feature 20	-0,17323	0,294746	1,638171	-0,13906	-0,05448	0,450808	1,206336	0,194635	0,356035	0,247768
Feature 21	-0,1616	-0,05541	-1,42172	-0,37017	-0,5072	0,026855	0,094698	-0,39858	-0,1171	-0,28428
Feature 22	-0,0272	0,115422	0,463164	0,103987	0,372111	-0,54031	0,06414	0,091198	-0,22234	-0,16562
Feature 23	0,302951	0,126874	-0,84593	0,276029	-0,7044	-0,22481	0,053766	-0,23834	-0,16443	-0,34591
Feature 24	1,44067	1,857408	1,873845	0,795327	0,882727	0,979612	1,934261	1,744096	1,484222	1,881657
Feature 25	0,020998	0,02941	0,020385	NaN	NaN	0,03	NaN	NaN	1,145986	0,034195
Feature 26	0,660191	0,43925	0,359137	NaN	NaN	0,819453	NaN	NaN	0,327679	0,47116
Feature 27	31,44104	14,93519	17,61727	NaN	NaN	27,31509	NaN	NaN	0,285936	13,77862
#	Bird 31	Bird 32	Bird 33	Bird 34	Bird 35	Bird 36	Bird 37	Bird 38	Bird 39	Bird 40
Feature 1	0,110668	0,060513	0,031764	0,020219	0,004257	0,003559	0,022901	0,012042	0,036641	0,002334
Feature 2	0,811713	0,869281	0,891019	0,882702	0,857678	0,903427	0,928197	0,76699	0,921472	0,878173
Feature 3	11514,72	11064,61	13045,54	13341,51	11462,77	12441,22	14811,06	8288,503	13983,98	8395,539
Feature 4	2256,444	1796,382	1466,298	1598,122	2037,075	2880,108	1623,848	2930,871	3221,42	9175,247
Feature 5	0,273331	0,291797	0,136616	0,172929	0,407498	0,240105	0,235717	0,618829	0,865765	0,329342
Feature 6	2158,619	1767,409	1680,77	1846,959	2623,762	3601,196	1865,424	2684,489	2616,091	4006,549
Feature 7	2705,648	2037,137	2312,2	2984,22	4186,305	5926,126	2622,842	2373,522	1676,865	5490,752
Feature 8	4,006254	6,810732	7,22843	6,030409	3,633777	2,21323	7,054968	5,250668	7,442716	1,953306
Feature 9	0,090139	0,02866	0,012329	0,025491	0,059437	0,127337	0,01162	0,049106	0,017173	0,243485
Feature 10	0,768017	0,7307	0,672288	0,714871	0,753068	0,776624	0,703317	0,77336	0,81956	0,835636
Feature 11	0,820476	1,907287	0,706595	0,315531	0,857681	4,256177	-0,32331	-0,02957	-2,11806	2,150251
Feature 12	-3,61231	-4,17313	-9,149	-6,46519	-5,1806	-2,91234	-8,9336	-2,09933	-6,16593	-1,6016
Feature 13	-2,24261	-0,6566	-1,34882	-2,18266	-1,18657	-1,19606	0,511391	0,046913	0,762216	-0,84742
Feature 14	-0,1683	-0,74912	3,351326	-0,7555	1,08493	0,214616	3,643165	-0,64356	-1,42683	1,405248
Feature 15	0,127242	0,096969	0,209123	0,361101	2,132496	0,052434	0,318727	1,328089	1,819261	1,413747
Feature 16	-0,29768	-1,01895	2,020767	0,638666	0,167706	0,445351	0,306325	0,874526	0,354179	0,936073
Feature 17	-1,13068	-2,07386	-1,4265	-0,3996	-0,38215	-0,18964	-2,80883	-1,81736	0,059989	-0,6787
Feature 18	-0,57561	-0,29418	-2,10776	0,035001	-0,7809	-0,79935	0,472526	-1,37763	-0,82054	-0,90385
Feature 19	0,204575	-0,40553	1,178244	0,351988	1,096186	0,565195	1,84758	1,475188	1,446061	0,293135
Feature 20	-0,25697	0,169785	0,900816	1,300858	-1,44493	0,899265	-0,658	0,075244	-1,13274	0,851203
Feature 21	-0,30182	-0,68341	-0,05183	0,232039	-0,51646	0,138667	-0,13621	-0,48733	0,609622	-0,59715
Feature 22	-0,49703	-0,78738	0,181664	-0,41941	0,409822	0,173097	-0,08363	0,443935	-0,42279	-0,04968
Feature 23	-0,3627	-0,28104	-0,61231	-0,68161	-0,79209	-0,47293	-0,81447	-0,42623	0,736982	-0,2709
Feature 24	1,734282	1,862147	0,976364	1,890642	1,563897	1,804376	0,833439	0,982735	1,490487	1,432007
Feature 25	0,033605	NaN	NaN	NaN	NaN	0,15839	NaN	NaN	NaN	NaN
Feature 26	0,683479	NaN	NaN	NaN	NaN	0,665336	NaN	NaN	NaN	NaN
Feature 27	20,33834	NaN	NaN	NaN	NaN	4,200617	NaN	NaN	NaN	NaN

#	Bird 41	Bird 42	Bird 43	Bird 44	Bird 45	Bird 46	Bird 47	Bird 48	Bird 49	Bird 50
Feature 1	0,028444	0,071472	0,016212	0,005593	0,040716	0,021848	0,063907	0,058066	0,033907	0,023109
Feature 2	0,894454	0,898568	0,919598	0,896067	0,898234	0,863934	0,861886	0,908184	0,883202	0,898801
Feature 3	13949,99	13596,91	14316,48	12718,07	12381,13	12002,05	11288,46	10548,1	12738,96	12088,07
Feature 4	1970,055	1789,352	1420,349	1572,786	1622,708	1704,151	1648,215	1549,465	1731,825	1409,92
Feature 5	0,294317	0,356935	0,107179	0,189507	0,177904	0,238651	0,250417	0,187834	0,360006	0,105282
Feature 6	2443,71	2101,173	2475,957	2580,656	2069,859	1807,911	1719,504	1571,023	1933,558	1911,923
Feature 7	3577,661	2745,979	4530,99	4419,232	2718,91	2726,384	1899,004	1816,716	2332,971	2696,268
Feature 8	3,774201	4,839979	3,599152	3,465115	4,541597	5,651097	6,576445	7,235826	5,5008	5,504188
Feature 9	0,11994	0,081385	0,017431	0,040316	0,075864	0,057734	0,047336	0,038307	0,04258	0,033094
Feature 10	0,757126	0,743415	0,668178	0,721392	0,758527	0,739632	0,731758	0,74805	0,746502	0,704186
Feature 11	1,133553	-0,04507	2,895092	3,160325	3,042046	2,953907	1,87668	2,60752	-1,84698	0,123155
Feature 12	-2,96579	-4,66604	-6,45123	-3,45609	-2,71708	-2,85024	-3,21576	-4,32621	-6,8902	-4,30505
Feature 13	-1,56219	-1,83348	-5,55302	-0,51612	-1,31498	-0,2588	-1,38102	-1,45909	-3,97754	-4,20705
Feature 14	0,233193	0,793006	2,32479	2,100822	2,363373	0,688338	1,20031	1,474937	1,642565	0,691533
Feature 15	0,560909	1,102124	3,562641	2,575528	-0,39467	1,62696	1,162666	1,097717	1,207426	1,558211
Feature 16	-1,02966	0,065064	0,874629	-0,12295	0,803837	0,95053	0,678534	0,416133	0,23899	0,801097
Feature 17	-1,42507	0,04652	-2,17125	-3,97947	-1,81392	0,005645	-0,41161	-0,26916	0,531786	0,789649
Feature 18	-0,45524	1,168362	-0,41056	-0,64764	-0,59554	-0,13334	-1,03478	0,489322	0,081111	0,81743
Feature 19	0,43986	1,876332	0,124731	1,636855	1,97371	1,025036	0,631365	1,517749	0,81007	1,102247
Feature 20	-0,21264	1,180162	0,119045	-0,34965	-0,93667	0,35973	-0,11125	0,464418	-0,69136	2,100127
Feature 21	-1,02596	0,840503	0,309246	-0,17073	0,724096	-0,24653	-0,90565	-0,66433	-0,74645	0,043655
Feature 22	-0,78178	0,609417	0,500809	0,950925	-1,04065	0,126111	-0,25755	-0,18097	0,098223	0,26217
Feature 23	-0,44372	0,051755	-0,07223	-1,91994	-0,18094	-0,30218	-0,34316	-0,28762	-0,53644	-0,78951
Feature 24	1,719539	1,790043	0,628662	1,814892	1,129333	1,862442	1,82136	1,070698	1,46885	1,791584
Feature 25	0,021604	0,029417	0,070208	NaN	0,018594	0,013197	NaN	0,025193	0,027596	NaN
Feature 26	0,669304	0,871586	0,949257	NaN	0,718953	0,463842	NaN	0,186332	0,940616	NaN
Feature 27	30,98032	29,629	13,52058	NaN	38,66564	35,14681	NaN	7,39626	34,08478	NaN

