



**ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ**

**Σχολή Εφαρμοσμένων Τεχνών
Ακουστικός Σχεδιασμός και Ψηφιακός Ήχος**

**Διπλωματική Εργασία:
Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές
εφαρμογές**

Πέτρος Αμοργιανός

Επιβλέπων καθηγητής Α΄
Δημούλας Χαράλαμπος

Επιβλέπων καθηγητής Β΄
Κωτσάκης Ρήγας



Breisach am Rhein, Ιούνιος 2026

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

© Copyright

Πέτρος Αμοργιανός Σχολή Εφαρμοσμένων Τεχνών (ΣΕΤ),
Ακουστικός Σχεδιασμός και Ψηφιακός Ήχος (ΑΣΠ), Breisach am Rhein

Διπλωματική Εργασία:
Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Πέτρος Αμοργιανός

Επιτροπή Εξέτασης: Δημούλας Χαράλαμπος, Καθηγητής
Κωτσάκης Ρήγας, Καθηγητής

Σημείωμα του συγγραφέα

Το παρόν δοκίμιο αποτελεί διπλωματική εργασία, η οποία συντάχθηκε για το τμήμα Ακουστικός Σχεδιασμός και Ψηφιακός Ήχος (ΑΣΠ) του Ελληνικού Ανοικτού Πανεπιστημίου, το οποίο έχει έδρα την Πάτρα και υποβλήθηκε προς εξέταση τον Σεπτέμβριο του 2026. Ο συγγραφέας είναι σε θέση να βεβαιώσει ότι το περιεχόμενο της παρούσας διπλωματικής εργασίας είναι αποτέλεσμα προσωπικής εργασίας. Επιπλέον ότι έχει γίνει η κατάλληλη αναφορά στην εργασία τρίτων, και όλων των πηγών που δανειστικό περιεχόμενο σύμφωνα με τους κανόνες της ακαδημαϊκής δεοντολογίας. Οι απόψεις που παρουσιάζονται στην παρούσα διπλωματική εργασία εκφράζουν αποκλειστικά τον συγγραφέα του παρόντος δοκιμίου.

Ευχαριστίες

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας, θα ήθελα ιδιαίτερα να εκφράσω τις θερμές μου ευχαριστίες προς τον καθηγητή κ. Χαράλαμπο Δημούλα για την πολύτιμη καθοδήγηση του, καθώς και την συνεχή υποστήριξη που μου παρείχε καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένεια μου, τους φίλους μου και όλους όσοι στάθηκαν δίπλα μου και με υποστήριξαν κατά τη διάρκεια αυτής της προσπάθειας.

Για το τέλος, θα ήθελα να εκφράσω μια ιδιαίτερη ευχαριστία, αλλά και μια μεγάλη συγγνώμη αξίζει στον εξάχρονο γιό μου για τον χρόνο που χάθηκε και δεν αξιοποιήθηκε για το παιχνίδι μας.

Περίληψη

Ο εντοπισμός της θέσης μιας ηχητικής πηγής αποτελεί μια διαδικασία, η οποία μπορεί να βρει εφαρμογή σε πολλούς τομείς της ανθρώπινης δραστηριότητας. Η οποία παρουσιάζει μεγάλη ανάπτυξη τις τελευταίες δεκαετίες, αλλά κυρίως σε ερευνητικό επίπεδο. Στην παρούσα εργασία πραγματοποιείται μια αναφορά σε θεωρητικό επίπεδο των διαδικασιών ηχητικού εντοπισμού από ρομποτικά συστήματα. Στη συνέχεια, σχεδιάστηκε και κατασκευάστηκε ένα ρομποτικό όχημα, το οποίο έχει ως βάση τον ενεργειακό εντοπισμό μιας ηχητικής πηγής, το οποίο υλοποιήθηκε με την υπολογιστική πλατφόρμα Arduino Mega. Το εν λόγω όχημα, μέσω δύο μικρόφωνων, έχει την δυνατότητα να συλλαμβάνει το ακουστικό σήμα και στην συνέχεια να κινείται προς την κατεύθυνση της μετάδοσης του, επιτυγχάνοντας μέσω αυτού του τρόπου τον εντοπισμό της ηχητικής πηγής.

Λέξεις κλειδιά:

Ηχητικός εντοπισμός, χωρικές τεχνολογίες ήχου, χωρική επεξεργασία ήχου, εντοπισμός ηχητικής πηγής

Abstract

Locating the location of a sound source is a process that can be applied in many areas of human activity. It has shown great development in recent decades, but mainly at the research level. In this paper, a theoretical reference is made to the sound localization processes by robotic systems. Subsequently, a robotic vehicle was designed and constructed, which is based on the energy localization of a sound source, which was implemented with the Arduino Mega computing platform. This vehicle, through two microphones, has the ability to capture the acoustic signal and then move in the direction of its transmission, achieving through this method the localization of the sound source.

Keywords:

Sound localization, spatial sound technologies, spatial sound processing, sound source localization

Πίνακας περιεχομένων

Κεφάλαιο 1	11
1. Θεωρητικό μέρος.....	11
1.1 Εισαγωγή	11
1.2 Ήχος	11
1.2.1 Ηχητικό πεδίο	13
1.2.2 Ο νόμος αντιστρόφου τετραγώνου	13
1.2.3 Κυματικά φαινόμενα ήχου.....	15
1.2.4 Θόρυβος και SNR.....	20
1.3 Μικρόφωνα	24
1.3.1 Εισαγωγή	24
1.3.2 Είδη μικροφώνων.....	25
1.3.3 Χαρακτηριστικά μικροφώνων.....	29
1.4 Αντίληψη του ήχου.....	35
1.4.1 Η ανατομία του ανθρώπινου οργάνου της ακοής	35
1.4.2 Χρονική αντίληψη του ηχητικού σήματος.....	38
1.4.3 Αντίληψη της κατεύθυνσης ηχητικού σήματος	39
1.5 Σύνοψη 1 ^{ου} Κεφαλαίου	43
Κεφάλαιο 2	45
2. Ανάλυση των τεχνικών ηχητικού εντοπισμού	45
2.1 Εισαγωγή	45
2.2 Ιστορική αναδρομή	45
2.3 Σύγχρονες μέθοδοι ηχητικού εντοπισμού	48
2.4 Μέθοδος εντοπισμού Ισχύος Λαμβανόμενου Σήματος (RSS)	49
2.5 Μέθοδος εντοπισμού Κατεύθυνσης Άφιξης (DOA).....	50
2.6 Μέθοδος εντοπισμού Γωνίας άφιξης (AOA)	54
2.6.1 Αλγόριθμος δισδιάστατου εντοπισμού θέσης με τριγωνισμό	57
2.6.2 Αλγόριθμος δισδιάστατου εντοπισμού θέσης με αντιστοίχιση	57
2.7 Μέθοδος εντοπισμού Χρόνου άφιξης (TOA)	59
2.8 Μέθοδος εντοπισμού Διαφοράς Χρόνου άφιξης (TDOA)	62
2.9 Μέθοδος εντοπισμού Διαφοράς Συχνότητας άφιξης (FDOA).....	65
2.10 Μέθοδος εντοπισμού Πολλαπλής Ταξινόμησης Σημάτων (MUSIC)	70
2.11 Μέθοδος ενεργειακού ηχητικού εντοπισμού (EBL)	73
2.12 Σύνοψη 2 ^{ου} κεφαλαίου.....	76
Κεφάλαιο 3	78
3. Σχεδίαση και παρουσίαση τεχνικών ηχητικού εντοπισμού στη ρομποτική	78
3.1 Εισαγωγή	78

3.2 Ιστορική αναδρομή	79
3.3 Παρουσίαση των προβλημάτων εντοπισμού ηχητικής πηγής	80
3.3.1 Παρουσίαση του μοντέλου διάδοσης	81
3.3.2 Παρουσίαση των ακουστικών χαρακτηριστικών	82
3.3.3 Παρουσίαση των διαδικασιών χαρτογράφησης	82
3.4 Παρουσίαση μεθόδων εντοπισμού ηχητικής πηγής από άκρο σε άκρο	83
3.4.1 Εκτίμηση κατεύθυνσης άφιξης από μια ηχητική πηγή	84
3.4.2 Εκτίμηση της δισδιάστατης κατεύθυνσης άφιξης από μια ηχητική πηγή	88
3.4.3 Εκτίμηση της κατεύθυνσης άφιξης από πολλές ηχητικές πηγές	92
3.4.4 Εκτίμηση της απόστασης της ηχητικής πηγής	105
3.5 Παρουσίαση μεθόδων παρακολούθησης της ηχητικής πηγής	107
3.5.1 Τα φίλτρα Kalman	107
3.5.2 Τα φίλτρα σωματιδίων	111
3.5.3 Διάφορες άλλες προσεγγίσεις	113
3.6 Σύνοψη 3 ^{ου} κεφαλαίου	114
Κεφάλαιο 4	115
4. Ανάπτυξη ρομποτικής εφαρμογής	115
4.1 Εισαγωγικό μέρος - Σκοπός	115
4.2 Ανάλυση υλικού του ρομποτικού οχήματος	115
4.3 Ανάλυση αλγόριθμου ρομποτικού οχήματος	137
4.4 Σχεδίαση και ανάπτυξη του ρομποτικού οχήματος	145
4.5 Παρουσίαση της ηχητικής πηγής	149
4.6 Παρουσίαση του χώρου διεξαγωγής του πειράματος	151
4.7 Μαθηματική μελέτη ηχητικού εντοπισμού	154
4.8 Σύνοψη 4 ^{ου} κεφαλαίου	162
Κεφάλαιο 5	163
5. Αξιολόγηση	163
5.1 Γενικά	163
5.2 Αξιολόγηση αποτελεσμάτων	163
5.2.1 Αξιολόγηση των στόχων ανάπτυξης	164
5.2.2 Αξιολόγηση χωρικού εντοπισμού	171
5.3 Προβλήματα υλοποίησης	171
5.4 Μελλοντικά ανοιχτά ζητήματα	172
6. Βιογραφικές αναφορές	173
6.1 Ελληνόγλωσσες Βιογραφικές αναφορές	173
6.2 Ξενόγλωσσες Βιογραφικές αναφορές	173
7. Ιστογραφία	176
8. Video	178

9. Data sheets	178
<i>Παράρτημα Α</i>	<i>179</i>
9.1 Αλγόριθμος 1.....	179
<i>Παράρτημα Β</i>	<i>188</i>
9.2 Αλγόριθμος 2.....	188

*« Αν οι άνθρωποι δεν πιστεύουν ότι τα μαθηματικά είναι απλά, είναι γιατί δεν έχουν συνειδητοποιήσει πόσο πολύπλοκη είναι η ζωή »
John von Neumann*

Κεφάλαιο 1

1. Θεωρητικό μέρος

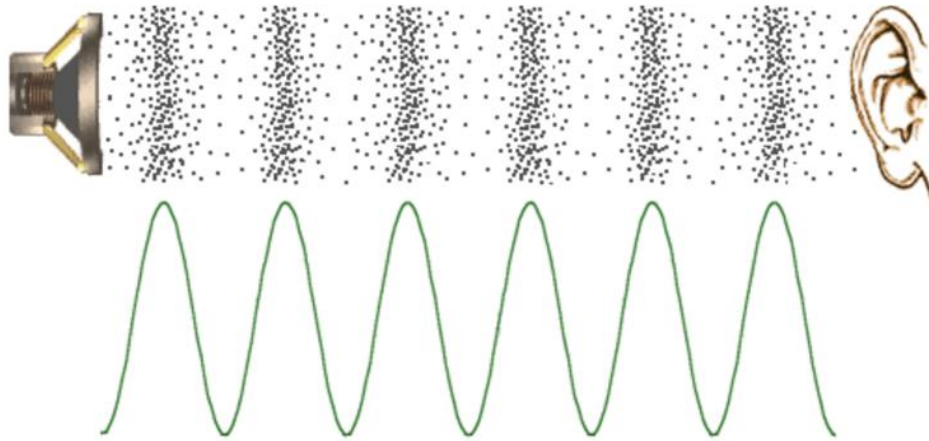
1.1 Εισαγωγή

Στο θεωρητικό μέρος της παρούσας εργασίας, παρατίθενται τα βασικά στοιχεία αναφορικά με τον ήχο ως έννοια, τον τρόπο που διαδίδεται και τη διαδικασία της επεξεργασίας του ως σήμα. Επιπλέον, γίνεται αναφορά στα φαινόμενα που παρατηρούνται κατά τη διάδοση του ήχου στον χώρο. Ιδιαίτερη έμφαση δίνεται, στην αντίληψη του ήχου από τον άνθρωπο και στον προσδιορισμό της θέσης της ηχητικής πηγής μέσα στο χώρο. Τέλος, γίνεται αναφορά στις μεθόδους δημιουργίας ηχητικών σημάτων με βάση χωρικά χαρακτηριστικά.

1.2 Ήχος

Ήχος ορίζεται η διαταραχή της ατμοσφαιρικής πίεσης, η οποία μπορεί να διεγείρει το ακουστικό αισθητήριο όργανο, κυρίως του ανθρώπου, το οποίο δίνει την δυνατότητα της ακοής. Μεταβολές στην ατμοσφαιρική πίεση μπορεί να παραχθούν από ταλαντωμένα σώματα, όπως για παράδειγμα, ένα τύμπανο, από διάφορα άλλα φαινόμενα, όπως από την σειρήνα του ασθενοφόρου, η οποία δημιουργεί κυκλική διακοπή της ροής του αέρα, από την τουρμπίνα του αεροπλάνου, η οποία δημιουργεί ανάδευση αέρα και τέλος από τους φυσητήρες που δημιουργούν απότομη μεταβολή της πίεσης του αέρα μέσω του συστήματος σωληνώσεων. Το ανθρώπινο ακουστικό αισθητήριο ανταποκρίνεται σε ένα ευρύ φάσμα συχνοτήτων από 16 Hz μέχρι και 20kHz, τα οποία όρια όμως δεν είναι σταθερά, καθώς εξαρτώνται από διάφορους παράγοντες, όπως η ηλικία του ανθρώπου, η οποία επηρεάζει αρκετά την ακουστική οξύτητα. Στην πραγματικότητα, όμως το ανθρώπινο αυτί παρουσιάζει μέγιστη ευαισθησία στο εύρος συχνοτήτων μεταξύ 1000Hz μέχρι και 6000Hz (Λουτρίδη, 2024).

Για τη διάδοση ενός ηχητικού κύματος, απαιτείται απαραίτητα κάποιο υλικό μέσο μεταξύ της πηγής του ήχου και του δέκτη. Τα ηχητικά κύματα δύναται να διαδοθούν σε στερεά, υγρά και αέρια, αλλά δεν μπορούν να διαδοθούν στο απόλυτο κενό, διότι σε αυτή την περίπτωση λείπει το υλικό διάδοσης. Όταν κάποιο ερέθισμα προκαλεί διαταραχή στο υλικό μέσο, αυτή μεταδίδεται στα μόρια της ύλης και στην συνέχεια αυτά ασκούν δυνάμεις στα γειτονικά τους μόρια. Μέσω αυτής της αλληλεπίδρασης, προκαλείται διαταραχή των μορίων, που ταξιδεύει μέσα στο μέσο, επιτυγχάνοντας τελικά τη μεταφορά του ηχητικού σήματος μέσα στο εκάστοτε υλικό (Φλώρος, 2013).