Πίνακας 2: Εξαχθέντα χαρακτηριστικά δεδομένων εκπαίδευσης - Κούκος*

#	Bird 1	Bird 2	Bird 3	Bird 4	Bird 5	Bird 6	Bird 7	Bird 8	Bird 9	Bird 10
Feature 1	0,047494	0,08238	0,126225	0,081248	0,021937	0,095621	0,035886	0,072905	0,073633	0,115997
Feature 2	0,755172	0,766962	0,764423	0,805764	0,829365	0,832258	0,868286	0,797092	0,783934	0,75367
Feature 3	6207,162	6815,783	1429,715	5441,673	3957,639	3172,942	2726,845	2954,571	1279,916	4845,343
Feature 4	689,9457	652,7252	657,2674	691,1654	613,6542	749,5405	668,4546	847,1549	656,7627	796,0977
Feature 5	0,016485	0,014494	0,005133	0,009523	0,04141	0,006541	0,017148	0,007046	0,008702	0,006296
Feature 6	834,2633	731,3327	676,3561	735,9036	1167,823	769,7464	865,8775	770,8273	731,8388	754,6831
Feature 7	2075,787	1501,996	1198,863	1418,462	3231,942	1259,188	2224,722	1413,537	1518,086	1154,282
Feature 8	8,597151	11,52636	15,13375	12,70724	5,362082	14,32599	8,064223	12,77825	11,95267	15,2908
Feature 9	0,01373	0,013211	0,00388	0,006876	0,021188	0,004524	0,009492	0,004534	0,005373	0,006586
Feature 10	0,638927	0,667654	0,635799	0,646312	0,653335	0,675995	0,661041	0,688524	0,622964	0,673712
Feature 11	7,761036	7,497013	10,18191	8,223228	8,06622	8,533668	10,14305	10,96499	9,311768	8,072705
Feature 12	1,592323	0,679883	1,627247	0,202455	-0,01932	0,118785	1,7611	1,719513	1,056416	1,635645
Feature 13	-0,47078	-0,97884	-1,88334	0,003803	-1,04185	-3,33673	-2,47481	-3,06304	-1,54992	-1,83358
Feature 14	-1,24866	-0,91297	-1,66488	-1,82349	0,687346	-2,30781	-0,06886	-1,43901	-1,22369	-2,58216
Feature 15	-0,63919	-0,49622	-0,50329	-0,7121	-1,25033	-1,44309	-1,13612	-1,05825	-0,25709	-0,95591
Feature 16	-0,61169	-0,80521	0,066306	-0,54568	-1,42548	-1,04556	-1,02848	-0,58493	-0,82946	-0,95251
Feature 17	-0,99338	-0,8355	-0,93321	-1,43924	-2,36071	-1,46395	-1,24737	-0,95868	-1,26539	-0,3134
Feature 18	-0,98614	-0,84985	-1,76049	-0,8797	-0,75303	-0,60268	-0,35927	0,597192	-0,97208	-0,36738
Feature 19	-0,06025	0,009203	-0,03207	0,196755	-0,18888	0,465812	0,164473	0,244411	0,373659	0,518569
Feature 20	0,926211	0,437658	0,219958	0,909001	-0,74902	0,170267	-0,29396	0,274694	0,124293	0,829892
Feature 21	1,176537	0,357366	0,096098	1,006938	0,526835	0,857605	0,848356	-0,20074	0,631433	0,412647
Feature 22	0,514842	0,760683	0,935538	0,434809	0,378695	0,698981	0,503055	0,386183	0,749855	-0,14993
Feature 23	0,553748	0,487346	0,639578	0,423266	0,416207	-0,47645	0,626254	0,126117	0,601357	-0,22524
Feature 24	1,682659	0,93546	1,169967	0,443786	1,107199	1,245715	1,425207	1,774122	0,668138	1,388183
Feature 25	NaN	0,0678	NaN	NaN	0,056984	0,074399	0,070794	NaN	0,057007	0,050385
Feature 26	NaN	0,946557	NaN	NaN	0,806159	0,510045	0,417214	NaN	0,762365	0,966138
Feature 27	NaN	13,96093	NaN	NaN	14,14707	6,855525	5,893388	NaN	13,37322	19,17492