Εικόνα 1.1 Σχηματική απεικόνιση της διάδοσης του ήχου (<https://users.sch.gr>).

Στην συνέχεια, πραγματοποιείται αναφορά στην ταχύτητα διάδοσης του ήχου, και στον τρόπο με τον οποίο αυτή είναι ανάλογη με το μήκος κύματος και με τη συχνότητα του ηχητικού σήματος. Η ταχύτητα του ήχου εκφράζεται από τον παρακάτω τύπο.

$$c = \lambda * f \quad (1.1)$$

όπου

c: η ταχύτητα διάδοσης του ήχου

λ: το μήκος κύματος

f: η συχνότητα

Η ταχύτητα με την οποία διαδίδεται το ηχητικό κύμα είναι ανάλογη με τα χαρακτηριστικά του μέσου διάδοσης. Ειδικότερα, κατά τη διάδοση του ηχητικού κύματος στον ατμοσφαιρικό αέρα, η ταχύτητα επηρεάζεται από την θερμοκρασία του μέσου διάδοσης. Στην περίπτωση αυτή, η ταχύτητα του ήχου υπολογίζεται με βάση την ακόλουθη σχέση

$$c = 332 * \sqrt{1 + \frac{\theta}{273}} \quad (1.2)$$

όπου

θ: η θερμοκρασία σε βαθμούς Κελσίου

Συνεπώς, για θερμοκρασία 20°C βαθμούς Κελσίου, η ταχύτητα διάδοσης του ήχου είναι

$$c=340 \text{ m/sec} \quad (\text{Φλώρος, 2013})$$

1.2.1 Ηχητικό πεδίο

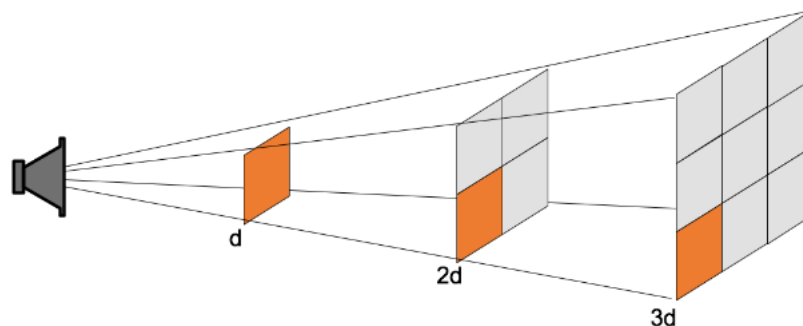
Ηχητικό πεδίο (sound field) ορίζεται η περιοχή του χώρου στην οποία διαδίδεται το ηχητικό κύμα. Το ηχητικό πεδίο ορίζεται αφενός από τα χαρακτηριστικά της ενεργής πηγής ήχου ή των ενεργών πηγών ήχου και αφετέρου από τα εμπόδια που ενεργούν πάνω στο ηχητικό κύμα κατά την πορεία του στον χώρο ή που καθορίζουν τον ίδιο το χώρο (Φλώρος, 2013).

Το ηχητικό πεδίο δύναται να είναι είτε διάχυτο είτε ελεύθερο. Στην περίπτωση του διάχυτου πεδίου γίνεται αναφορά στην συμπεριφορά του ηχητικού κύματος σε κλειστό χώρο. Αναλύοντας περαιτέρω την έννοια του ηχητικού πεδίου, παρουσιάζονται οι ακόλουθες κατηγορίες.

1. Ελεύθερο ηχητικό πεδίο είναι η περιοχή του χώρου στη οποία το ηχητικό κύμα έχει ελεύθερη διάδοση.
2. Απευθείας ηχητικό πεδίο είναι η περιοχή του χώρου στην οποία το ηχητικό κύμα δεν υπόκειται σε ανακλάσεις.
3. Εγγύς ηχητικό πεδίο είναι η περιοχή του χώρου κοντά στην ηχητική πηγή στην οποία η ηχητική πίεση και η ταχύτητα του ήχου δεν είναι ισοφασικές.
4. Άπω ηχητικό πεδίο είναι η περιοχή του χώρου μεταξύ του εγγύς ηχητικού πεδίου μέχρι και το άπειρο.
5. Αντηχητικό ηχητικό πεδίο είναι η περιοχή του χώρου όπου μέρος του ηχητικού κύματος που εκπέμπεται από την ηχητική πηγή λαμβάνει μια ανάκλαση τουλάχιστον. (Βλαχόπουλος, 2022)

1.2.2 Ο νόμος αντιστρόφου τετραγώνου

Έστω ότι υπάρχει μια ιστροπική πηγή ήχου S σε ελεύθερο ηχητικό πεδίο, η οποία επιδρά ισοδύναμα προς όλες τις κατευθύνσεις με ακουστική ισχύ P_{ac} . Η ισχύς που παράγεται από την πηγή παραμένει σταθερή σε όλες τις δυνητικές αποστάσεις που μπορεί να βρεθεί ο ακροατής, αντίθετα όμως, η ένταση του ήχου μειώνεται σύμφωνα με το τετράγωνο της απόστασης, λόγω του ότι η ακουστική ισχύς κατανέμεται σε μεγαλύτερη επιφάνεια. Αντίστοιχα μειώνεται και η ακουστική πίεση, δεδομένου ότι η ακουστική ένταση για το επίπεδο κύμα είναι ανάλογη του τετραγώνου της ενεργούς τιμής της ακουστικής πίεσης. (Λουτρίδης, 2024)



Εικόνα 1.2 Απεικόνιση του νόμου του αντίστροφου τετραγώνου, στον οποίο ισχύει ότι η ένταση της ηχητικής ισχύς μειώνεται ανά μονάδα επιφάνειας, καθώς αυξάνεται το τετράγωνο της απόστασης d (Velzen, 2012).

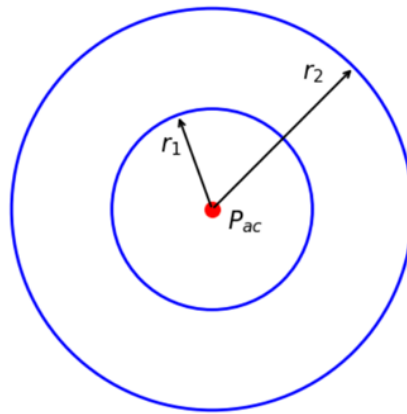
Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Για παράδειγμα, έστω για τον υπολογισμό της έντασης του ήχου που επικρατεί σε δύο επιφάνειες δύο σφαιρών με ακτίνες r_1 και r_2 , όπως παρουσιάζεται στη παρακάτω εικόνα. Η ένταση στην επιφάνεια κάθε σφαίρας με ακτίνα r υπολογίζεται από τον παρακάτω τύπο

$$I = \frac{P_{ac}}{4\pi r^2} \quad (1.3)$$

και

$$I = \frac{p^2}{\rho_0 c} \quad (1.4)$$



Εικόνα 1.3 Απεικόνιση του νόμου αντιστρόφου τετραγώνου για δύο σφαιρικές επιφάνειες. (Λουτρίδης, 2024)

Με γνωστή τη στάθμη του ήχου σε απόσταση r_1 από την πηγή, δύναται να υπολογιστεί η στάθμη σε απόσταση r_2 , σύμφωνα με τον παρακάτω τύπο

$$L_{p2} = L_{p1} + 20 \log_{10} \left(\frac{r_1}{r_2} \right) \quad (1.5)$$

Βάσει του προηγούμενου τύπου, και για λόγο αποστάσεων $\frac{r_2}{r_1} = 2$ μπορεί να υπολογιστεί $L_{p2} = L_{p1} - 6$. Συνεπώς, η στάθμη της ηχητικής πίεσης ελαττώνεται, με κάθε διπλασιασμό της απόστασης από την ηχητική πηγή κατά 6 dB.

1.2.3 Κυματικά φαινόμενα ήχου

Ο ήχος δύναται να διαδίδεται ελεύθερα σε ηχητικό πεδίο χωρίς εμπόδια και με απουσία ανακλάσεων. Εντούτοις, σε πραγματικές συνθήκες, όταν διαδίδεται ο ήχος υφίσταται τουλάχιστον μια ανάκλαση που πραγματοποιείται από το έδαφος, όταν δηλαδή ο ήχος διαδίδεται κοντά στην γήινη επιφάνεια. Ο ήχος διαδίδεται στον αέρα ως διάμηκες κύμα. Η ταχύτητα διάδοσης ορίζεται από τον παρακάτω τύπο

$$c = \sqrt{\frac{\gamma RT}{M}} \quad (1.6)$$

όπου

γ : 1,4 για διατομικά αέρια

M : μοριακό βάρος αερίου

R : σταθερά ίση με 8314 J/(K*kmol)

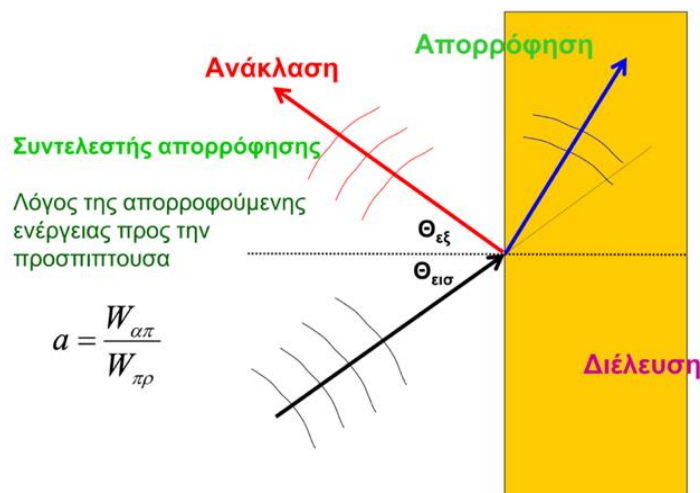
T : θερμοκρασία

Βάσει του ορισμού της ταχύτητας διάδοσης, προκύπτει το συμπέρασμα ότι αυτή εξαρτάται από το μέσο στο οποίο διαδίδεται το ηχητικό κύμα, αλλά και από την θερμοκρασία. (Λουτρίδης, 2024)

Ακολουθως, πραγματοποιείται μια συνοπτική αναφορά στα κυματικά φαινόμενα του ήχου.

Εξασθένηση του ηχητικού κύματος.

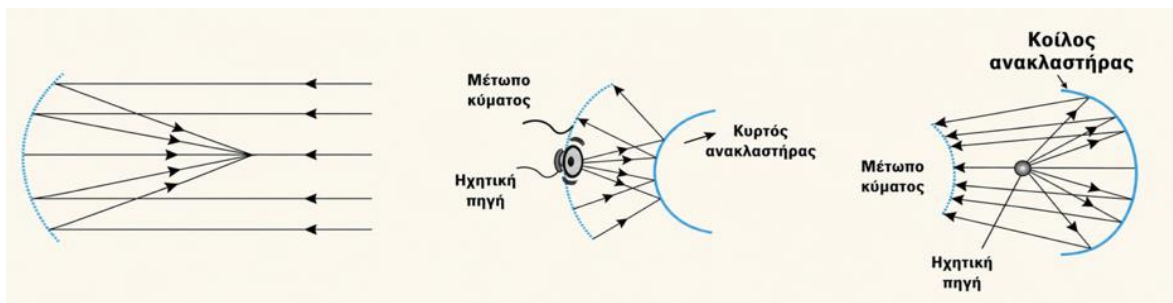
Το ηχητικό κύμα κατά τη διάδοση του στον αέρα εξασθενεί, λόγω της απορρόφησης της ηχητικής του ενέργειας από τα στοιχεία του ατμοσφαιρικού αέρα. Η εξασθένηση αυτή εξαρτάται σε μεγάλο βαθμό από την θερμοκρασία, η οποία επηρεάζει κυρίως τις υψηλές συχνότητες, αλλά και από την απόσταση. Σημειώνεται ότι για κάθε διπλασιασμό της απόστασης από την ηχητική πηγή, η στάθμη της ηχητικής πίεσης μειώνεται κατά 6 dB. Αυτό ισχύει μόνο στο ελεύθερο πεδίο, στο οποίο το ηχητικό κύμα διαδίδεται προς όλες τις κατευθύνσεις, χωρίς την παρουσία κάποιου εμποδίου. (Λουτρίδης, 2024)



Εικόνα 1.4 Γραφική απεικόνιση της ανάκλασης και απορρόφησης του ήχου σε ηχητικό εμπόδιο. (Παρουσιάσεις, Μεσογειακό πανεπιστήμιο)

Ανάκλαση του ήχου (reflection)

Ανάκλαση του ήχου παρατηρείται όταν, κατά την διάδοση του, το ηχητικό κύμα συναντήσει κάποιο εμπόδιο, το οποίο έχει διαστάσεις μεγαλύτερες από $\lambda/4$ σε σχέση με το μήκος κύματος του διαδιδόμενου ήχου. Για παράδειγμα, όταν υπάρχει επίπεδο εμπόδιο, η γωνία που σχηματίζεται από το ηχητικό κύμα που προσπίπτει στο εμπόδιο είναι ίση με τη γωνία ανάκλασης του ανακλώμενου ηχητικού κύματος. Επιπλέον, η ενέργεια που έχει το ανακλώμενο ηχητικό κύμα είναι πάντοτε μικρότερη από την ενέργεια του προσπίπτοντος ηχητικού κύματος. Το μέγεθος αυτής της διαφοράς, δηλαδή το ποσοστό της απώλειας της ενέργειας, εξαρτάται τόσο από το υλικό του εμποδίου, που δημιουργεί το φαινόμενο της ανάκλασης, όσο και από τη συχνότητα του προσπίπτοντος ηχητικού κύματος. (Φλώρος, 2013)

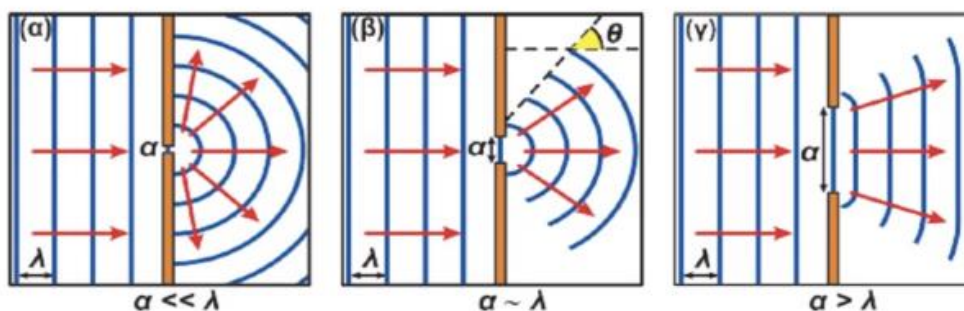


Εικόνα 1.5 Απεικόνιση του φαινομένου της ανάκλασης σε κυρτά εμπόδια (διαχυτές). (www.brainkart.com).