#	Bird 11	Bird 12	Bird 13	Bird 14	Bird 15	Bird 16	Bird 17	Bird 18	Bird 19	Bird 20
Feature 1	0,032983	0,037641	0,067312	0,009734	0,100602	0,062548	0,048949	0,090932	0,017899	0,050002
Feature 2	0,830443	0,864303	0,771784	0,80427	0,712996	0,786151	0,775899	0,717532	0,761745	0,693582
Feature 3	1869,464	8025,752	6241,087	3974,124	4369,24	3126,105	5289,143	5027,94	4466,535	1876,9
Feature 4	638,6361	657,5828	623,7059	663,7862	659,0338	693,0159	661,5572	663,4918	818,6823	697,0954
Feature 5	0,022237	0,013623	0,010822	0,064507	0,018428	0,014153	0,019373	0,014221	0,035665	0,011722
Feature 6	853,147	816,9875	721,3622	1503,99	719,8383	761,6628	817,4939	746,385	1115,394	792,2191
Feature 7	2183,213	2033,341	1739,468	3968,916	1406,156	1615,053	1876,58	1470,791	2945,449	1770,409
Feature 8	8,073803	8,893323	10,4346	4,200636	12,00438	11,13163	9,08739	12,20242	5,922256	10,17786
Feature 9	0,020383	0,007834	0,008835	0,034801	0,012126	0,00781	0,019533	0,006791	0,019204	0,009214
Feature 10	0,675669	0,669638	0,633435	0,684382	0,674905	0,67698	0,666918	0,625735	0,690025	0,674249
Feature 11	7,075527	9,372586	8,608981	8,107407	7,579976	9,542195	7,014941	6,234758	9,335279	8,9237
Feature 12	0,851411	2,23591	1,615747	0,063546	0,105951	-0,84914	0,824443	-0,84281	-0,81334	0,754379
Feature 13	-0,79522	-1,70039	-0,90531	-0,68528	-0,23091	-1,01122	-1,01004	-1,58892	-1,79726	-1,48123
Feature 14	-0,77643	-1,90684	-0,81135	0,857642	-0,02922	-0,24124	-0,84264	-2,3183	-0,03143	-1,37001
Feature 15	-0,55346	-1,19529	-0,13498	-1,28926	-1,03322	-1,56587	-0,38096	-1,11398	-2,20909	-0,58795
Feature 16	-1,26186	-0,26557	-0,49888	-1,47346	-0,58599	0,055953	-1,11816	-2,53638	-0,87934	-0,83128
Feature 17	-1,49648	-2,25909	-1,2791	-1,24933	-1,56935	-1,52044	-1,07839	-2,04115	-1,23118	-0,96644
Feature 18	-0,92573	-0,98424	-1,24303	-0,44595	-1,14945	-0,35678	-0,73261	-0,95186	-0,69643	-0,58842
Feature 19	-0,23942	-0,24996	-0,60331	-0,56124	0,137004	0,165162	-0,21467	-0,6657	-0,18793	-0,00947
Feature 20	0,037013	0,024446	0,149401	0,386592	0,205309	0,179733	0,390303	0,231594	-0,12485	0,642189
Feature 21	0,148699	0,364952	0,618664	0,624719	0,400368	0,765385	0,590806	-0,22049	-0,54113	0,79318
Feature 22	0,267056	0,715846	0,782894	-0,03762	0,443464	-0,15624	0,678181	0,702581	-0,45137	0,610374
Feature 23	0,371891	0,607466	0,650636	0,62614	-0,15132	0,797071	0,619357	-0,19074	0,085769	0,216156
Feature 24	1,200151	0,990635	1,150069	1,117707	1,125455	1,315202	1,134515	1,067154	1,545644	0,99212
Feature 25	NaN	0,040816	NaN	NaN	NaN	NaN	NaN	0,047415	NaN	0,09839
Feature 26	NaN	0,455015	NaN	NaN	NaN	NaN	NaN	0,770613	NaN	0,842615
Feature 27	NaN	11,14786	NaN	NaN	NaN	NaN	NaN	16,25253	NaN	8,564031
#	Bird 21	Bird 22	Bird 23	Bird 24	Bird 25	Bird 26	Bird 27	Bird 28	Bird 29	Bird 30
Feature 1	0,032982	0,074514	0,040843	0,029321	0,085346	0,094149	0,026699	0,020643	0,028268	0,084585
Feature 2	0,726703	0,800546	0,802885	0,771715	0,796738	0,795524	0,857798	0,846814	0,790698	0,792
Feature 3	5137,569	7126,08	3651,317	5411,506	7330,913	6915,596	3832,083	10458,5	4491,161	7804,647
Feature 4	712,1728	697,8104	632,622	889,5484	694,3617	640,9492	718,3763	664,375	666,4568	728,2597
Feature 5	0,020275	0,007896	0,022139	0,018359	0,006991	0,007133	0,027416	0,033585	0,01957	0,008712
Feature 6	847,6726	751,7417	903,2196	1034,268	705,2858	657,1455	1004,081	1100,63	929,6639	748,4644
Feature 7	1977,701	1382,009	2508,525	2125,579	1288,724	1228,534	2429,952	2981,57	2426,934	1373,173
Feature 8	8,842282	13,08894	7,107619	8,336099	14,00055	14,66026	7,018664	5,857794	7,348311	12,98025
Feature 9	0,019763	0,006126	0,013738	0,012923	0,00626	0,004932	0,02587	0,020119	0,010682	0,005516
Feature 10	0,70017	0,648112	0,639869	0,687811	0,6841	0,68782	0,662315	0,683717	0,668766	0,683379
Feature 11	7,120461	8,719297	8,211724	8,422913	8,651392	9,069871	6,289533	8,846798	9,664827	10,84611
Feature 12	1,172081	0,051716	1,926553	-0,2708	1,411713	0,738453	2,599559	-0,07286	1,417932	-0,70032
Feature 13	-0,49629	-2,67564	-1,02748	-2,12747	-0,73709	-0,62796	0,467069	-0,92897	-2,86189	-1,85647
Feature 14	-0,96462	-1,87662	-1,56459	-2,26214	-1,62674	-1,29686	-1,69397	-0,38126	-0,84842	-0,27851
Feature 15	-0,83352	-1,08094	-0,3693	-0,44161	-0,85735	-0,68988	-2,08907	-1,10287	-1,0552	-1,18021
Feature 16	-1,21272	-1,37166	-1,20698	-0,26572	-0,78889	-0,08802	-1,5808	-0,4063	-0,60745	0,024541
Feature 17	-0,93716	-1,81632	-1,98868	-0,10275	-0,89403	-1,47141	-0,43889	-1,8298	-1,85762	-1,17056
Feature 18	-0,33048	-0,80679	-0,58946	0,731277	-0,90851	-1,01845	0,539679	-0,54876	-1,01015	-0,31494
Feature 19	0,215721	-0,28141	-0,96976	-0,15417	0,136734	-0,64862	0,637963	-0,3054	-0,48024	0,53426
Feature 20	0,559981	-0,13585	-0,39968	0,320753	0,591357	-0,02483	0,436661	-0,10373	0,050503	0,415186
Feature 21	0,609066	0,24421	0,198057	0,179024	0,384058	0,360223	0,244504	0,440202	0,933018	0,938784
Feature 22	0,296383	0,205909	-0,05469	-0,29642	0,633522	-0,23671	0,370087	0,377908	0,760868	0,751309
Feature 23	0,017679	-0,14427	0,417104	-0,13466	0,253039	0,86975	0,494705	0,21068	0,462519	-0,12658
Feature 24	0,636964	0,954103	1,458851	0,937943	0,762768	1,202004	0,758643	1,26363	0,573729	1,12723
Feature 25	NaN	NaN	0,069002	NaN	NaN	NaN	0,057596	NaN	NaN	NaN
Feature 26	NaN	NaN	0,30061	NaN	NaN	NaN	0,699039	NaN	NaN	NaN
Feature 27	NaN	NaN	4,356524	NaN	NaN	NaN	12,13685	NaN	NaN	NaN