Στην περίπτωση που το εμπόδιο είναι κυρτή επιφάνεια και όχι επίπεδη, όπως αποτυπώνεται στο προηγούμενο σχήμα, ο ήχος ανακλάται προς όλες τις κατευθύνσεις, το οποίο έχει ως αποτέλεσμα να προκαλείται διάχυση του ήχου. (Φλώρος, 2013)

Περίθλαση του ηχητικού κύματος

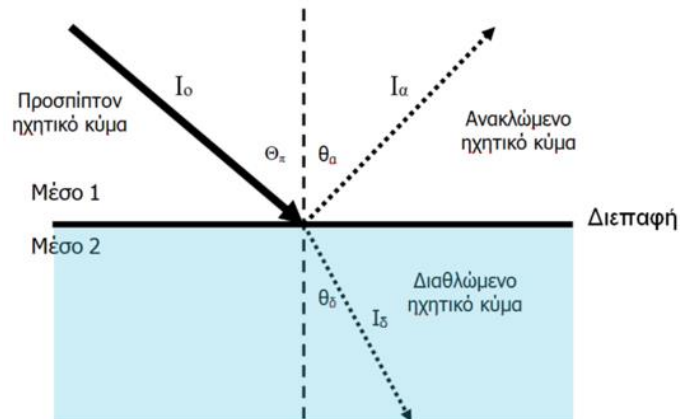
Κατά την πρόσπτωση του ηχητικού κύματος σε εμπόδιο, το οποίο έχει διαστάσεις μικρότερες από $\lambda/4$, τα ηχητικά κύματα κάμπτονται και υπερπηδούν το εμπόδιο. Το φαινόμενο αυτό, όπου τα ηχητικά κύματα υπερπηδούν το εμπόδιο ονομάζεται περίθλαση. Επιπλέον, σε αυτή την περίπτωση το εμπόδιο λειτουργεί ως δευτερογενής ηχητική πηγή. Η περίθλαση απεικονίζεται στο σχήμα που ακολουθεί. (Λουτρίδης, 2024)



Εικόνα 1.6 Απεικόνιση του φαινομένου της περίθλασης (https://www.youtube.com/watch?v=gHoHAz_pC8&t=97s).

Διάθλαση του ήχου

Κατά τη μετάβαση του ηχητικού κύματος από ένα μέσο διάδοσης σε ένα άλλο μέσο, η κατεύθυνση του μεταβάλλεται λόγω ότι υπάρχει αλλαγή στην ταχύτητα διάδοσης του, η οποία ταχύτητα εξαρτάται από το μέσο που διαδίδεται. Το φαινόμενο αυτό της μεταβολής της κατεύθυνσης του ήχου ονομάζεται διάθλαση του ήχου. Ως παράδειγμα αναφέρεται, ότι το φαινόμενο της διάθλασης του ήχου παρατηρείται όταν το ηχητικό κύμα μεταβεί από ένα μέσο διάδοσης χαμηλότερης πυκνότητας (π.χ. αέρα) σε ένα μέσο διάδοσης υψηλότερης πυκνότητας (π.χ. νερό), όπως απεικονίζεται στο σχήμα που ακολουθεί. (Φλώρος, 2013)

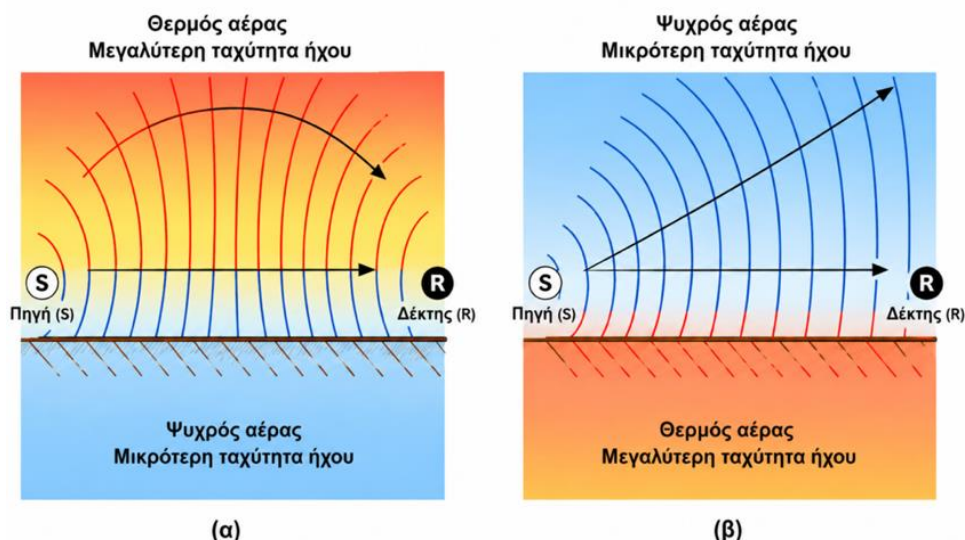


Εικόνα 1.7 Απεικόνιση του φαινομένου της διάθλασης (Καραϊσκος, 2015)

Βάσει της ανάλυσης του φαινομένου της διάθλασης στην προηγούμενη εικόνα, και θεωρώντας ότι η ταχύτητα διάδοσης στο πρώτο μέσο είναι c_1 και η ταχύτητα διάδοσης στο δεύτερο μέσο είναι c_2 . Η γωνία που δημιουργεί το προσπίπτον ηχητικό κύμα θ_π και η γωνία που δημιουργεί το διαθλώμενο ηχητικό κύμα θ_δ συνδέονται από την παρακάτω σχέση. (Φλώρος, 2013)

$$\frac{\sin \theta_\pi}{c_1} = \frac{\sin \theta_\delta}{c_2} \quad (1.7)$$

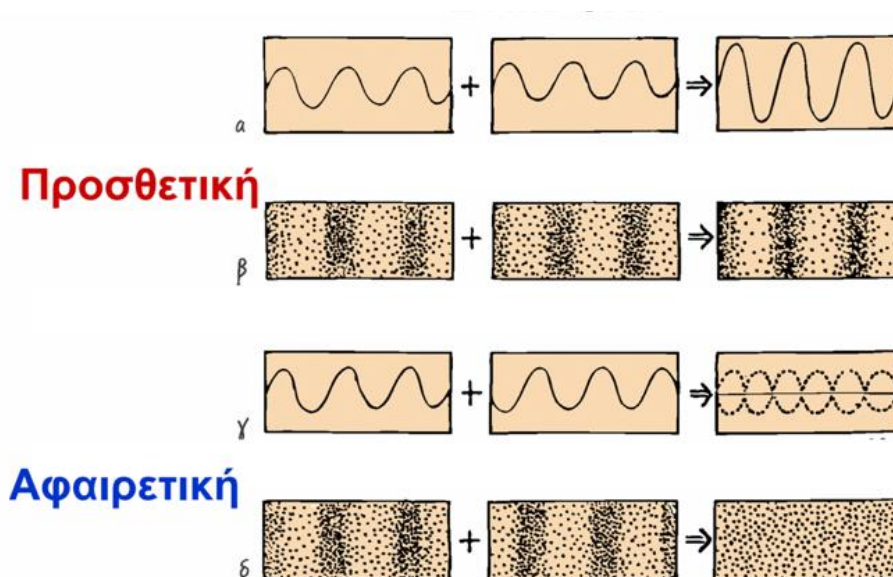
Επιπλέον, διάθλαση των ηχητικών κυμάτων, ενδέχεται να παρατηρηθεί και εντός του ίδιου μέσου, λόγω μεταβολής της θερμοκρασίας. Ως παράδειγμα αναφέρεται, ότι όταν μια ηχητική πηγή εκπέμπει ηχητικό κύμα σε υψηλότερη θερμοκρασιακή περιοχή κοντά στο έδαφος σε σχέση με τη θερμοκρασία του ατμοσφαιρικού αέρα, η ταχύτητα διάδοσης μειώνεται αντίστοιχα με την ελάττωση της θερμοκρασίας, και το ηχητικό κύμα έχει ανοδική κατεύθυνση. Αντιθέτως όταν η περιοχή κοντά το έδαφος είναι πιο ψυχρή από τον ατμοσφαιρικό αέρα, η θερμοκρασιακή διαφορά κατευθύνει τα ηχητικά κύματα προς το έδαφος, όπως αποτυπώνεται στην εικόνα που ακολουθεί. (Λουτρίδης, 2024)



Εικόνα 1.8 Απεικόνιση του φαινομένου της διάθλασης λόγω της μεταβολής της θερμοκρασίας στον ατμοσφαιρικό αέρα (Montalbano, 2012).

Συμβολή των ηχητικών κυμάτων

Στην περίπτωση, όπου δύο διαφορετικές ηχητικές πηγές εκπέμπουν ηχητικά κύματα στο ίδιο μέσο, και τα κύματα συναντηθούν σε κάποιο σημείο της περιοχής του μέσου. Τότε παράγεται ένα νέο ηχητικό κύμα, το οποίο προκύπτει από την άθροιση ή την αφαίρεση των αρχικών ηχητικών κυμάτων. Το φαινόμενο αυτό ονομάζεται συμβολή των ηχητικών κυμάτων και αποτυπώνεται στη παρακάτω εικόνα (Ευταξίας, 2024).

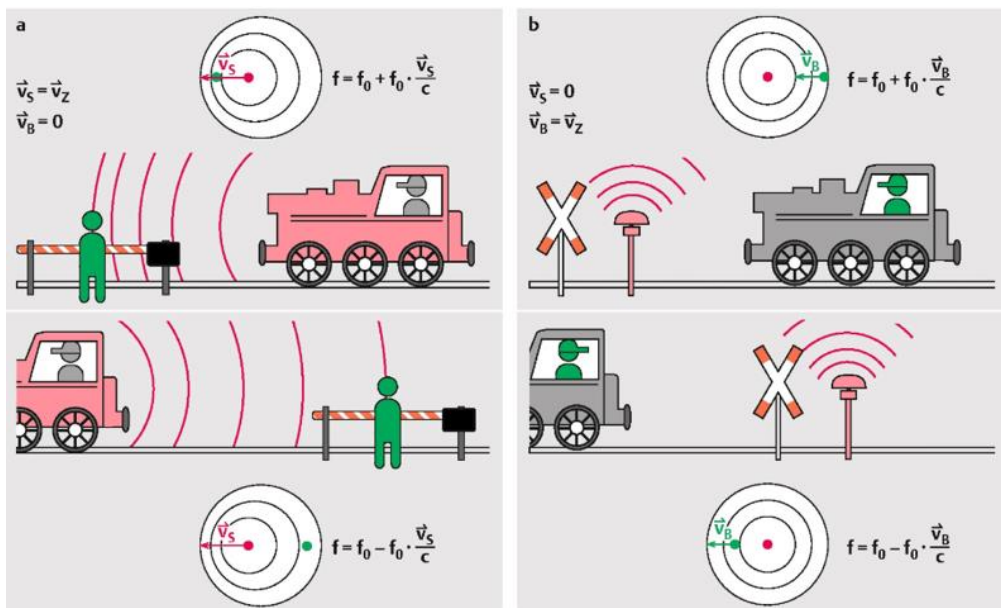


Εικόνα 1.9 Απεικόνιση του φαινομένου της συμβολής. (Ευταξίας, 2024)

Φαινόμενο Doppler

Στην περίπτωση στην οποία υπάρχει κάποιος ακίνητος παρατηρητής-δέκτης σε μια αποβάθρα σταθμού και προσεγγίζει ένα τρένο που κινείται με σταθερή ταχύτητα, ο ήχος που αντιλαμβάνεται ο παρατηρητής είναι οξύτερος, δηλαδή ο αντιλαμβανόμενος ήχος παρουσιάζει μεγαλύτερη συχνότητα. Αντίθετα, στην περίπτωση που το τρένο απομακρύνεται από την αποβάθρα και κατ' επέκταση από τον δέκτη, ο δέκτης αντιλαμβάνεται τον ήχο με μειούμενη συχνότητα. Από την άλλη πλευρά, ο ήχος που αντιλαμβάνεται ο μηχανοδηγός, ο οποίος βρίσκεται πάνω στο τρένο έχει σταθερή συχνότητα και για τις δύο περιπτώσεις.

Συνοψίζοντας, το φαινόμενο Doppler αναφέρεται στη μεταβολή του μήκους κύματος του ήχου που παράγεται από μια ηχητική πηγή, λόγω της σχετικής κίνησης μεταξύ μιας ηχητικής πηγής και του παρατηρητή.



Εικόνα 1.10 Απεικόνιση του φαινομένου doppler για τις δύο περιπτώσεις α. Κινούμενη πηγή ακίνητος παρατηρητής και β. Ακίνητη πηγή και κινούμενος παρατηρητής. (Amann-Vesti Beatrice, 2012)

Θεωρώντας ότι μια ηχητική πηγή εκπέμπει σε μια συχνότητα f_0 και με μήκος κύματος σχετικό με την συχνότητα αυτή λ_0 , παίρνουμε την εξίσωση

$$f_0 = \frac{c}{\lambda_0} \quad (1.8)$$

όπου

c: ταχύτητα του ήχου

Μια ηχητική πηγή εκπέμπει ηχητικά κύματα προς έναν δέκτη. Στην περίπτωση, στην οποία ο δέκτης απομακρύνεται από την ηχητική πηγή με ταχύτητα v , το μήκος κύματος του ήχου αυξάνει, ενώ στην περίπτωση, στην οποία ο δέκτης πλησιάζει την ηχητική πηγή, το μήκος κύματος του ήχου μειώνεται. Και για τις δύο περιπτώσεις, η σχέση της συχνότητας f_r που καταγράφει ο δέκτης, εκφράζεται από τον παρακάτω τύπο

$$f_r = \frac{c \pm v_r}{c \mp v_s} f_0 \quad (1.9)$$

όπου

f_r : η συχνότητα του ήχου που αντιλαμβάνεται ο παρατηρητής

c : ταχύτητα του ήχου

v_r : ταχύτητα παρατηρητή

v_s : ταχύτητα ηχητικής πηγής

και για τα πρόσημα ισχύει το (+) για όταν προσεγγίζουν και (-) όταν απομακρύνονται. (Λουτρίδης, 2024)

1.2.4 Θόρυβος και SNR

Αξίζει να σημειωθεί ότι ο ήχος συνίσταται στις μεταβολές της πίεσης του ατμοσφαιρικού αέρα, οι οποίες είναι η αιτία που διεγείρεται η ακοή. Στην συνέχεια, τα ηχητικά σήματα διακρίνονται σε τέσσερις κατηγορίες: απλός ήχο, σύνθετος ήχο, θόρυβος και κρότος.

Σε αυτό το σημείο, εστιάζοντας κυρίως στην τρίτη κατηγορία, ως θόρυβοι δύναται να οριστούν οι ήχοι που δημιουργούνται από ακανόνιστες και μη περιοδικές μεταβολές της πίεσης του ατμοσφαιρικού αέρα και οι οποίες συνήθως προξενούν στον ακροατή ένα δυσάρεστο ή ενοχλητικό αίσθημα. Ο θόρυβος αποτελεί μια έννοια, κατά βάση υποκειμενική και σχετική με διάφορους παράγοντες. Αυτός είναι ο λόγος, ότι το ίδιο ηχητικό σήμα άλλοτε θεωρείται θόρυβος και άλλοτε ως ένας επιθυμητός ήχος.

Ακουστός θόρυβος ορίζεται το ηχητικό σήμα, το οποίο εκδηλώνεται σε συχνότητες που γίνονται αντιληπτές από το ανθρώπινο αυτί. Πηγές που παράγουν θόρυβο μπορούν να εντοπιστούν παντού στο περιβάλλον, όπως σε χώρους εργασίας, στη φύση, καθώς και μέσα στο οικιακό χώρο.