#	Bird 31	Bird 32	Bird 33	Bird 34	Bird 35	Bird 36	Bird 37	Bird 38	Bird 39	Bird 40
Feature 1	0,081229	0,089391	0,176087	0,053356	0,124693	0,012557	0,04877	0,117579	0,01162	0,046387
Feature 2	0,825061	0,845823	0,747197	0,7375	0,788565	0,787263	0,812005	0,807947	0,552846	0,77577
Feature 3	2725,228	9916,569	3386,814	936,8652	2298,739	900,9611	1506,574	2528,232	2083,47	789,9794
Feature 4	634,2622	710,827	694,2145	654,2393	682,5437	636,5333	834,4116	636,0917	2180,742	712,2359
Feature 5	0,008712	0,007897	0,008277	0,011226	0,00496	0,041072	0,014035	0,00874	0,882541	0,017892
Feature 6	694,7883	722,781	671,9882	765,3361	663,5453	1259,461	873,6528	658,3412	2174,169	882,9966
Feature 7	1420,256	1284,566	1235,626	1835,383	1050,905	3501,641	1673,491	1298,52	2158,799	2048,381
Feature 8	12,71327	14,03455	13,98641	9,849935	17,18966	4,891816	10,70571	13,98343	7,493239	8,698247
Feature 9	0,006392	0,005397	0,009361	0,006556	0,004121	0,019856	0,007447	0,00544	0,020171	0,009103
Feature 10	0,678483	0,687434	0,665888	0,6326	0,667239	0,640675	0,671895	0,654303	0,724444	0,672784
Feature 11	8,901225	8,705114	7,773074	10,66433	8,874431	8,122033	9,109911	7,760314	4,433875	10,46341
Feature 12	1,660457	0,957699	1,499732	1,987634	1,625334	2,662521	-1,41405	1,100478	-5,73729	-0,58722
Feature 13	-1,62858	-1,346	-1,59196	-2,27817	-1,32748	-1,55398	-1,66795	0,450995	0,680437	-2,89194
Feature 14	-1,11265	-1,69974	-1,44098	-0,75	-1,46836	-0,86973	-1,34062	-1,94107	2,172885	0,15577
Feature 15	0,200488	-0,17095	-0,43611	-1,39774	-0,97134	-1,55093	-2,43245	-0,03226	-0,84971	-0,89237
Feature 16	-1,10934	-0,49248	-0,78085	-0,70319	-0,39472	-0,87485	-0,28357	-0,3101	-0,54932	-0,8965
Feature 17	-1,1854	-1,85248	-1,43959	-1,31995	-1,17428	-2,23405	-0,46532	-1,86309	0,713014	-0,83805
Feature 18	-0,66688	-0,07728	-1,33806	-0,73096	-0,89591	-1,35461	-0,37825	-1,05129	-0,01347	0,110098
Feature 19	-0,45525	0,042884	-0,27086	-0,69401	-0,04954	-0,22106	0,525724	-1,01528	-0,71876	-0,49019
Feature 20	0,426214	-0,12748	0,20667	-0,40668	0,313151	0,368835	0,30403	0,232868	1,068376	0,687977
Feature 21	0,005629	1,22063	0,145636	0,204033	0,511362	0,620369	-0,08693	0,392333	0,477368	1,154193
Feature 22	0,255125	0,194524	0,074829	0,26355	0,587146	0,464776	-0,04095	0,68375	-0,31264	0,440247
Feature 23	0,319777	0,62167	0,172148	0,242048	-0,21816	0,791287	-0,5054	1,00711	0,180663	0,726744
Feature 24	1,023914	1,477562	1,691569	1,281871	-1,204895	1,274843	1,607816	0,419511	0,447616	1,031369
Feature 25	0,025193	NaN	0,070816	NaN	0,053401	0,054014	0,06059	NaN	0,0378	0,103787
Feature 26	0,386601	NaN	0,440754	NaN	0,343994	0,347505	0,706588	NaN	0,243015	0,711275
Feature 27	15,34572	NaN	6,223908	NaN	6,441679	6,433658	11,66188	NaN	6,428885	6,85323
#	Bird 41	Bird 42	Bird 43	Bird 44	Bird 45	Bird 46	Bird 47	Bird 48	Bird 49	Bird 50
Feature 1	0,079105	0,028644	0,067577	0,082441	0,077734	0,024992	0,104527	0,050129	0,05782	0,049329
Feature 2	0,804154	0,784506	0,791667	0,677116	0,819201	0,599666	0,802057	0,83953	0,759067	0,79702
Feature 3	2088,028	4046,573	1955,88	1782,099	1836,987	1844,227	1339,138	2069,652	962,9493	5328,564
Feature 4	650,2018	707,841	711,563	677,1183	721,5305	692,385	656,4473	694,1724	689,5041	672,6182
Feature 5	0,009525	0,021251	0,009534	0,008081	0,00999	0,019295	0,004766	0,009612	0,064187	0,04987
Feature 6	730,617	979,8319	794,8671	736,8065	766,171	943,275	670,2637	771,7894	1121,653	974,7347
Feature 7	1535,288	2375,656	1680,202	1510,281	1473,412	2367,704	1136,303	1708,667	2498,045	2339,493
Feature 8	11,79541	7,485734	10,79587	12,06336	12,2396	7,52266	15,83252	10,59376	5,812215	6,861117
Feature 9	0,006509	0,010257	0,005759	0,004925	0,006955	0,012525	0,004768	0,005781	0,060263	0,021915
Feature 10	0,637735	0,676506	0,629273	0,620331	0,672523	0,69213	0,668007	0,689734	0,704385	0,658273
Feature 11	7,906289	9,112215	9,680364	9,456286	8,272642	9,011656	9,479075	11,01972	4,698341	5,896595
Feature 12	2,043229	-0,47819	1,5942	1,836986	0,27481	1,358655	2,142526	1,64781	1,340608	0,434416
Feature 13	-1,18437	-4,31813	-1,92584	-2,1432	-1,33255	-1,98103	-1,92401	-3,25036	-0,32424	0,488913
Feature 14	-2,63569	-1,01767	-1,68651	-0,67102	-1,8199	-1,45263	-1,51295	-0,74566	-1,34963	0,753501
Feature 15	-0,12417	-1,52552	-1,16029	-1,65922	-0,85582	-1,00563	-0,07933	-0,68874	-1,05366	-2,42788
Feature 16	-0,18516	-2,01284	-0,79021	-0,44534	-0,95672	-0,66038	-0,4387	-0,31745	-0,80967	-2,02283
Feature 17	-1,94211	-2,7949	-1,37733	-1,73638	-0,76339	-1,19504	-1,49108	-1,53955	-0,58165	-1,16617
Feature 18	-1,29798	-0,72204	-0,56807	-0,8089	-1,14327	-0,79814	-1,35596	-1,0004	-0,55968	0,073799
Feature 19	0,411214	-0,00794	-0,48304	-0,21255	0,482176	0,131424	-0,25124	-0,07818	-0,23262	-0,05851
Feature 20	-0,03385	-0,25236	0,531199	0,059201	0,127207	0,212314	0,508388	0,350353	0,195238	0,51125
Feature 21	0,036541	0,34619	0,967153	1,400751	0,616015	0,840965	0,574442	1,017524	0,604617	0,606309
Feature 22	0,928749	0,562386	0,456435	0,393667	0,259641	0,375131	0,455764	0,350944	0,620608	0,877349
Feature 23	0,556377	0,252116	0,028465	-0,15888	0,046698	0,22655	0,377741	0,453531	0,415323	0,99641
Feature 24	0,92993	0,941785	0,779044	0,569568	1,005257	0,55688	1,261139	1,509473	1,376469	1,401105
Feature 25	NaN	NaN	0,054603	0,061202	NaN	0,067188	1,531202	NaN	NaN	0,039002
Feature 26	NaN	NaN	0,876779	0,833754	NaN	0,473872	0,328936	NaN	NaN	0,339647
Feature 27	NaN	NaN	16,05729	13,62303	NaN	7,052898	0,214822	NaN	NaN	8,708404