Γενικά, ο ακουστός θόρυβος επηρεάζει αρνητικά την ποιότητα ζωής των ανθρώπων. Οι επιπτώσεις του στην υγεία των ανθρώπων σχετίζονται με τη στάθμη της ακουστικής πίεσης, τη διάρκεια έκθεσης και τη συχνότητα επανάληψης του φαινομένου. (www.wikiwand.com)

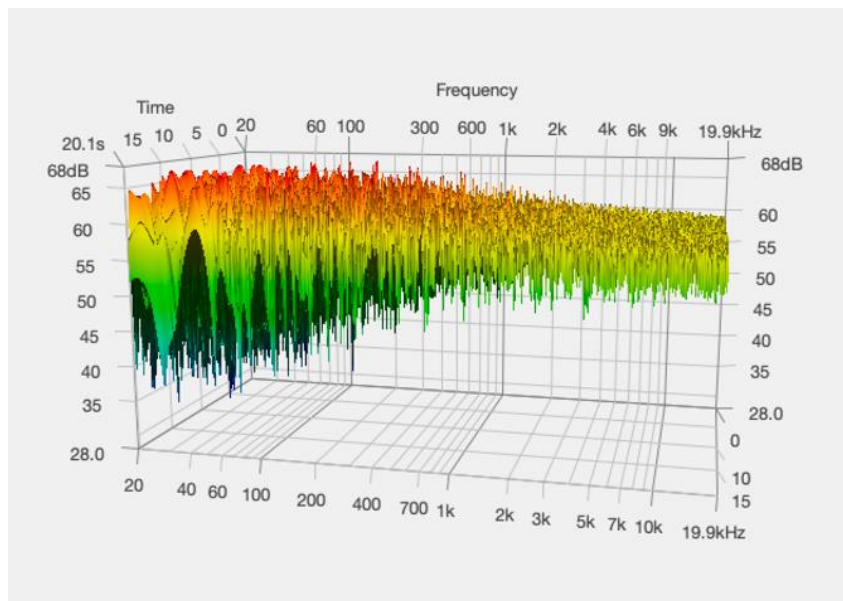
Ανάλογα με τη φασματική του ισχύ, ο θόρυβος διακρίνεται σε τρεις κατηγορίες τον ροζ θόρυβο, τον λευκό θόρυβο και τον καφέ ή κόκκινο θόρυβο.

Λευκός ή ευρυζωνικός θόρυβος ορίζεται ο θόρυβος, ο οποίος καλύπτει με την ίδια ισχύ όλες τις συχνότητες του ακουστικού φάσματος.

Λευκός θόρυβος



Εικόνα 1.11 Απεικόνιση του σήματος λευκού θορύβου στο πεδίο του χρόνου μέσω του λογισμικού Audacity.

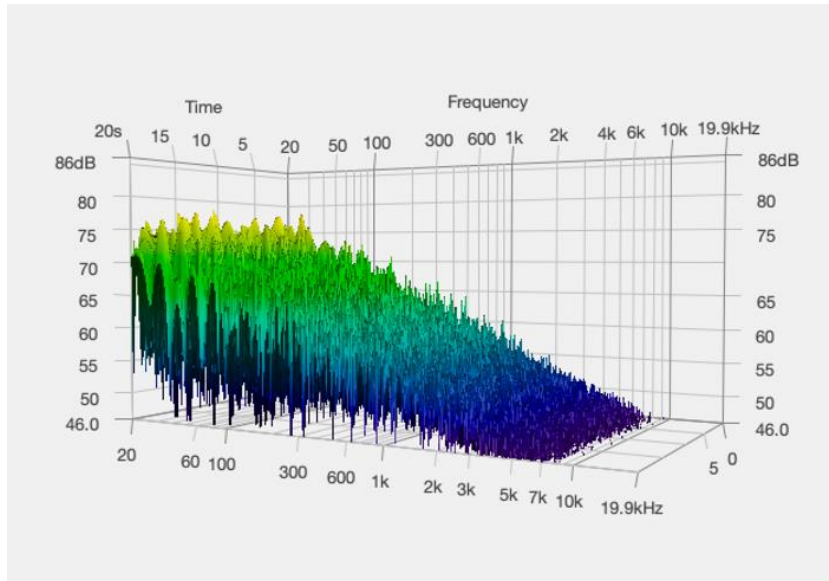


Εικόνα 1.12 Απεικόνιση του φασματογραφήματος του λευκού θορύβου μέσω του λογισμικού REW.

Ροζ θόρυβος ορίζεται ως ο ήχος που προκύπτει από μια τυχαία σύνθεση όλων των συχνοτήτων του ακουστικού φάσματος, παρουσιάζοντας ωστόσο μεγαλύτερη ισχύ στις χαμηλές συχνότητες. Χαρακτηρίζεται από την ίδια συνολική στάθμη ακουστικής πίεσης σε κάθε οκτάβα, όπως απεικονίζεται στα γραφήματα που ακολουθούν.



Εικόνα 1.13 Απεικόνιση του σήματος ροζ θορύβου στο πεδίο του χρόνου με το λογισμικό Audacity.

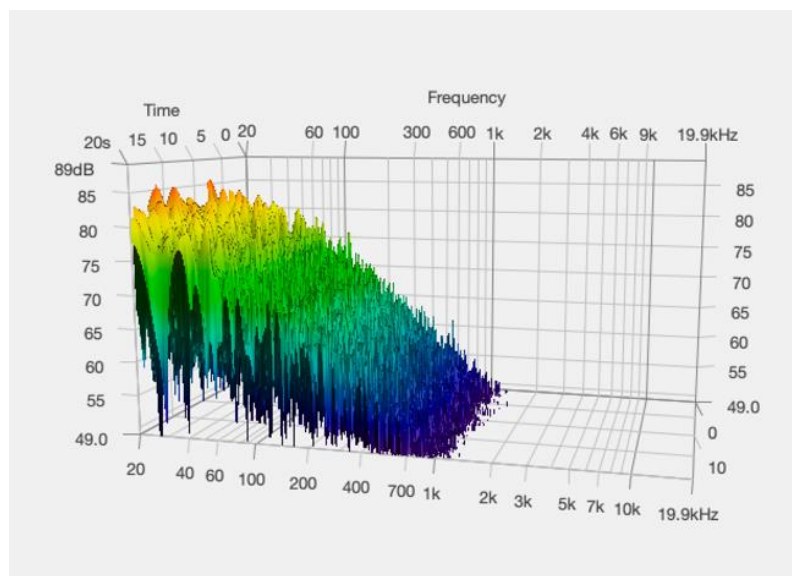


Εικόνα 1.14 Απεικόνιση του φασματογραφήματος του ροζ θορύβου με το λογισμικό REW.

Καφέ θόρυβος ορίζεται ως ο ήχος που προκύπτει από την τυχαία σύνθεση όλων των συχνοτήτων του ακουστικού φάσματος, παρουσιάζοντας ωστόσο μεγαλύτερη ισχύ στις χαμηλές συχνότητες σε σχέση με το ροζ θόρυβο, όπως απεικονίζεται στα γραφήματα που ακολουθούν (www.orionstrom.gr).



Εικόνα 1.15 Απεικόνιση του σήματος καφέ θορύβου στο πεδίο του χρόνου με το λογισμικό Audacity.



Εικόνα 1.16 Απεικόνιση του φασματογραφήματος του καφέ θορύβου με το λογισμικό REW.

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Στα ηχητικά συστήματα και στην ακουστική, κατά κοινή άποψη, η ισχύς του θορύβου αξιολογείται σε σχέση με την ισχύ των επιθυμητών ηχητικών σημάτων. Η σχέση αυτή εκφράζεται ως λόγος σήματος προς θόρυβο (Signal to Noise Ratio – SNR), ο οποίος λόγος σήματος προς θόρυβο ορίζεται από την ακόλουθη σχέση

$$SNR = \frac{P_{signal}}{P_{noise}} \quad (1.10)$$

όπου

P: μέση ισχύς των σημάτων

Συνήθως, ο λόγος σήματος προς θόρυβο εκφράζεται σε ντεσιμπέλ (dB) και ορίζεται από τον παρακάτω τύπο

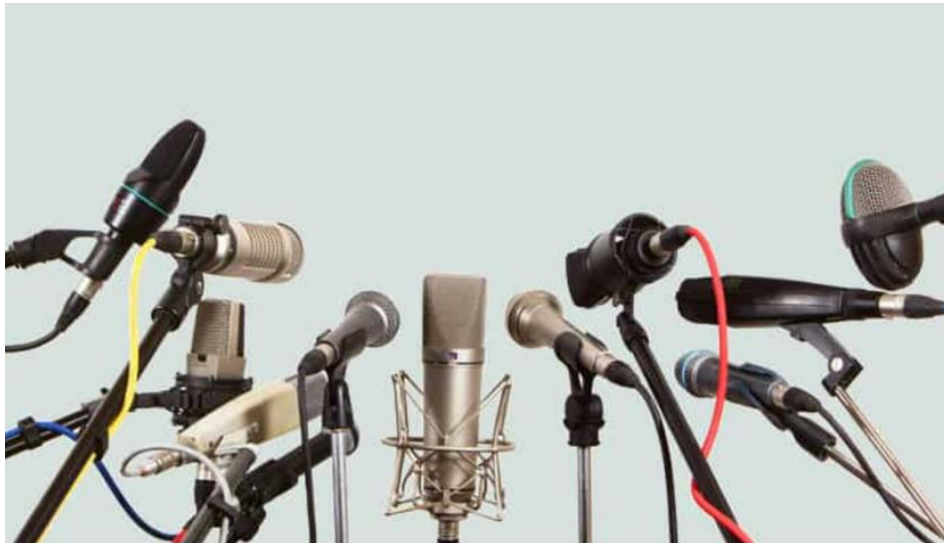
$$SNR_{(dB)} = 10 * \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) \quad (1.11)$$

Στην περίπτωση που ο λόγος σήματος προς θόρυβο είναι μεγαλύτερος από 0 dB, το οποίο συνεπάγεται ότι ο αντίστοιχος λόγος είναι μεγαλύτερος από τη μονάδα, τότε η ισχύς του θορύβου είναι μικρότερη από την ισχύ του επιθυμητού ηχητικού σήματος. (www.orionstrom.gr)

1.3 Μικρόφωνα

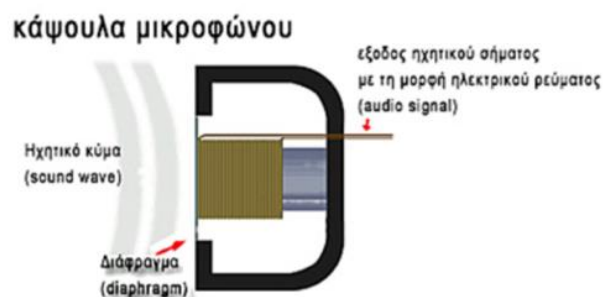
1.3.1 Εισαγωγή

Το μικρόφωνο είναι ένας ηλεκτροακουστικός μετατροπέας, ο οποίος μετατρέπει τα ηχητικά σήματα που λαμβάνει στην είσοδο του σε αντίστοιχα ηλεκτρικά σήματα. Το μικρόφωνο χρησιμοποιείται σε πολλές εφαρμογές όπως σε συστήματα επικοινωνίας, σε ηχητικές εγκαταστάσεις, σε στούντιο ηχογράφησης.



Εικόνα 1.17 Απεικόνιση διαφόρων τύπων μικροφώνων. (<https://performancehigh.net>)

Για την λειτουργία του μικροφώνου, το βασικότερο εξάρτημα είναι το διάφραγμα, το οποίο είναι μια λεπτή μεμβράνη, κατασκευασμένη από υλικό ευαίσθητο στις δονήσεις, που προκαλούνται από την πίεση του εκάστοτε ηχητικού κύματος. Μέσω του ηλεκτρονικού εξαρτήματος που λέγεται κάψουλα, το διάφραγμα ταλαντώνεται από τα ηχητικά κύματα που λαμβάνει. Το οποίο αντιγράφει τις κατάλληλες ενεργειακές διακυμάνσεις που δημιουργεί το ηχητικό σήμα και τις μετατρέπει σε κατάλληλο ηλεκτρικό σήμα. (Λουτρίδης, 2024)

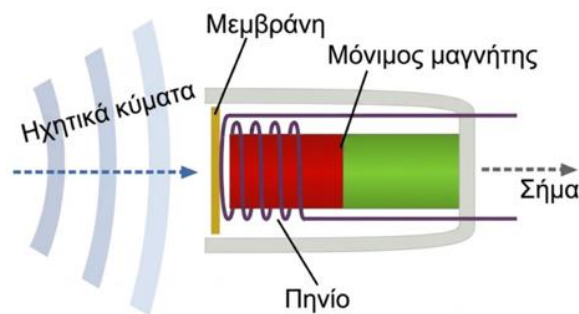


Εικόνα 1.18 Απεικόνιση του διαφράγματος και της κάψουλας του μικροφώνου. (Λουτρίδης, 2024)

1.3.2 Είδη μικροφώνων

1. Δυναμικά μικρόφωνα

Στα δυναμικά μικρόφωνα, τα ηχητικά κύματα διαπερνούν το προστατευτικό κάλυμμα, προκαλώντας κίνηση τόσο στο πηνίο λεπτού σύρματος, όσο και στο διάφραγμα. Το πηνίο λεπτού σύρματος εισέρχεται σε ομογενές μαγνητικό πεδίο μόνιμου μαγνήτη, ενός μαγνητικού κυκλώματος από μαλακό σίδηρο. Το πηνίο μήκους l κινείται με ταχύτητα u μέσα με μαγνητικό πεδίο με ένταση B . Το οποίο έχει σαν αποτέλεσμα την επαγωγή τάσης $V = Blu$ στα άκρα του πηνίου.



Εικόνα 1.19 Απεικόνιση της εσωτερικής δομής του δυναμικού μικροφώνου. (<https://blog.teufel.de>)

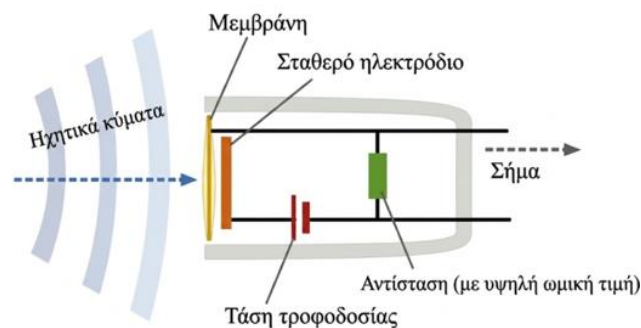
Τα δυναμικά μικρόφωνα αποτελούν γενικά απλές, οικονομικές και αρκετά ανθεκτικές κατασκευές, και τα οποία είναι ικανά για να παράγουν ηχητικό σήμα εξαιρετικής ποιότητας με ικανοποιητική λεπτομέρεια. Δεν είναι πολύ ευαίσθητα, ώστε να επηρεάζονται από σήματα θορύβου και παρουσιάζουν μεγάλη αντοχή σε υψηλά επίπεδα ηχητικής πίεσης SPL. Επιπλέον, δεν απαιτούν εξωτερική τροφοδοσία και δεν επηρεάζονται από μεταβολές της θερμοκρασίας και της υγρασίας. Ωστόσο, παρουσιάζουν ένα βασικό μειονέκτημα, ότι παράγουν στην έξοδο τους αρκετά ασθενές σήμα, με αποτέλεσμα να απαιτείται προενίσχυση. (Κοντογιάννης, 2019)



Εικόνα 1.20 Απεικόνιση ενός τυπικού πυκνωτικού μικροφώνου. (<https://mikrofon-tests.de>)

2. Πυκνωτικά μικρόφωνα

Τα πυκνωτικά μικρόφωνα βασίζονται στη λειτουργία μιας διάταξης πυκνωτή, ο οποίος παρουσιάζει ευαισθησία στα ηχητικά σήματα. Ο πυκνωτής αποτελείται από μια διάταξη δύο παράλληλων μεταλλικών πλακών ίσου μεγέθους, οι οποίες αποθηκεύουν ηλεκτρικό φορτίο, καθορίζοντας τη χωρητικότητα του. Όταν η διάταξη του πυκνωτή είναι φορτισμένη, τότε ανάμεσα στις πλάκες δημιουργείται ηλεκτρικό πεδίο, η ένταση του οποίου είναι αντιστρόφως ανάλογη της απόστασης μεταξύ των πλακών. Οποιαδήποτε μεταβολή της απόστασης μεταξύ των πλακών, μεταβάλλει αντίστοιχα την χωρητικότητα του πυκνωτή, όσο και την διαφορά δυναμικού στα άκρα του. Η μια εκ των δύο πλακών αποτελεί το διάφραγμα του μικροφώνου, το οποίο συνήθως είναι μεταλλικό με εξαιρετικά μικρό πάχος. Όταν τα ηχητικά σήματα προσπίπτουν στο μεταλλικό διάφραγμα, τότε το θέτουν σε ταλάντωση, με αποτέλεσμα τη μεταβολή της χωρητικότητας του πυκνωτή. Μέσω αυτής της διαδικασίας και με βάση την αρχή λειτουργίας του πυκνωτή, το ακουστικό σήμα που λαμβάνει το μικρόφωνο μετατρέπεται σε ηλεκτρικό σήμα.



Εικόνα 1.21 Απεικόνιση της εσωτερικής δομής ενός πυκνωτικού μικροφώνου (www.mikrofon.guru)

Τα πυκνωτικά μικρόφωνα απαιτούν εξωτερική τροφοδοσία ώστε να διατηρηθεί η τάση πόλωσης του πυκνωτή. Σε ορισμένες περιπτώσεις, έχουν έναν προενισχυτή ενσωματωμένο επειδή παρουσιάζουν υψηλή εσωτερική αντίσταση. Εξαιτίας των ηλεκτρονικών κυκλωμάτων του προενισχυτή και της τροφοδοσίας, παράγεται ένα μικρό ποσό ηλεκτρονικού θορύβου, ενώ παράλληλα υπάρχει όριο στο επίπεδο του σήματος που μπορούν να αντέξουν τα ηλεκτρονικά κυκλώματα, για τα οποία υπάρχουν αντίστοιχες πληροφορίες στα εγχειρίδια λειτουργίας των μικροφώνων. Επιπλέον, τα πυκνωτικά μικρόφωνα επηρεάζονται από τις μεταβολές της θερμοκρασίας και της υγρασίας, και είναι αρκετά ακριβά. Ωστόσο διαθέτουν υψηλή ευαισθησία, παράγουν καθαρό και δυνατό ηχητικό σήμα και παρουσιάζουν καλύτερη απόκριση συχνοτήτων, ιδίως στις υψηλές συχνότητες.



Εικόνα 1.22 Απεικόνιση ενός τυπικού πυκνωτικού μικροφώνου. (<https://mikrofon-tests.de>)

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Η ευαισθησία ενός πυκνωτικού μικροφώνου προσδιορίζεται από τον τύπο

$$S_m = \frac{a^2 V_s}{8 T t X_0} \quad (1.12)$$

όπου

a : ακτίνα (m)

V_s : τάση τροφοδοσίας (V)

T : μηχανική τάση (N/m²)

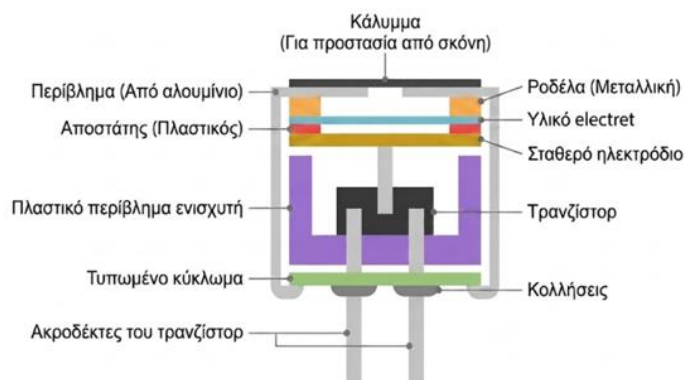
t : πάχος διαφράγματος (m)

X_0 : απόσταση μεταξύ των πλακών (m)

Η θεμελιώδης συχνότητα συντονισμού του διαφράγματος, για δεδομένη πυκνότητα ρ του υλικού προσδιορίζεται από τον τύπο

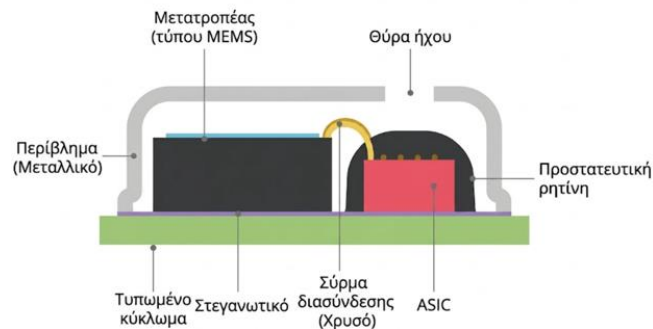
$$f_r = \frac{0,383}{a} \sqrt{\frac{T}{\rho}} \quad (1.13)$$

Μια εξέλιξη στα μικρόφωνα πυκνωτή αποτελεί η προσθήκη ενός στρώματος διηλεκτρικού στο διάφραγμα, το οποίο είναι φορτισμένο με ηλεκτρικό φορτίο. Τα συγκεκριμένα μικρόφωνα ονομάζονται ηλεκτρές και το κύριο χαρακτηριστικό τους είναι ότι δεν απαιτούν εξωτερική τροφοδοσία για τη λειτουργία τους. Στην πραγματικότητα, ωστόσο, απαιτείται μια μικρή τάση τροφοδοσία 5V για την λειτουργία των ηλεκτρονικών τους κυκλωμάτων. Το πιο συνηθισμένο διηλεκτρικό που χρησιμοποιείται είναι το Teflon ($\epsilon_r=2$), το οποίο παρουσιάζει επιφανειακή πυκνότητα φορτίου 10^{-4} C/m². Το μικρόφωνο ηλεκτρές μπορεί να διατηρήσει το φορτίο του για πολύ μεγάλο χρονικό διάστημα. (Λουτρίδης, 2024)

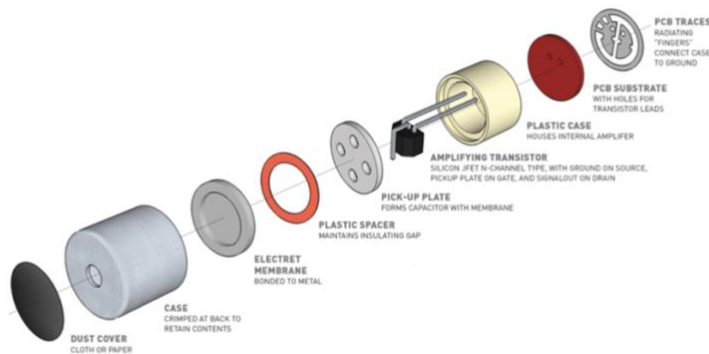


Εικόνα 1.23 Απεικόνιση της εσωτερικής δομής ενός ηλεκτρές μικροφώνου. (www.arrow.de)

Τα ηλεκτρες μικρόφωνα μπορούν να κατασκευαστούν σε εξαιρετικά μικρό μέγεθος, σύμφωνα με την τεχνολογία MEMS (Micro-Electro-Mechanical Systems), και βρίσκουν εφαρμογή σε φορητές συσκευές όπως τα Tablets και τα κινητά τηλέφωνα (Λουτρίδης, 2024).



Εικόνα 1.24 Απεικόνιση της εσωτερικής δομής ενός μικροφώνου τεχνολογίας MEMS. (www.arrow.de)



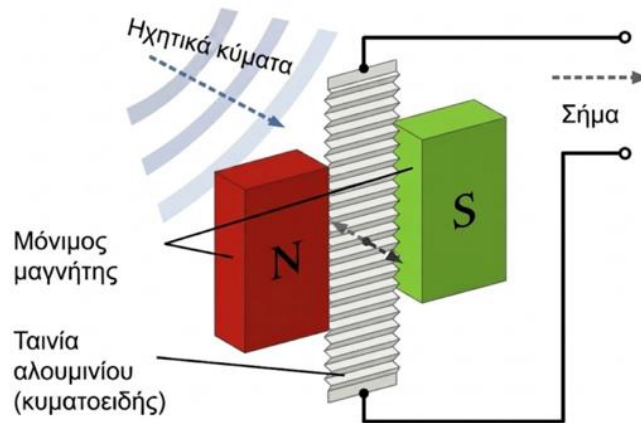
Εικόνα 1.25 Απεικόνιση των τμημάτων ενός μικροφώνου τεχνολογίας MEMS. (<https://elektro.turanis.de>)



Εικόνα 1.26 Απεικόνιση ενός τυπικού μικροφώνου τεχνολογίας MEMS. (<https://elektro.turanis.de>)

3. Μικρόφωνα Ταινίας

Τα μικρόφωνα ταινίας αποτελούνται από μια ελαφριά, λεπτή και πτυχωμένη μεταλλική ταινία, η οποία διαρρέεται από ηλεκτρικό ρεύμα. Η μεταλλική ταινία αυτή αιωρείται εντός μαγνητικού πεδίου και λειτουργεί ως διάφραγμα. Κατά τη διέγερση από ηχητικό σήμα, ταλαντώνεται και κινείται μέσα στο μαγνητικό πεδίο στο οποίο βρίσκεται, παράγοντας ηλεκτρική τάση στα άκρα της.



Εικόνα 1.27 Απεικόνιση της εσωτερικής δομής ενός μικροφώνου ταινίας. (www.mikrofon.guru)

Τα μικρόφωνα ταινίας απαιτούν προενίσχυση, διότι παράγουν σήμα χαμηλής στάθμης, παρά το γεγονός ότι είναι αρκετά ευαίσθητα στις μεταβολές της ηχητικής πίεσης. Επιπλέον, παρουσιάζουν ικανοποιητική απόκριση στις υψηλές συχνότητες και στα μεταβατικά σήματα. Ωστόσο, εμφανίζουν το μειονέκτημα, ότι είναι ιδιαίτερα ευαίσθητα σε κραδασμούς και φυσήματα.



Εικόνα 1.28 Απεικόνιση ενός τυπικού μικροφώνου ταινίας. (<https://www.hollyland.com>)

1.3.3 Χαρακτηριστικά μικροφώνων

Κάθε μικρόφωνο προσδιορίζεται από ορισμένα τεχνικά χαρακτηριστικά, τα οποία καθορίζουν το είδος των εφαρμογών στις οποίες μπορεί να χρησιμοποιηθεί. Ορισμένα από τα βασικότερα χαρακτηριστικά αναλύονται στην συνέχεια.

Η **κατευθυντικότητα** ή **πολική απόκριση** του μικροφώνου αποτελεί ένα μέγεθος που εκφράζει το ποσοστό μεταβολής της ευαισθησίας του σε συνάρτηση με τη διεύθυνση προέλευσης του λαμβανόμενου ηχητικού σήματος. Συγκεκριμένα, η κατευθυντικότητα ορίζεται ως ο λόγος της ευαισθησίας $M(\theta)$ για μια δεδομένη γωνία θ σε σχέση με τη μέγιστη ευαισθησία M_0 , και περιγράφεται από τον τύπο

$$D = \frac{M(\theta)}{M_0} \quad (1.14)$$

Η πολική απόκριση κάθε μικροφώνου αποτυπώνεται σε ένα πολικό διάγραμμα, το οποίο αποτελεί συνήθως τρισδιάστατη απεικόνιση. Στα πολικά διαγράμματα αυτά, το πολικό κέντρο συμπίπτει με το κέντρο του διαφράγματος, ενώ η διεύθυνση αναφοράς ορίζεται από μια ευθεία που περνάει από το κέντρο του και είναι κάθετη προς το διάφραγμα. Η κατευθυντικότητα ενός μικροφώνου εξαρτάται από τα χαρακτηριστικά της κατασκευής του.

Στην συνέχεια, παρουσιάζονται οι διάφορες κατηγορίες μικροφώνων με βάση την κατευθυντικότητά τους, όπως προκύπτει από τον τύπο που προαναφέρθηκε

α. Πανκατευθυντικό: $\frac{M}{M_0} = 1$

β. Δικατευθυντικό τάξης n: $\frac{M}{M_0} = (\cos \theta)^n$

γ. Μονοκατευθυντικό τάξης n: $\frac{M}{M_0} = (1 - \beta + \beta \cos \theta) (\cos \theta)^n, 0 < \beta < 1$

Η κατηγορία μονοκατευθυντικό με n=1 χωρίζεται σε επιπλέον υποκατηγορίες

γ1. Καρδιοειδές: $\beta = 0,5$ και $\frac{M}{M_0} = \frac{1}{2} + \frac{1}{2} \cos \theta$

γ2. Super-καρδιοειδές: $\beta = 0,67$ και $\frac{M}{M_0} = \frac{1}{3} + \frac{2}{3} \cos \theta$

γ3. Υπερκαρδιοειδές: $\beta = 0,75$ και $\frac{M}{M_0} = \frac{1}{4} + \frac{3}{4} \cos \theta$

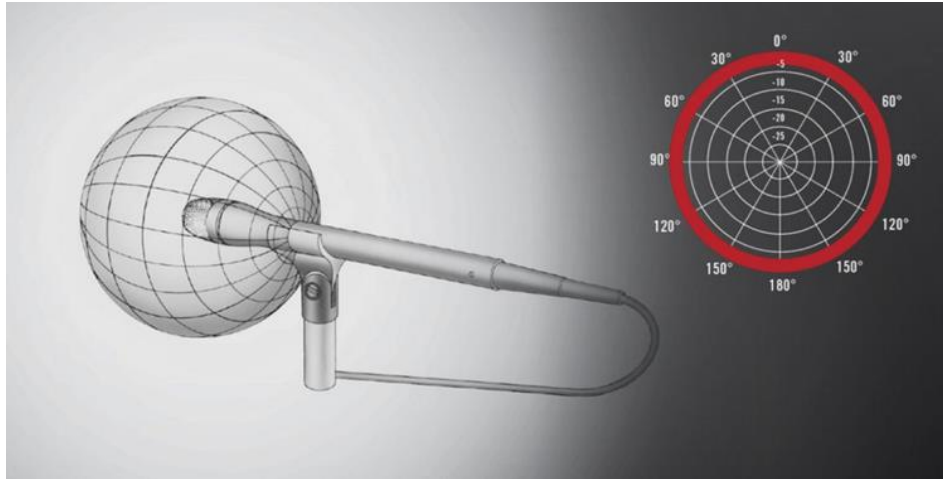
Ένα επιπλέον μέγεθος που σχετίζεται με την κατευθυντικότητα και κρίνεται σκόπιμο να αναφερθεί είναι ο παράγοντας απόστασης DF (Distance Factor). Το μέγεθος αυτό προσδιορίζει την μέγιστη απόσταση στην οποία δύναται να τοποθετηθεί ένα κατευθυντικό μικρόφωνο, ώστε η έξοδός του να είναι ίση με εκείνη ενός πανκατευθυντικού μικροφώνου σε μια ορισμένη γωνία. Ο παράγοντας απόστασης υπολογίζεται από τον παρακάτω τύπο.

$$DF = \sqrt{\frac{1}{D}} \quad (1.15)$$

α. Πανκατευθυντικό μικρόφωνο

Ένα πανκατευθυντικό μικρόφωνο παρουσιάζει την ίδια πολική απόκριση για όλες τις γωνίες πρόσπτωσης του ηχητικού σήματος. Στην πραγματικότητα, όμως, ισχύει μόνο για τις συχνότητες στις οποίες το $\lambda/2$ είναι μεγαλύτερο από την διάμετρο του διαφράγματος. Επιπλέον, στις πολύ υψηλές συχνότητες εμφανίζει μειωμένη ευαισθησία στην οπίσθια πλευρά του διαφράγματος, δηλαδή σε γωνία 180°, και κυρίως σε αυτές τις συχνότητες δημιουργείται ακουστική σκιά. Σε θεωρητικό επίπεδο, ο λόγος μεταξύ εμπρόσθιας προς οπίσθιας απόκρισης πρέπει να είναι ίσος με τη μονάδα.