Πίνακας 3: Εξαχθέντα χαρακτηριστικά δεδομένων εκπαίδευσης - Δεκαοχτούρα*

#	Bird 1	Bird 2	Bird 3	Bird 4	Bird 5	Bird 6	Bird 7	Bird 8	Bird 9	Bird 10
Feature 1	0,019496	0,025884	0,019273	0,050568	0,006324	0,018513	0,069034	0,125439	0,042631	0,056984
Feature 2	0,674779	0,524104	0,569712	0,660066	0,532468	0,714286	0,573222	0,77551	0,568678	0,535
Feature 3	814,5694	1363,002	1386,111	6463,14	4898,928	10440,58	4281,668	9190,099	2294,586	3045,019
Feature 4	515,2408	680,3988	630,8556	602,5512	664,5012	4759,847	625,1358	558,3913	697,0954	705,2965
Feature 5	0,030424	0,012145	0,024127	0,01021	0,060506	0,782772	0,005449	0,005754	0,009595	0,006682
Feature 6	998,0281	785,7673	943,2101	703,1428	1619,251	3141,372	646,8274	613,7264	754,7709	722,2018
Feature 7	3010,853	1931,444	2670,279	1766,789	4202,601	2277,829	1284,33	1133,842	1589,147	1405,762
Feature 8	5,840695	9,345167	6,637963	10,21898	3,916711	3,698243	14,12555	15,96226	11,39115	12,91236
Feature 9	0,016042	0,006573	0,013738	0,005766	0,029955	0,041972	0,003335	0,003185	0,005967	0,003956
Feature 10	0,635641	0,692308	0,673578	0,654406	0,70457	0,868642	0,680403	0,645721	0,679624	0,657915
Feature 11	8,511168	10,84332	10,17213	11,11845	8,574406	-1,13065	12,13155	11,59486	10,35186	11,28256
Feature 12	1,901412	3,6078	1,769485	2,265155	2,286171	-1,38476	2,113496	-0,38793	-0,11031	1,410565
Feature 13	-0,4215	-2,10897	-1,7082	-2,36212	-2,14311	0,186759	-2,99153	-1,57807	-1,48792	-3,1234
Feature 14	0,899365	-1,69742	0,366706	-0,89024	-0,94562	-0,19532	-0,31352	0,716229	-0,21758	-1,11424
Feature 15	-0,90081	-1,54659	-0,86648	-1,19895	-1,01873	-0,111	-0,02645	-1,51123	-1,858	-1,26543
Feature 16	0,138098	-0,03551	-0,60927	0,397364	-0,2843	-1,33412	-0,033	-0,14558	-0,7512	-0,74726
Feature 17	-0,70104	-1,33736	-1,39953	-0,41954	-2,02342	-1,38938	-1,79635	-1,08408	-1,59014	-1,7632
Feature 18	-1,29717	-1,02188	-0,62269	-0,98604	-1,37735	-0,92868	-1,0239	-1,12621	-0,88446	-0,52422
Feature 19	-1,64865	-0,32515	-0,35695	-1,56798	-0,13718	0,267298	-0,48389	-0,40323	-0,17116	-0,11031
Feature 20	-1,06968	0,192175	-0,06559	-0,80922	0,576964	0,296483	-0,37696	-0,25632	0,542309	-0,00775
Feature 21	-0,49592	0,395188	0,491253	0,251599	0,646433	0,889627	0,189474	-0,08139	0,63925	0,651132
Feature 22	-0,44258	-0,02451	0,358446	-0,20523	0,111961	0,638232	0,425634	0,195825	0,333773	0,592983
Feature 23	0,394754	0,12122	0,433454	-0,27132	0,348775	-0,29677	0,588306	0,320196	0,441919	0,292553
Feature 24	0,670679	0,923689	1,896597	1,86357	0,943398	0,598365	0,932548	1,918747	0,762829	1,233106
Feature 25	0,095397	0,079206	NaN	0,107392	NaN	NaN	NaN	NaN	0,148186	0,148798
Feature 26	0,970749	0,488885	NaN	0,974953	NaN	NaN	NaN	NaN	0,95732	0,676776
Feature 27	10,17591	6,172299	NaN	9,078424	NaN	NaN	NaN	NaN	6,460262	4,548283
#	Bird 11	Bird 12	Bird 13	Bird 14	Bird 15	Bird 16	Bird 17	Bird 18	Bird 19	Bird 20
Feature 1	0,013473	0,054111	0,08364	0,031615	0,1102	0,097663	0,098421	0,006432	0,006035	0,024705
Feature 2	0,810198	0,601227	0,786509	0,723529	0,542029	0,610619	0,619355	0,83631	0,926706	0,75
Feature 3	2036,493	709,8889	2100,222	722,8923	528,9155	969,8648	1398,799	7003,594	7082,44	7670,02
Feature 4	525,3345	663,2395	625,1148	587,0321	626,0611	588,7985	577,2749	789,4526	676,151	567,8962
Feature 5	0,04986	0,009137	0,006715	0,016526	0,00469	0,004592	0,022675	0,086862	0,064146	0,030692
Feature 6	1349,925	703,8884	702,6694	793,1346	608,2026	601,8956	667,1011	1874,319	1625,436	955,9732
Feature 7	3854,518	1601,744	1427,394	2221,499	1171,076	1147,165	1386,24	4694,528	4295,761	2832,29
Feature 8	4,384868	11,31333	12,72849	8,101453	15,48219	15,97904	12,83922	3,386215	3,817628	6,180445
Feature 9	0,023821	0,005362	0,00437	0,009535	0,00294	0,002662	0,006488	0,058426	0,03317	0,024355
Feature 10	0,642103	0,671832	0,673726	0,632287	0,639257	0,626926	0,641686	0,710988	0,712098	0,658147
Feature 11	7,748244	11,29865	11,10584	10,19138	12,01835	10,61875	9,405571	6,764711	9,065921	5,810098
Feature 12	2,798398	1,065881	2,148855	1,672458	1,348673	1,862884	-1,91712	0,283242	1,808633	2,847795
Feature 13	-1,06672	-2,6665	-2,61235	-1,63751	-2,5799	-0,74601	0,81687	-0,66835	-2,1158	1,943688
Feature 14	-0,21455	-0,09111	-1,20686	0,315203	-0,11487	1,103298	0,044968	-0,29792	-0,16275	0,654237
Feature 15	-0,53778	-1,06334	-0,38994	-0,77472	-0,74331	-1,25182	-1,32673	-0,48128	-0,55765	-0,23474
Feature 16	0,356719	-0,92526	-0,21121	-0,28083	-0,2961	-1,6045	-0,05995	-0,71729	-0,10357	-0,8719
Feature 17	-1,45781	-1,44242	-1,60903	-1,44833	-1,2971	-2,09278	-0,96435	-0,90545	-1,33215	-1,06668
Feature 18	-1,71548	-0,5073	-0,99118	-1,0543	-0,96907	-1,22638	-0,46977	-0,78797	-0,75168	-0,92125
Feature 19	-1,40594	-0,62534	-0,2098	-0,61247	-1,08721	-0,80909	-0,90555	-1,07011	-0,81921	-0,50551
Feature 20	-0,76386	-0,52999	0,332967	-0,39898	-0,74193	-0,26455	-0,24938	-0,66041	-1,05763	0,008631
Feature 21	-0,09635	0,286985	0,251975	-0,07651	0,21288	0,119939	0,017978	0,124688	-0,34652	0,473819
Feature 22	-0,06195	0,361353	0,297962	0,157157	0,330837	0,121711	0,275042	0,726123	0,243807	0,776667
Feature 23	0,354278	0,29907	0,466823	0,529799	0,220142	0,450583	0,435058	0,512126	0,52361	0,866089
Feature 24	0,957566	1,62363	0,97721	1,739551	1,751552	0,976884	1,925461	1,138588	1,783597	0,683053
Feature 25	NaN	0,085805	0,167982	0,107392	NaN	0,040181	0,125397	0,026984	0,0522	NaN
Feature 26	NaN	0,426444	0,636599	0,971755	NaN	0,447447	0,323662	0,29701	0,377891	NaN
Feature 27	NaN	4,969918	3,78969	9,04865	NaN	11,13568	2,581101	11,00684	7,239353	NaN

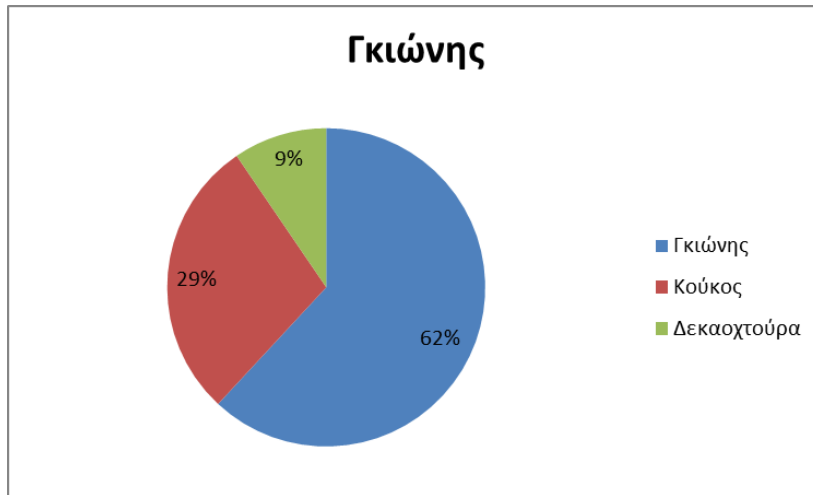
#	Bird 21	Bird 22	Bird 23	Bird 24	Bird 25	Bird 26	Bird 27	Bird 28	Bird 29	Bird 30
Feature 1	0,004471	0,028044	0,05785	0,036713	0,023176	0,02471	0,050423	0,044912	0,028337	0,010853
Feature 2	0,817391	0,55611	0,58209	0,592503	0,635522	0,637681	0,641115	0,555866	0,683089	0,664589
Feature 3	4267,314	1553,391	1749,533	2045,824	2759,279	2344,086	1695,045	1353,294	4363,375	2734,797
Feature 4	1085,324	506,1144	499,8058	509,2266	525,9022	529,5402	585,476	595,3594	543,8816	559,2745
Feature 5	0,077843	0,026748	0,013266	0,01999	0,028627	0,031087	0,011495	0,008988	0,026359	0,041218
Feature 6	1926,749	936,7947	703,5666	820,2478	974,2422	1020,1	701,3353	670,8204	918,0527	1195,391
Feature 7	4528,114	2789,939	1980,621	2412,256	2863,32	2990,067	1827,944	1671,275	2773,214	3493,477
Feature 8	3,52116	6,347128	9,158068	7,429002	6,160283	5,868881	9,954305	10,87276	6,392339	4,924362
Feature 9	0,042084	0,014542	0,007461	0,011082	0,015549	0,016775	0,006836	0,004908	0,014522	0,021977
Feature 10	0,750782	0,600763	0,572242	0,61942	0,65581	0,6252	0,648729	0,655207	0,619209	0,685964
Feature 11	8,996677	9,494495	9,823895	9,904819	9,709236	9,564909	10,02356	11,09187	9,394978	9,159425
Feature 12	-0,3989	1,338475	1,484667	1,502535	1,29296	1,356703	1,206754	4,097292	1,414713	2,146501
Feature 13	-3,41259	-1,66183	-1,30264	-1,3027	-1,82757	-1,6714	-1,06127	-1,81297	-1,06744	-1,20602
Feature 14	-0,4517	0,726942	1,00601	0,96279	0,579106	0,667722	0,383934	-1,69005	1,036675	0,535006
Feature 15	-1,33744	-0,60252	-0,69998	-0,59881	-0,64185	-0,75028	-1,12958	-1,3233	-0,81579	-0,68864
Feature 16	-0,7061	-0,51431	-0,66316	-0,22777	-0,38864	-0,52234	-0,5837	0,630543	-0,56735	-0,46506
Feature 17	-0,93911	-1,18047	-1,38838	-0,82328	-1,0764	-0,96692	-1,50364	-0,82358	-1,19941	-1,42064
Feature 18	-0,41039	-0,53511	-0,98124	-0,55608	-0,48808	-0,28555	-1,09692	-1,3478	-0,75093	-0,82126
Feature 19	-0,80081	-0,92685	-1,00867	-0,9813	-0,78512	-0,78626	-0,99268	-1,25145	-0,88174	-0,80339
Feature 20	-0,37103	-1,13113	-0,71476	-0,98483	-0,97339	-1,06752	-0,65554	-0,44432	-0,7496	-0,63412
Feature 21	0,585699	-0,51148	-0,30681	-0,37098	-0,39815	-0,39954	-0,06702	0,39218	-0,11603	0,037523
Feature 22	0,205252	-0,14287	-0,41166	-0,1862	-0,03235	-0,01187	0,182156	0,212125	0,0558	0,040662
Feature 23	-0,15302	0,418257	0,190626	0,274731	0,55063	0,572535	0,370694	0,18171	0,375291	0,135496
Feature 24	1,085135	1,95421	0	0,958576	0,937963	1,85902	1,365008	1,3624	1,878167	1,115947
Feature 25	0,075601	NaN	NaN	NaN	0,145782	0,094195	0,053379	0,052812	0,131995	NaN
Feature 26	0,257795	NaN	NaN	NaN	0,262796	0,851428	0,225115	0,208444	0,840607	NaN
Feature 27	3,409945	NaN	NaN	NaN	1,802662	9,038997	4,217314	3,94693	6,368452	NaN
#	Bird 31	Bird 32	Bird 33	Bird 34	Bird 35	Bird 36	Bird 37	Bird 38	Bird 39	Bird 40
Feature 1	0,124908	0,0495	0,08621	0,072919	0,110882	0,055371	0,018233	0,024044	0,019294	0,026159
Feature 2	0,579845	0,678175	0,694444	0,821053	0,490775	0,513889	0,573394	0,778169	0,571174	0,486865
Feature 3	2830,296	1743,382	577,1127	3278,036	753,2141	1345,04	2279,508	6405,376	1274,513	698,9316
Feature 4	609,7009	585,8966	535,9749	565,2045	637,837	607,4718	566,4242	525,3765	766,9521	633,8416
Feature 5	0,004414	0,01788	0,010854	0,00913	0,003185	0,008314	0,029975	0,026278	0,023686	0,017232
Feature 6	575,65	777,8897	637,4926	682,2251	594,0406	669,132	998,6864	930,3434	886,1722	799,6155
Feature 7	1167,016	2039,975	1393,726	1589,894	997,7017	1495,965	3002,424	2792,543	2540,959	2084,234
Feature 8	15,59096	8,750314	12,98811	11,43169	18,00659	12,13895	5,85497	6,3389	6,981523	8,605293
Feature 9	0,00271	0,013081	0,005242	0,006335	0,001932	0,005044	0,016437	0,014603	0,01729	0,013064
Feature 10	0,650637	0,647535	0,602213	0,640096	0,66396	0,656074	0,610725	0,644677	0,686074	0,68887
Feature 11	11,99854	7,2925	10,38661	9,760165	12,34977	10,89588	7,80339	9,1448	8,552863	8,340908
Feature 12	2,420372	1,993036	-0,76521	0,805362	3,036491	0,629958	3,625749	2,496192	1,385258	1,188997
Feature 13	-1,96561	-0,48166	-1,10562	-0,58853	-2,77234	-1,73077	0,562324	0,230937	-0,91111	-0,42438
Feature 14	-0,55438	-0,63193	0,775271	-0,86176	-1,1962	0,317969	-0,44633	0,157938	-0,93735	-1,10074
Feature 15	-1,07213	-0,40024	-1,4283	-0,67094	-0,91485	-0,97965	-1,51865	-2,23948	-0,37784	0,01424
Feature 16	0,045176	-1,27181	-0,62098	-0,0275	-0,23865	-0,84123	-0,99884	-0,2162	-0,32138	-1,39928
Feature 17	-1,31328	-1,0148	-0,30293	-1,5688	-1,77119	-1,72608	-1,32949	-0,34317	-0,26661	-1,51074
Feature 18	-1,08382	-1,06365	-1,09022	-0,87938	-0,98614	-0,70742	-1,09524	-0,67035	-0,39322	-0,27924
Feature 19	-0,5032	-0,09541	-1,35447	-0,80697	-0,22502	-0,61557	-1,31789	-0,49111	-0,99547	-0,83267
Feature 20	-0,30863	-0,30076	-0,38425	-0,18668	-0,07287	-0,11301	-0,99826	0,010855	-1,0723	-0,07026
Feature 21	-0,20103	0,340979	-0,32542	-0,20611	0,143796	0,437037	-0,59175	-0,06269	-0,74676	0,140857
Feature 22	-0,2519	0,067799	-0,16484	0,143368	0,15461	-0,08652	-0,05239	-0,36364	-0,16447	0,283799
Feature 23	0,161272	0,690166	0,172556	0,631513	0,434341	0,164063	0,248077	-0,02064	0,001869	0,273064
Feature 24	1,734167	1,77636	1,939417	0,972545	1,153392	1,879635	1,625274	1,926202	1,678744	0,593786
Feature 25	0,053401	0,069002	0,058186	0,091791	NaN	NaN	NaN	0,088821	0,447007	0,033605
Feature 26	0,611911	0,321329	0,668855	0,389821	NaN	NaN	NaN	0,406692	0,37838	0,728261
Feature 27	11,45872	4,656786	11,49513	4,246812	NaN	NaN	NaN	4,578792	0,846474	21,67092