Τα πανκατευθυντικά μικρόφωνα χαρακτηρίζονται από αυξημένη ευαισθησία. Για το λόγο αυτό, χρησιμοποιούνται κυρίως σε στούντιο και σε ελεγχόμενους ακουστικά χώρους, προκειμένου να αποφεύγονται οι ανεπιθύμητοι ήχοι που μπορεί να συλλάβουν τα μικρόφωνα λόγω της υψηλής ευαισθησίας τους.

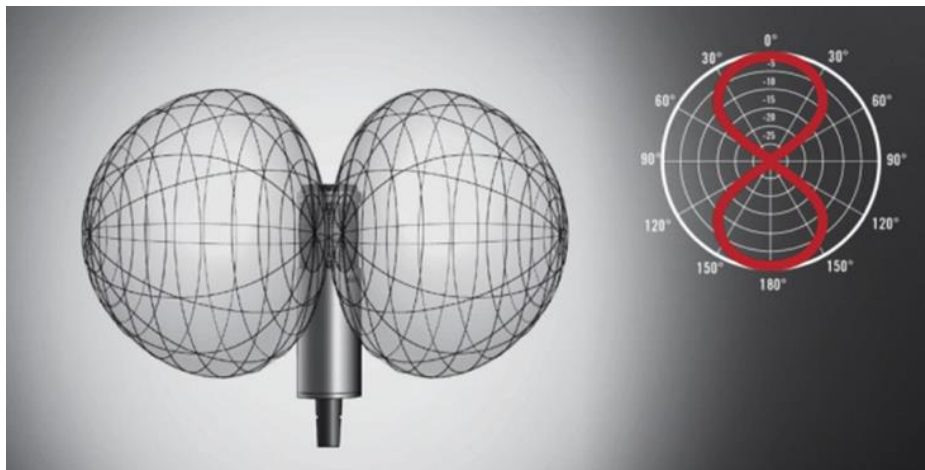


Εικόνα 1.29 Απεικόνιση πολικού διαγράμματος πανκατευθυντικού μικροφώνου. (www.stageaid.de)

β. Δικατευθυντικό μικρόφωνο

Τα δικατευθυντικά μικρόφωνα έχουν ως κύριο χαρακτηριστικό το οκτωειδές πολικό τους διάγραμμα, το οποίο δηλώνει ότι παρουσιάζουν μέγιστη ευαισθησία στις 0° και στις 180° μοίρες και ελάχιστη στις 90° μοίρες και από τις δύο πλευρές, πάντα σε σχέση με την διεύθυνση αναφοράς. Ωστόσο, στην πραγματικότητα, σχηματίζουν δύο ισοδύναμους λοβούς αναφορικά με το διάφραγμα, με εύρος γωνίας 90° για τον κάθε λοβό.

Τα δικατευθυντικά μικρόφωνα χρησιμοποιούνται για τη λήψη ηχητικών σημάτων από δύο διαφορετικές πηγές, οι οποίες βρίσκονται κυρίως η μια απέναντι από την άλλη. Αν και δεν παρουσιάζουν ευρεία χρήση, παρόλα αυτά όμως χρησιμοποιούνται σε συγκεκριμένες τεχνικές στερεοφωνίας, όπως για παράδειγμα η τεχνική M-S (Mid-Side).



Εικόνα 1.30 Απεικόνιση πολικού διαγράμματος δικατευθυντικού μικροφώνου. (www.stageaid.de)

γ. Μονοκατευθυντικό μικρόφωνο.

Το μονοκατευθυντικό μικρόφωνο παρουσιάζει μέγιστη ευαισθησία σε μια συγκεκριμένη διεύθυνση. Υπάρχουν τρία είδη μονοκατευθυντικών μικροφώνων: τα καρδιοειδές (cardioid), τα super-καρδιοειδές (super-cardioid) και τα υπερκαρδιοειδές (hyper-cardioid).

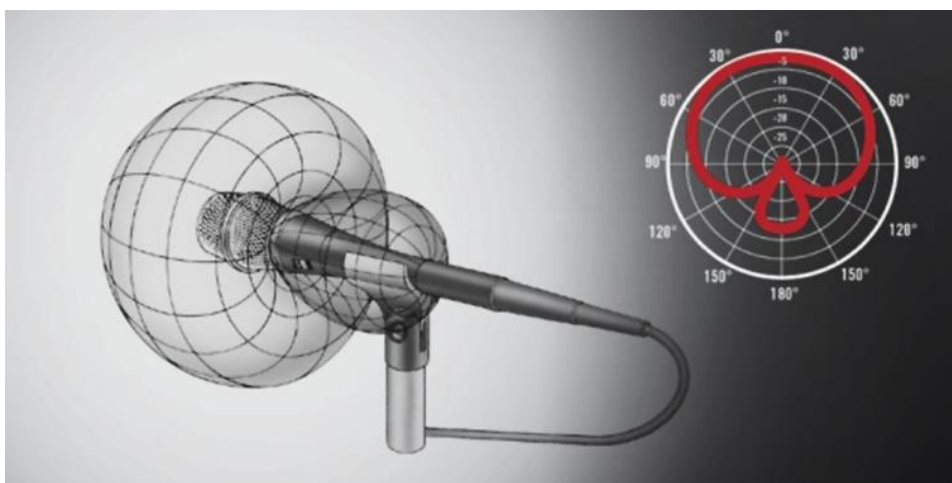
Τα καρδιοειδές μικρόφωνα παρουσιάζουν πολική απόκριση σε σχήμα καρδιάς, δηλαδή εμφανίζουν μέγιστη ευαισθησία στις 0° και ελάχιστη στις 180° . Στην πράξη, η ενεργή γωνία που μπορούν να κάνουν λήψη του ηχητικού σήματος είναι 130° .

Τα καρδιοειδή μικρόφωνα χρησιμοποιούνται σε εφαρμογές σε μη ελεγχόμενους ακουστικά χώρους καθώς έχουν τη δυνατότητα να απομονώνουν το επιθυμητό ηχητικό σήμα από τους ανεπιθύμητους ήχους, όπως ενδεικτικά είναι ο θόρυβος από το περιβάλλον.



Εικόνα 1.31 Απεικόνιση πολικού διαγράμματος μονοκατευθυντικού καρδιοειδές μικροφώνου. (www.stageaid.de)

Τα super-καρδιοειδές και τα υπερκαρδιοειδές μικρόφωνα είναι δύο τύποι μονοκατευθυντικών μικροφώνων, οι οποίοι παρουσιάζουν μικρότερη σε εύρος γωνίας λήψης για τα ηχητικά σήματα που λαμβάνουν, ωστόσο δίνουν μεγαλύτερη έμφαση στα ηχητικά σήματα που λαμβάνουν στις 0° . Συγκεκριμένα, η γωνία λήψης για τα super-καρδιοειδές είναι στις 115° και για τα υπερκαρδιοειδές είναι στις 105° . Επιπλέον, εμφανίζουν αυξημένη ευαισθησία και στις 180° , ενώ αντίστοιχα παρουσιάζουν ελάχιστη ευαισθησία, τα super-καρδιοειδές στις 126° και τα υπερκαρδιοειδές στις 110° .

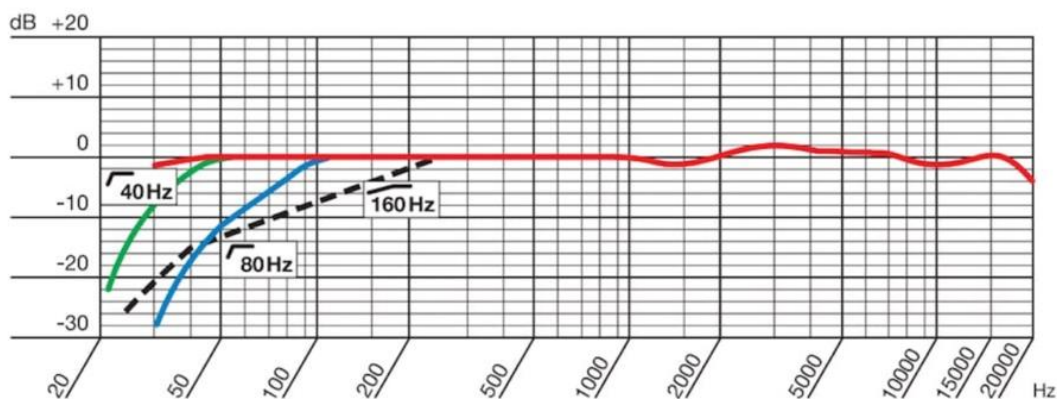


Εικόνα 1.32 Απεικόνιση πολικού διαγράμματος μονοκατευθυντικού super-καρδιοειδές και του υπερκαρδιοειδές μικροφώνου. (www.stageaid.de)

Η **απόκριση συχνότητας** ορίζεται ως το ηλεκτρικό σήμα εξόδου του μικροφώνου σε σχέση με τη μέγιστη ευαισθησία του μικροφώνου, η οποία συνήθως αναφέρεται στην μεσαία περιοχή των συχνοτήτων. Στην μεσαία αυτή περιοχή των συχνοτήτων, θεωρείται ότι η στάθμη είναι 0 dB και στα όρια του εύρους των συχνοτήτων θεωρείται ότι η στάθμη είναι -3 dB.

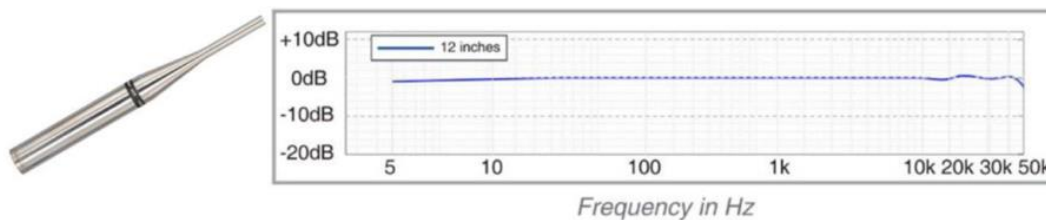
Στην πράξη, τα δυναμικά μικρόφωνα παρουσιάζουν τυπική απόκριση συχνότητας από 50 Hz έως και 10 kHz, ενώ τα πυκνωτικά μικρόφωνα εκτείνονται από 5 Hz έως 18 kHz. Ωστόσο, η καμπύλη της απόκρισης συχνότητας είναι επιθυμητό να είναι όσο το δυνατό πιο επίπεδη, με όσο το δυνατό λιγότερους συντονισμούς στο ακουστικό φάσμα. Στην πραγματικότητα όμως, τα μικρόφωνα παρουσιάζουν σε συγκεκριμένες περιοχές του φάσματος, διακύμανση της ευαισθησίας.

Το διάγραμμα της συχνотικής απόκρισης του μικροφώνου αποτελείται από μια γραφική παράσταση που παρουσιάζει κατά πόσο το μικρόφωνο επηρεάζει το εκάστοτε ηχητικό σήμα. Τα μικρόφωνα που έχουν επίπεδη απόκριση συχνότητας καλύπτουν ευρύτερο φάσμα συχνοτήτων, και δύναται να χρησιμοποιηθούν σε ποικίλες ηχητικές πηγές χωρίς να εισάγουν χρωματισμό. Αντίθετα όμως, ορισμένα μικρόφωνα έχουν διαμορφωμένη απόκριση συχνότητας, είτε ενισχυμένη απόκριση στις υψηλές συχνότητες για την αντιστάθμιση των απωλειών λόγω απόστασης από την πηγή του ήχου, είτε εξασθενημένη απόκριση στις χαμηλές συχνότητες με στόχο την αποφυγή ανεπιθύμητων υποηχητικών θορύβων και τις συνέπειες που θα επιφέρουν.



Εικόνα 1.33 Απεικόνιση της απόκρισης συχνότητας ενός τυχαίου μικροφώνου.

(<https://microfono.rocks>)



Εικόνα 1.34 Απεικόνιση της απόκρισης συχνότητας του μικροφώνου Earthworks M50.

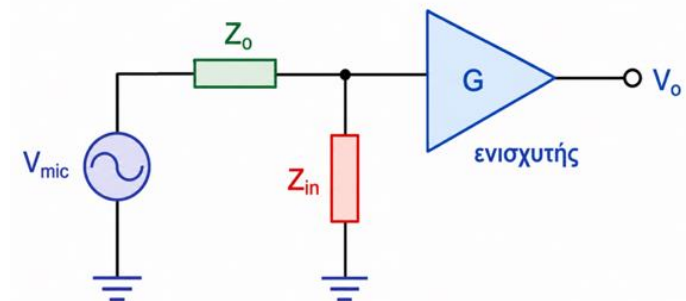
(<https://microfono.rocks>)

Η **ευαισθησία ή δυναμική απόκριση** ορίζεται ως η ικανότητα του μικροφώνου να καταγράφει είτε ασθενές ή σε μεγάλη απόσταση ηχητικά σήματα. Η δυναμική απόκριση του μετρείται σε mV/Pa για τη συχνότητα του 1 kHz. Ουσιαστικά εκφράζει τη μεταβολή της τάσης στην έξοδο του μικροφώνου για μεταβολή της ακουστικής πίεσης στο διάφραγμα του κατά 1 Pa.