#	Bird 41	Bird 42	Bird 43	Bird 44	Bird 45	Bird 46	Bird 47	Bird 48	Bird 49	Bird 50
Feature 1	0,024225	0,052069	0,068036	0,009154	0,18735	0,098187	0,015357	0,025054	0,124776	0,059164
Feature 2	0,492913	0,630996	0,762092	0,751707	0,574513	0,5472	0,787149	0,617048	0,611529	0,527344
Feature 3	591,4578	477,9632	1163,292	10135,71	771,8493	676,8974	6880,412	5226,406	1308,838	1340,273
Feature 4	634,9771	513,18	551,445	710,0279	524,0936	585,5601	574,0365	629,0892	596,8945	618,9114
Feature 5	0,016376	0,010435	0,01194	0,047364	0,003203	0,006701	0,04336	0,017406	0,003947	0,007157
Feature 6	808,8055	663,0445	736,7364	1285,213	537,9255	609,3946	1284,982	836,1164	602,4554	665,4859
Feature 7	2143,44	1746,05	1959,668	3552,183	1074,258	1280,838	3567,949	2221,978	1091,762	1428,233
Feature 8	8,381095	10,40953	10,08835	4,803674	18,62735	14,16247	4,785001	8,012905	16,70966	12,76014
Feature 9	0,009903	0,005989	0,006566	0,023571	0,001827	0,003896	0,023544	0,013684	0,002398	0,004428
Feature 10	0,692084	0,608795	0,643291	0,715359	0,640046	0,668771	0,641003	0,679257	0,665244	0,667426
Feature 11	10,43603	10,81793	10,63369	9,52447	11,90174	11,7042	9,225165	8,378744	12,126	10,42648
Feature 12	0,808383	1,577944	0,239517	0,64491	1,618459	0,286823	1,289258	1,385839	2,243307	0,550093
Feature 13	-1,44005	-2,11902	-1,80817	-2,72892	-2,43758	-1,66495	-2,27127	-2,03689	-2,86906	-1,45894
Feature 14	0,345615	0,488288	0,205028	0,158943	-0,22639	0,916956	0,024475	-2,16805	-0,32769	-0,63019
Feature 15	-1,47564	-0,17427	-0,67734	-0,66829	-0,22543	-0,98131	-0,63317	-1,03552	-0,23035	-2,27439
Feature 16	-0,43137	-0,26594	-0,61122	-0,75059	-0,42486	-0,30714	-0,2365	-0,49112	-0,31941	-0,45679
Feature 17	-1,13566	-1,78707	-1,45019	-1,63647	-2,05339	-1,09103	-1,27921	-1,20037	-2,06426	-1,65826
Feature 18	-0,86229	-1,00817	-0,6823	-0,69583	-1,33612	-0,87334	-0,81472	-1,38714	-1,11712	-1,54567
Feature 19	-0,48935	-0,36132	-1,32427	-0,81554	-1,5948	-0,54914	-0,68721	-1,01663	-0,49172	-0,77621
Feature 20	-0,23986	-0,41337	-0,66583	-1,16881	-1,02712	-0,38848	-0,2476	-0,28175	-0,47843	-0,19589
Feature 21	0,195279	-0,39903	0,046531	-0,53368	-0,42838	-0,12742	0,449401	-0,15115	0,035238	0,247378
Feature 22	0,315441	-0,25563	0,141419	0,002375	-0,14838	0,099556	0,240567	-0,06329	0,344245	-0,20137
Feature 23	0,419612	0,262658	0,342911	0,384598	0,022542	0,413724	0,384404	0,077888	0,541435	-0,03707
Feature 24	0,449008	1,920927	0,953694	0,80209	1,715129	1,485782	1,907221	1,910262	1,666877	0,447445
Feature 25	0,110408	0,146395	0,144	0,041995	0,057583	NaN	NaN	0,03839	0,1278	0,040204
Feature 26	0,790923	0,236819	0,880295	0,329968	0,767774	NaN	NaN	0,27139	0,690774	0,64473
Feature 27	7,163624	1,617678	6,113162	7,857221	13,33326	NaN	NaN	7,069294	5,405097	16,03643

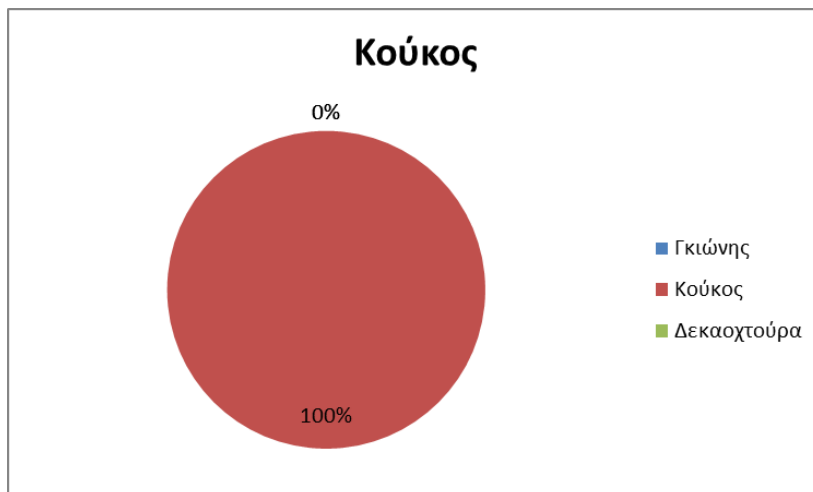
*Η ταξινόμηση των πουλιών στους πίνακες των εξαχθέντων χαρακτηριστικών (βλ. Παράρτημα 3) δεν αντιστοιχεί στην ταξινόμηση των πουλιών στους πίνακες των δεδομένων εκπαίδευσης (βλ. Παράρτημα 1). Στο Παράρτημα 3 η ταξινόμηση έγινε με τη σειρά που εισήχθησαν τα δεδομένα στο σύστημα, δηλαδή με βάση την ονομασία του εκάστοτε αρχείου ήχου. Στο Παράρτημα 1 η ταξινόμηση αρχικώς έγινε αλφαβητικά, βάσει του είδους του καλέσματος, και εν συνεχεία αριθμητικά, ως προς τον κωδικό καταχώρησης στην πλατφόρμα του xeno-canto.

Παράρτημα 4

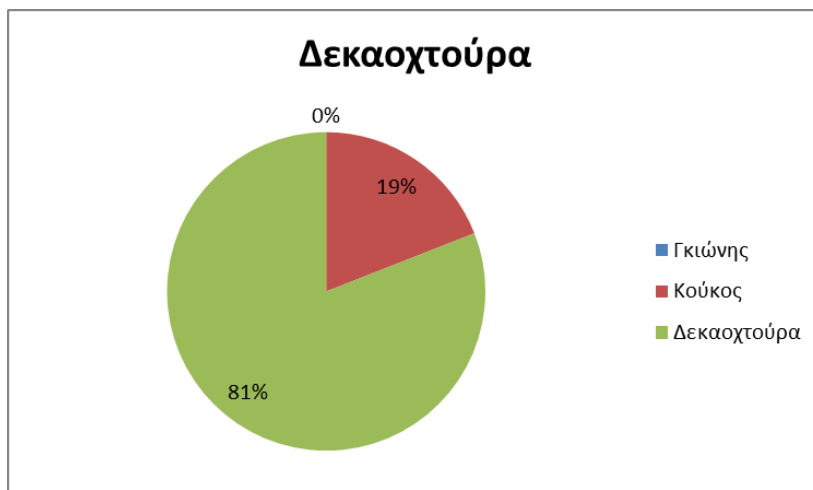
Εικόνα 1: Αναγνωρίσεις όταν τα δεδομένα δοκιμής ήταν Γκιώνη



Εικόνα 2: Αναγνωρίσεις όταν τα δεδομένα δοκιμής ήταν Κούκος



Εικόνα 3: Αναγνωρίσεις όταν τα δεδομένα δοκιμής ήταν Δεκαοχτούρας



Παράρτημα 5

Εικόνα 1: Classification Learner – Εισαγωγή δεδομένων

birdsdata																	
VarName11	VarName12	VarName13	VarName14	VarName15	VarName16	VarName17	VarName18	VarName19	VarName20	VarName21	VarName22	VarName23	VarName24	NaN	NaN1	NaN2	otus
Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Categorical
92	9.1122	-0.4782	-4.3181	-1.0177	-1.5255	-2.0128	-2.7949	-0.7220	-0.0079	-0.2524	0.3462	0.5624	0.2521	0.9418	NaN	NaN	cuckoo
93	9.6804	1.5942	-1.9258	-1.6865	-1.1603	-0.7902	-1.3773	-0.5681	-0.4830	0.5312	0.9672	0.4564	0.0285	0.7790	0.0546	0.8768	16.0573 cuckoo
94	9.4563	1.8370	-2.1432	-0.6710	-1.6592	-0.4453	-1.7364	-0.8089	-0.2125	0.0592	1.4008	0.3937	-0.1589	0.5696	0.0612	0.8338	13.6230 cuckoo
95	8.2726	0.2748	-1.3326	-1.8199	-0.8558	-0.9567	-0.7634	-1.1433	0.4822	0.1272	0.1610	0.2596	0.0467	1.0053	NaN	NaN	cuckoo
96	9.0117	1.3587	-1.9810	-1.4526	-1.0056	-0.6604	-1.1950	-0.7981	0.1314	0.2123	0.8410	0.3751	0.2265	0.5569	0.0672	0.4739	7.0529 cuckoo
97	9.4791	2.1425	-1.9240	-1.5129	-0.0793	-0.4387	-1.4911	-1.3560	-0.2512	0.5084	0.5744	0.4558	0.3777	1.2611	1.5312	0.3289	0.2148 cuckoo
98	11.0197	1.6478	-3.2504	-0.7457	-0.6887	-0.3174	-1.5395	-1.0004	-0.0782	0.3504	1.0175	0.3509	0.4535	1.5095	NaN	NaN	cuckoo
99	4.6983	1.3406	-0.3242	-1.3496	-1.0537	-0.8097	-0.5816	-0.5597	-0.2326	0.1952	0.6046	0.6206	0.4153	1.3765	NaN	NaN	cuckoo
100	5.8966	0.4344	0.4889	0.7535	-2.4279	-2.0228	-1.1662	0.0738	-0.0585	0.5113	0.6063	0.8773	0.9964	1.4011	0.0390	0.3396	8.7084 cuckoo
101	8.5112	1.9014	-0.4215	0.8994	-0.9008	0.1381	-0.7010	-1.2972	-1.6487	-1.0697	-0.4959	-0.4426	0.3948	0.6707	0.0954	0.9707	10.1759 decacto
102	10.8433	3.6078	-2.1090	-1.6974	-1.5466	-0.0355	-1.3374	-1.0219	-0.3252	0.1922	0.3952	-0.0245	0.1212	0.9237	0.0792	0.4889	6.1723 decacto
103	10.1721	1.7695	-1.7082	0.3667	-0.8665	-0.6093	-1.3995	-0.6227	-0.3570	-0.0656	0.4913	0.3584	0.4335	1.8966	NaN	NaN	decacto
104	11.1184	2.2652	-2.3621	-0.8902	-1.1989	0.3974	-0.4195	-0.9860	-1.5680	-0.8092	0.2516	-0.2052	-0.2713	1.8636	0.1074	0.9750	9.0784 decacto
105	8.5744	2.2862	-2.1431	-0.9456	-1.0187	-0.2843	-2.0234	-1.3774	-0.1372	0.5770	0.6464	0.1120	0.3488	0.9434	NaN	NaN	decacto
106	-1.1306	-1.3848	0.1868	-0.1953	-0.1110	-1.3341	-1.3894	-0.9287	0.2673	0.2965	0.8896	0.6382	-0.2968	0.5984	NaN	NaN	decacto
107	12.1316	2.1135	-2.9915	-0.3135	-0.0265	-0.0330	-1.7963	-1.0239	-0.4839	-0.3770	0.1895	0.4256	0.5883	0.9325	NaN	NaN	decacto
108	11.5949	-0.3879	-1.5781	0.7162	-1.5112	-0.1456	-1.0841	-1.1262	-0.4032	-0.2563	-0.0814	0.1958	0.3202	1.9187	NaN	NaN	decacto
109	10.3519	-0.1103	-1.4879	-0.2176	-1.8580	-0.7512	-1.5901	-0.8845	-0.1712	0.5423	0.6393	0.3338	0.4419	0.7628	0.1482	0.9573	6.4603 decacto
110	11.2826	1.4106	-3.1234	-1.1142	-1.2654	-0.7473	-1.7632	-0.5242	-0.1103	-0.0078	0.6511	0.5930	0.2926	1.2331	0.1488	0.6768	4.5483 decacto
111	7.7482	2.7984	-1.0667	-0.2146	-0.5378	0.3567	-1.4578	-1.7155	-1.4059	-0.7639	0.2870	0.3614	0.2991	1.6236	0.0858	0.4264	4.9699 decacto
112	11.2886	1.0659	-2.6665	-0.0911	-1.0633	-0.9253	-1.4424	-0.5073	-0.6253	-0.5300	0.2870	0.3614	0.2991	1.6236	0.0858	0.4264	4.9699 decacto

Εικόνα 2: Classification Learner – Επιλογή μεταβλητών και διασταυρούμενη επικύρωση

Name	Type	Range	Import as
VarName1	double	0.0019946 .. 0.18...	Predictor
VarName2	double	0.486865 .. 0.928...	Predictor
VarName3	double	477.963 .. 14811.1	Predictor
VarName4	double	499.806 .. 9175.25	Predictor
VarName5	double	0.00318533 .. 0.9...	Predictor
VarName6	double	537.825 .. 4006.55	Predictor
VarName7	double	997.702 .. 5926.13	Predictor
VarName8	double	1.95331 .. 18.6274	Predictor
VarName9	double	0.00182729 .. 0.2...	Predictor
VarName10	double	0.572242 .. 0.868...	Predictor
VarName11	double	-2.11806 .. 12.3498	Predictor
VarName12	double	-9.62965 .. 4.09729	Predictor
VarName13	double	-5.55302 .. 1.94369	Predictor
VarName14	double	-2.63569 .. 3.64316	Predictor
VarName15	double	-2.43245 .. 4.4681	Predictor
VarName16	double	-2.53638 .. 3.06169	Predictor
VarName17	double	-4.06131 .. 0.7896...	Predictor
VarName18	double	-2.10776 .. 1.16836	Predictor
VarName19	double	-1.64865 .. 1.97371	Predictor
VarName20	double	-1.44493 .. 2.65511	Predictor
VarName21	double	-1.42172 .. 1.40075	Predictor
VarName22	double	-1.04065 .. 1.38904	Predictor
VarName23	double	-1.91994 .. 1.00711	Predictor
VarName24	double	0 .. 1.95421	Predictor
NaN	double	0.0131973 .. 1.5312	Predictor
NaN1	double	0.186332 .. 0.974...	Predictor
NaN2	double	0.214822 .. 38.6656	Predictor
otus	categorical	3 unique	Response

☒ Cross-Validation

Protects against overfitting by partitioning the data set into folds and estimating accuracy on each fold.

Cross-validation folds: 5 folds

☐ Holdout Validation

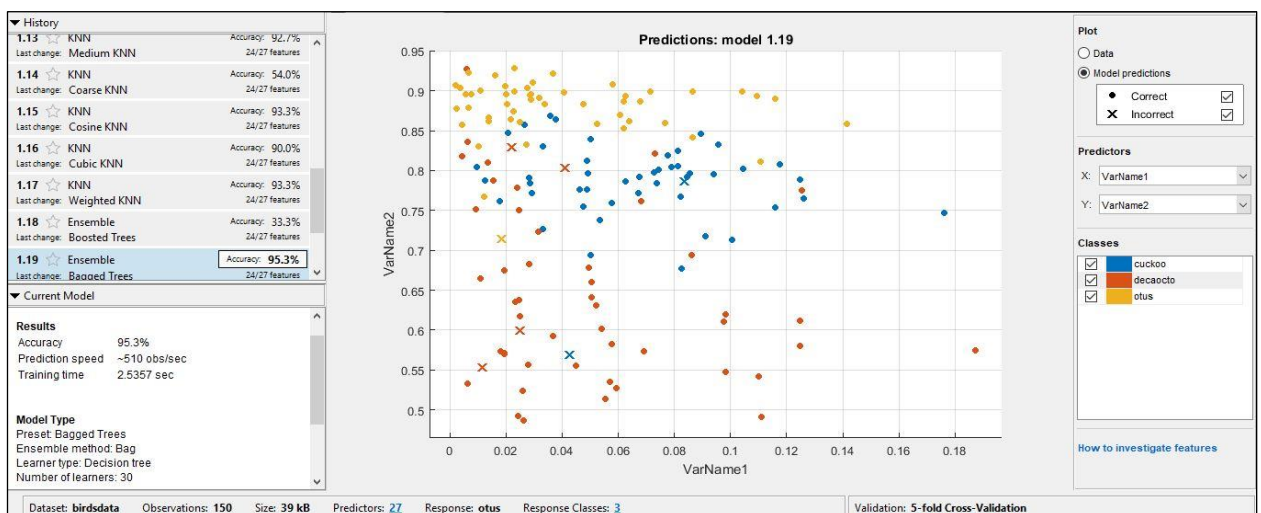
Recommended for large data sets.

Percent held out: 25%

☐ No Validation

No protection against overfitting.

Εικόνα 3: Classification Learner – Επιλογή / εκπαίδευση ταξινομητή



Παράρτημα 6

Εικόνα 1: birdMelos – Άνοιγμα εφαρμογής



Εικόνα 2: birdMelos – Εισαγωγή αρχείου ήχου στο σύστημα και έναρξη αναγνώρισης



Εικόνα 3: birdMelos – Ολοκλήρωση αναγνώρισης, πληροφορίες για το είδος του πουλιού