Εικόνα 1.35 Απεικόνιση δύο επιπέδων της εξόδου του μικροφώνου σε συνάρτηση με την ευαισθησία. (www.sound-toa.com)

Η **αντίσταση εξόδου** του μικροφώνου προσδιορίζει τεχνικά τη δυνατότητα που έχει το μικρόφωνο να συνεργαστεί με την επόμενη βαθμίδα στο σύστημα ήχου. Βάση του ισοδύναμου κυκλώματος, το μικρόφωνο προσομοιώνεται σε μια πηγή τάσης σε σειρά με μια σύνθετη αντίσταση, όπως απεικονίζεται στο σχήμα που ακολουθεί.



Εικόνα 1.36 Απεικόνιση του ηλεκτρικού μοντέλου σύνδεσης μικροφώνου με ενισχυτή. (Λουτρίδης, 2024)

Η τάσης εξόδου του ενισχυτή για το λαμβανόμενο σήμα του μικροφώνου εκφράζεται από τον παρακάτω τύπο

$$V_o = G \frac{Z_{in}}{Z_o + Z_{in}} V_{mic} \quad (1.16)$$

όπου

G : κέρδος του ενισχυτή

Z_{in} : αντίσταση εισόδου ενισχυτή

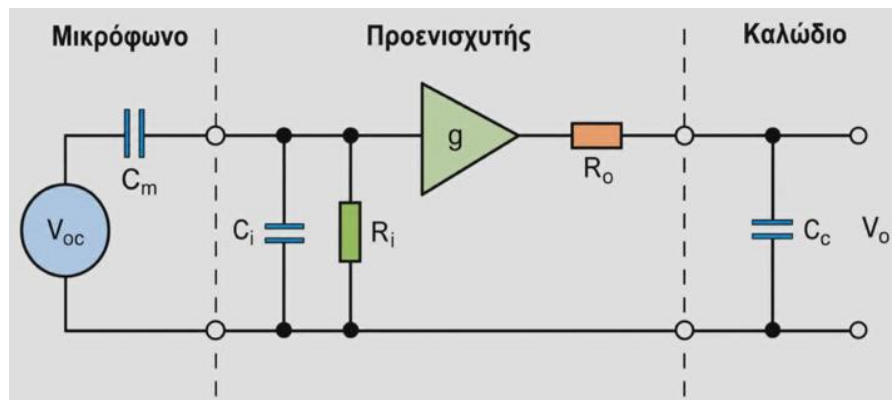
Z_o : αντίσταση εξόδου μικροφώνου

V_{mic} : τάση μικροφώνου

Προκειμένου να ελαχιστοποιηθούν οι απώλειες μεταξύ των βαθμίδων, απαιτείται να ικανοποιείται η σχέση $Z_{in} \gg Z_{mic}$. Στην πράξη όμως, ο προενισχυτής οφείλει να ενισχύσει το σήμα που λαμβάνει από το μικρόφωνο, με υψηλό συντελεστή ενίσχυσης, χωρίς παράλληλα να κάνει ενίσχυση και του θορύβου. Επιπλέον ο ενισχυτής, δεν πρέπει να παρέχει τροφοδοσία στο μικρόφωνο. Συνεπώς, για την επίτευξη βέλτιστης προσαρμογής, ο ενισχυτής οφείλει να έχει μεγαλύτερη αντίσταση εισόδου σε σχέση με την αντίσταση εξόδου του μικροφώνου. (Λουτρίδης, 2024)

a/a	Τύπος μικροφώνου	Συχνότητα	Αντίσταση εξόδου	Ενδιάμεση βαθμίδα
1	Δυναμικό μικρόφωνο	1 kHz	350Ω	Δεν χρειάζεται
2	Πυκνωτικά μικρόφωνα	Ανεξάρτητη συχνότητας	Μεγάλη τιμή	JFET σε συνδεσμολογία ακόλουθου πηγής
3	Μικρόφωνα ταινίας	Ανεξάρτητη συχνότητας	0,1Ω	Χρήση μετασχηματιστή

Πίνακας 1.1 Απεικόνιση τιμών αντίστασης εξόδου σε σχέση με τους τύπους των μικροφώνων.



Εικόνα 1.37 Απεικόνιση της αντίστασης εξόδου του μικροφώνου και σύνδεση με τις βαθμίδες που ακολουθούν. (Microphone Handbook, Brüel&Kjær, 2015)

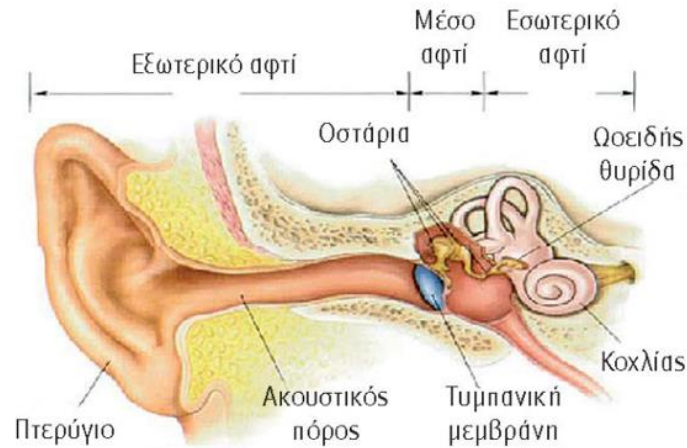
1.4 Αντίληψη του ήχου

Στην συνέχεια της παρούσας εργασίας παρουσιάζονται τα διάφορα τμήματα του αισθητηρίου οργάνου που διαθέτει ο άνθρωπος, για την ακοή. Ακολουθως, αναλύεται ο ρόλος του αισθητηρίου οργάνου της ακοής στη λήψη των ηχητικών σημάτων και την μετέπειτα μεταβίβαση στο κύριο ανθρώπινο όργανο, τον εγκέφαλο. Με σκοπό να γίνει καθορισμός και να κατανοηθεί ο τρόπος με τον οποίο αντιλαμβανόμαστε το ηχητικό σήμα.

1.4.1 Η ανατομία του ανθρώπινου οργάνου της ακοής

Το αισθητήριο όργανο που διαθέτει ο άνθρωπος για την ακοή διαιρείται ανατομικά σε τρία τμήματα, τα οποία είναι αντίστοιχα από έξω προς τα μέσα

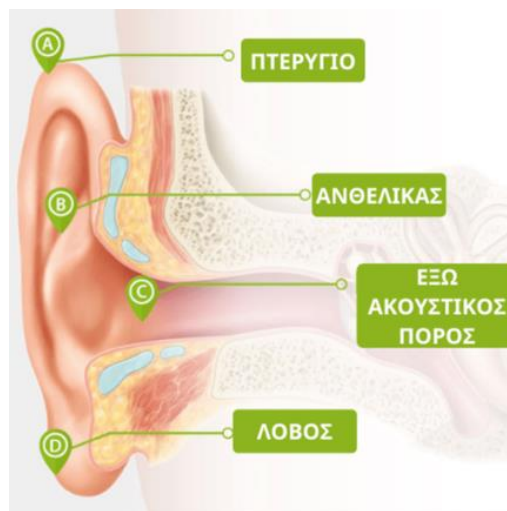
- α. Έξω ούς ή αυτί
- β. Μέσος ούς ή αυτί
- γ. Έσω ούς ή αυτί



Εικόνα 1.38 Απεικόνιση της τομής του αισθητηρίου οργάνου της ανθρώπινης ακοής (www.hygeia.gr)

α. Έξω αυτί

Το έξω αυτί αποτελείται από το πτερύγιο και τον ακουστικό πόρο. Το πτερύγιο είναι το τμήμα του αυτιού στην έξω πλαϊνή πλευρά του κεφαλιού. Το οποίο βοηθάει στην λήψη των ηχητικών κυμάτων αλλά και στην αντίστοιχη αντίληψη της κατεύθυνσης που έχουν τα ηχητικά κύματα που λαμβάνονται. Ανάλογα με την κατεύθυνση που λαμβάνονται τα ηχητικά κύματα και σε συνδυασμό με την ανάκλαση και την περίθλαση στο πτερύγιο διαμορφώνεται μια ιδιαίτερη ηχητική υπογραφή.



Εικόνα 1.39 Απεικόνιση του εξωτερικού τμήματος του ανθρώπινου αυτιού. (www.akoustikatheodorou.gr)

Στη συνέχεια, παρουσιάζεται ο ακουστικός πόρος, ο οποίος έχει μέση διάμετρο 0.7 εκατοστά και περίπου μήκος 3 εκατοστών. Η κύρια λειτουργία του ακουστικού πόρου είναι να οδεύσει τα κύματα του ήχου στην τυμπανική μεμβράνη και παράλληλα να τα ενισχύσει μέσω των συντονισμών που υφίστανται στον ηχητικό σωλήνα, κυρίως στην συχνοτική περιοχή από 2 kHz έως 5 kHz. Ωστόσο η ενίσχυση που συμβαίνει στον ηχητικό σωλήνα δεν είναι επαρκής, καθώς το ηχητικό κύμα μεταβαίνει από ένα μέσο χαμηλής ακουστικής αντίστασης, τον αέρα, σε ένα μέσο υψηλής ακουστικής αντίστασης, το υγρό του κοχλίου. Η προσαρμογή, σε αυτό το στάδιο, επιτυγχάνεται μέσω του μέσου αυτιού και από τα οστάρια που περιλαμβάνει.

β. Μέσο αυτί

Το μέσο αυτί αποτελείται από δύο τμήματα, τη τυμπανική μεμβράνη και τα τρία οστάρια, τη σφήρα, τον άκμονα και τον αναβολέα. Η τυμπανική μεμβράνη είναι σχήματος οβάλ με διάμετρο ογδόντα χιλιοστά. Η σφήρα, το πρώτο οστόριο, συνδέεται στην τυμπανική μεμβράνη. Ο αναβολέας, το τρίτο οστόριο συνδέεται στο οβάλ παράθυρο. Το οβάλ παράθυρο είναι μια μεμβράνη στο έσω τμήμα του αυτιού με μέγεθος της τάξης των τριών χιλιοστών.



Εικόνα 1.40 Απεικόνιση του μέσου τμήματος του ανθρώπινου αυτιού. (www.akoustikatheodorou.gr)

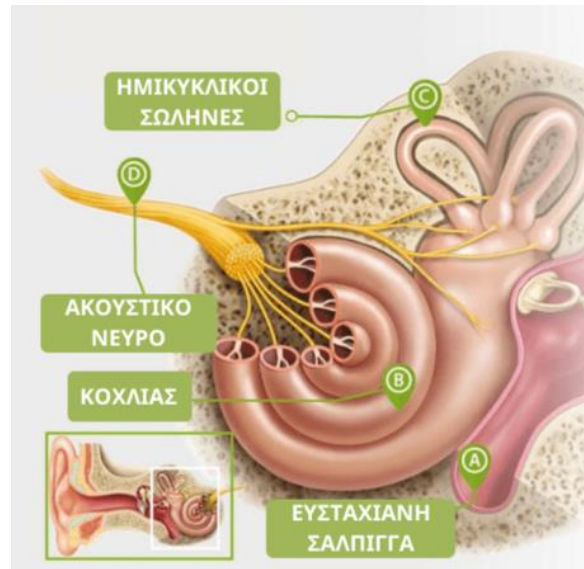
Η διαφορά της ακουστικής αντίστασης μεταξύ της τυμπανικής μεμβράνης και του οβάλ παράθυρου παράγει ένα μηχανικό κέρδος της τάξης των 27 φορές. Επιπλέον και η διαφορά στα μήκη των δύο οσταρίων, της σφήρας και του άκμονα, παρέχουν ένα επιπλέον κέρδος, εξαιτίας του μοχλισμού, της τάξης 1,3 με 3,1. Συνδυαστικά, το συνολικό κέρδος κυμαίνεται από τριανταπέντε έως και 84 φορές, το οποίο επιτυγχάνεται εξολοκλήρου στο μέσο αυτί. Αξίζει να σημειωθεί ότι η διαφορά ακουστικής αντίστασης του αέρα που διαδίδεται το ηχητικό κύμα και του υγρού του κοχλίου στο έσω αυτί είναι της τάξης 4000:1. Συνεπώς, η απευθείας λήψη των ηχητικών κυμάτων από τον κοχλία, θα είχε ως αποτέλεσμα την ανάκλαση του μεγαλύτερου ποσοστού της ηχητικής ενέργειας στη διεπιφάνεια των δύο αυτών μέσων. Παράλληλα, ένα σύστημα μυών ελέγχει την κίνηση της σφήρας και του άκμονα επιτρέποντας τη μεταβολή της ευαισθησίας. Το σύστημα αυτό των μυών χρησιμοποιείται κυρίως για την προστασία του κοχλίου από τις υψηλές στάθμες του εισερχόμενου ακουστικού σήματος, περίπου για άνω των 90 dB, μειώνοντας το μηχανικό κέρδος και τη μεταφορά των δονήσεων προς τον κοχλία, ώστε να προστατευτεί.

Τέλος, στο μέσο αυτί εντοπίζεται η ευσταχιακή σάλπιγγα, η οποία συνδέει το μέσο αυτί με το ρινοφάρυγγα. Η δομή αυτή επιτρέπει την εξισορρόπηση της πίεσης μεταξύ της τυμπανικής κοιλότητας και του εξωτερικού περιβάλλοντος.

γ. Εσωτερικό αυτί

Στο τελικό τμήμα εντοπίζεται το εσωτερικό αυτί, το οποίο αποτελείται από δύο μέρη τον κοχλία και τις απολήξεις του ακουστικού νεύρου. Ο κοχλίας έχει σπειροειδή μορφή και είναι γεμάτος από υγρό. Επιπλέον, ο κοχλίας διαχωρίζεται από δύο διαφορετικές μεμβράνες, την βασική μεμβράνη και την

μεμβράνη του Reissner. Οι δονήσεις που προκαλούνται από τα οστάρια του μέσου αυτιού μεταδίδονται στο υγρό του κοχλίου, προκαλώντας αντίστοιχα τη δόνηση της βασικής μεμβράνης. Σε αυτή την μεμβράνη εντοπίζονται 15000 εξωτερικά τριχίδια και 3500 εσωτερικά τριχίδια, όπου αντίστοιχα κάθε τριχίδιο έχει προσαρμοσμένο στην βάση του 40 έως 140 μετατροπείς. Οι μετατροπείς αυτοί, μετατρέπουν τις μηχανικές δονήσεις σε ηλεκτρικούς παλμούς, καθιστώντας τα εν λόγω τριχίδια ως βασικούς ηλεκτρομηχανικούς μετατροπείς του ακουστικού οργάνου.



Εικόνα 1.41 Απεικόνιση του εσωτερικού τμήματος του ανθρώπινου αυτιού.
(www.akoustikatheodorou.gr)

Η διαταραχή που προκαλείται από τα ηχητικά κύματα, διαδίδεται κατά μήκος της βασικής μεμβράνης του κοχλίου, και αναλόγως με την συχνότητα, εμφανίζει μέγιστο πλάτος ταλάντωσης σε διαφορετικό κάθε φορά σημείο της βασικής μεμβράνης. Με τον τρόπο αυτό, ο κοχλίας διαχωρίζει της συχνοτικές συνιστώσες που απαρτίζουν το ηχητικό κύμα, λειτουργώντας ως μηχανικός αναλυτής του ακουστικού φάσματος.

Στη συνέχεια, τα ηχητικά κύματα από τον κοχλίο μεταδίδονται στο ακουστικό νεύρο, με κωδικοποίηση μέσω μιας σειράς παλμών. Η στάθμη της ακουστικής πίεσης του ηχητικού κύματος κωδικοποιείται με ανάλογη αύξηση της συχνότητας μετάδοσης των παλμών αυτών. (Λουτρίδης, 2024)

1.4.2 Χρονική αντίληψη του ηχητικού σήματος

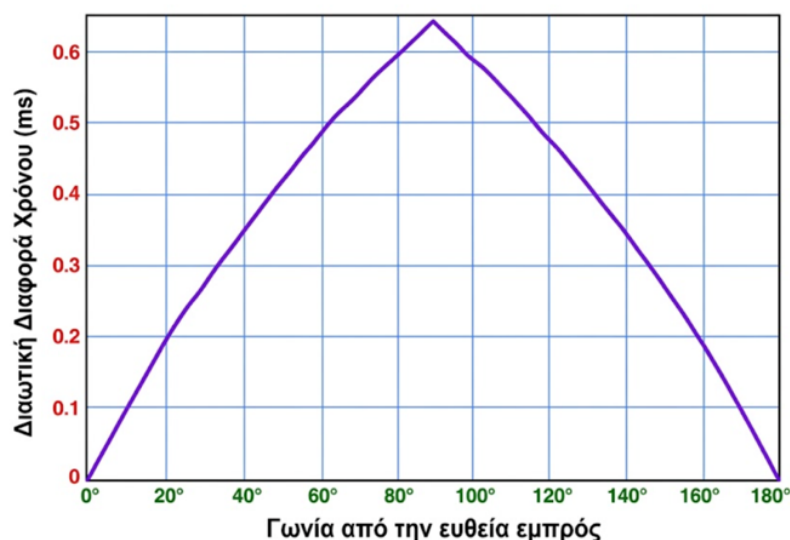
Η ικανότητα ακοής στο χρόνο, συσχετίζεται με την αντίληψη της μεταβολής στην ακουστική περιβάλλουσα, δηλαδή στην υποθετική καμπύλη που συνδέει τις κορυφές του σήματος σε σχέση με τη στάθμη της ακουστικής πίεσης του ηχητικού σήματος. Η αντιληπτική ικανότητα εντοπισμού της διαφοράς μεταξύ δύο ηχητικών σημάτων σύντομης διάρκειας είναι μεταξύ δύο με τριών μικροδευτερολέπτων. Τα σύντομα ηχητικά σήματα δεν γίνονται αντιληπτά ως δύο ξεχωριστά ηχητικά σήματα, αλλά ο ακροατής αντιλαμβάνεται ότι ο αρχικός ήχος έχει μεταβληθεί. Συνεπώς, η αντίληψη από τον ακροατή της χρονικής καθυστέρησης μεταξύ δύο ηχητικών σημάτων είναι εφικτή για καθυστέρηση μεγαλύτερη των δύο μικροδευτερολέπτων. (Λουτρίδης, 2024)

1.4.3 Αντίληψη της κατεύθυνσης ηχητικού σήματος

Η ακουστική αντίληψη του ανθρώπου βασίζεται στην αμφιωτική λειτουργία, η οποία διαδραματίζει καθοριστικό ρόλο για την αντίληψη της κατεύθυνσης του ηχητικού κύματος που λαμβάνει ο δέκτης. Συγκεκριμένα, ο ανθρώπινος εγκέφαλος, μέσω των ηχητικών σημάτων που λαμβάνονται στα δύο αυτιά, έχει την δυνατότητα να εκτιμά την ακριβείς θέση της ηχητικής πηγής στο χώρο. Για την εκτίμηση της θέσης μιας πηγής ήχου στο χώρο είναι κατά βάση υπεύθυνοι δύο μηχανισμοί, από την μια πλευρά είναι ο μηχανισμός που καθορίζει την χρονική διαφορά άφιξης του ηχητικού σήματος μεταξύ των δύο αυτιών και από την άλλη πλευρά είναι ο μηχανισμός που καθορίζει την διαφορά στη στάθμη της ακουστικής πίεσης του λαμβανόμενου ηχητικού σήματος στα δύο όργανα της ανθρώπινης ακοής. (Λουτρίδης, 2024)

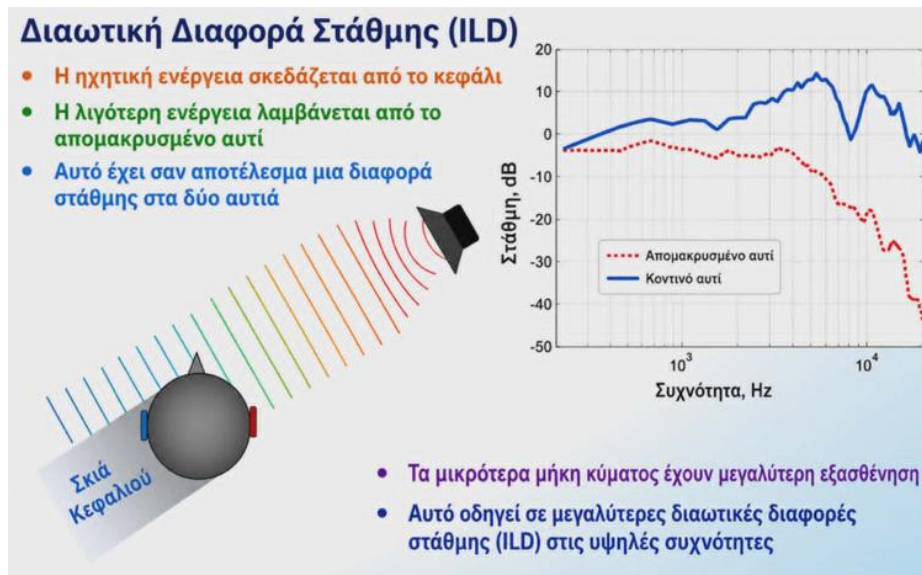
Ο μηχανισμός της ανθρώπινης ακοής, ο οποίος μπορεί να καθορίσει την χωρική αντίληψη των σημάτων ήχου, περιγράφεται στη δυική θεωρία (Duplex Theory) του Rayleigh. Σύμφωνα, με την οποία κάθε αυτί, βάση ότι έχει διαφορετική απόσταση από την πηγή ήχου, ανιχνεύει διαφορετικές ηχητικές τιμές πίεσης. (Φλώρος, 2003)

Η διαφορά που παρουσιάζεται στη χρονική άφιξη των ηχητικών κυμάτων στα δύο αυτιά, αποτελεί το μηχανισμό προσδιορισμού της θέσης της ηχητικής πηγής στις χαμηλές συχνότητες, οι οποίες συχνότητες ορίζονται κάτω από 1500 Hz, και ο οποίος ονομάζεται διαωτική διαφορά χρόνου (Interaural Time Difference, ITD). Η τιμή της διαωτικής διαφοράς του χρόνου παρουσιάζει άμεση σχέση με την θέση της ηχητικής πηγής, τη σκίαση που προκαλεί το κεφάλι, καθώς και τα χαρακτηριστικά του ίδιου του ηχητικού κύματος. Ο διαχωρισμός μεταξύ χαμηλών και υψηλών συχνοτήτων πραγματοποιείται με βάση τη μέση διάμετρο του κεφαλιού του ανθρώπου, η οποία μέση διάμετρος του κεφαλιού ορίζεται στα 23 εκατοστά. Στις χαμηλές συχνότητες, η στάθμη του ηχητικού σήματος παραμένει σχεδόν ίδια και στα δύο αυτιά, καθώς το μήκος κύματος του ηχητικού σήματος είναι αρκετά μεγαλύτερο από την μέση διάμετρο του ανθρώπινου κεφαλιού. Η διαωτική διαφορά χρόνου εκτείνεται από τα 0 ms για πηγές που βρίσκονται ευθεία εμπρός ή πίσω από το κεφάλι έως και τα 65 ms στην περίπτωση που η ηχητική πηγή βρίσκεται σε γωνία 90° σε συσχέτιση με τον άξονα αναφοράς. Κατά συνέπεια, στις χαμηλές συχνότητες αξιοποιείται η πληροφορία, η οποία αφορά την σχετική φάση. Αξίζει να σημειωθεί ότι το ανθρώπινο ακουστικό σύστημα έχει τη δυνατότητα εντοπισμού διαφορών της τάξης των 10 με 20 μ s, γεγονός που αποδίδεται σε γωνίες της τάξης της 1 με 2 μοιρών. (Λουτρίδης, 2024)

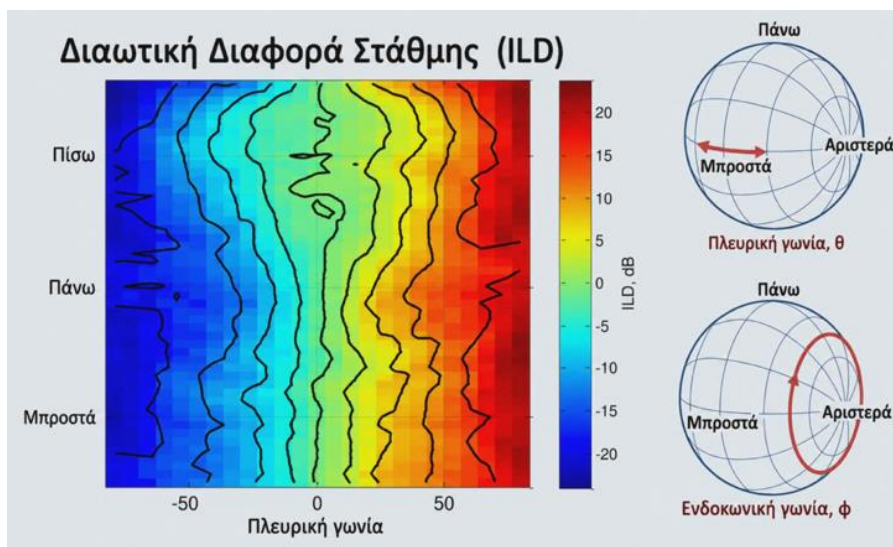


Εικόνα 1.42 Απεικόνιση της μεταβολής των τιμών ITD σε συνάρτηση με τη γωνία τοποθέτησης της πηγής ήχου. (Scharine, 2009)

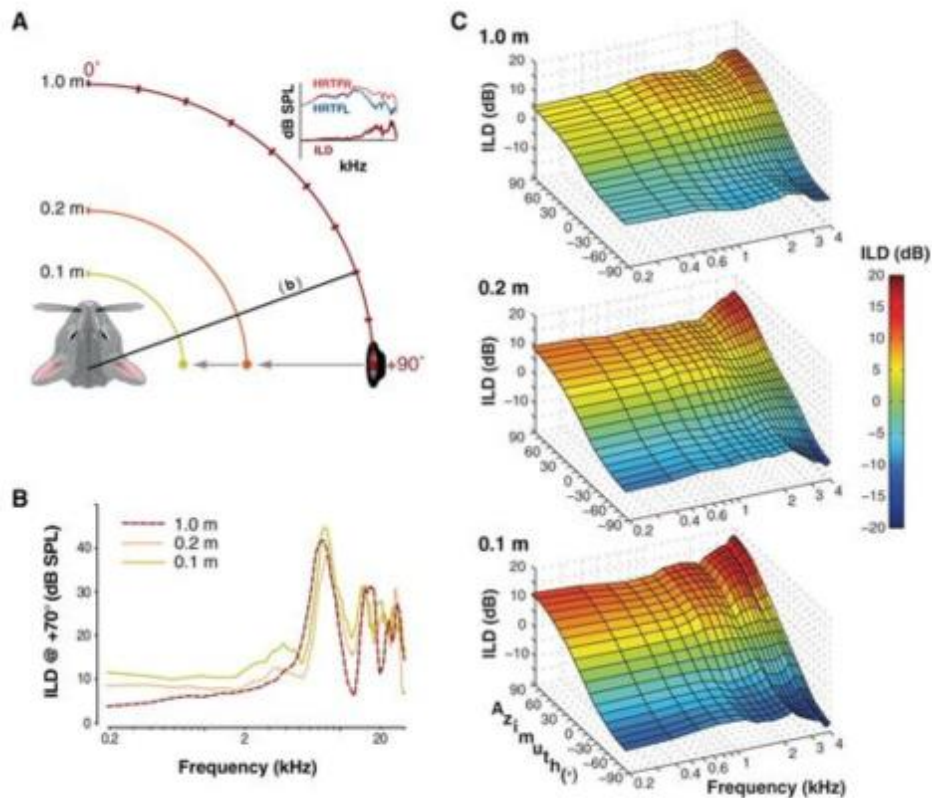
Η διαφορά της στάθμης των ηχητικών σημάτων, η οποία προκύπτει από τη μείωση της πίεσης ανάλογα με την απόσταση που έχει το κάθε αυτί από την ηχητική πηγή ονομάζεται διαωτική διαφορά στάθμης (Interaural Level Difference, ILD). Το φαινόμενο αυτό είναι κυρίαρχο στα ηχητικά σήματα των υψηλών συχνοτήτων, όπου τα μήκη κύματος είναι αρκετά μικρά συγκριτικά με το μέγεθος που έχει το ανθρώπινο κεφάλι. Λόγω της ακουστικής σκιάς που προκαλείται από το κεφάλι, οι διαφορές της ηχητικής στάθμης ενδέχεται να είναι μεγάλες, φτάνοντας μέχρι και 20 dB. (Μουρτζόπουλος, 2016)



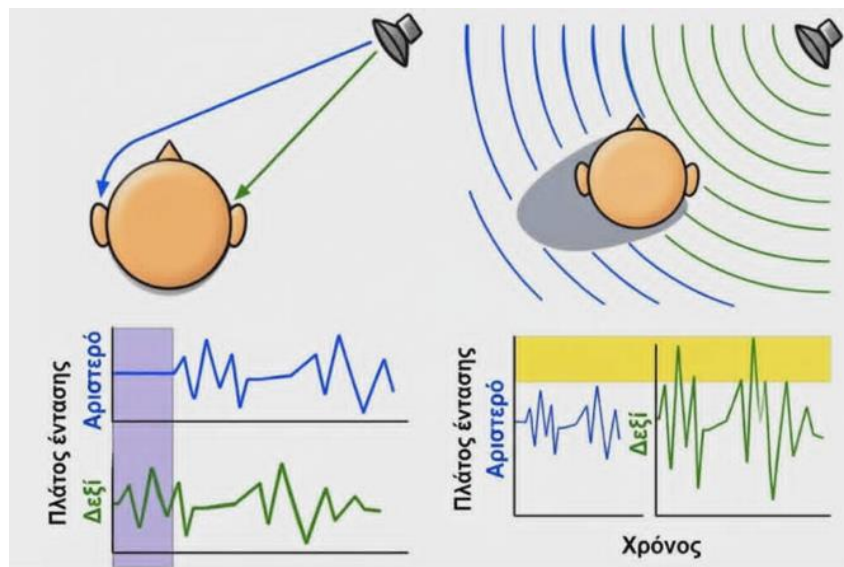
Εικόνα 1.43 Απεικόνιση του φαινομένου της ηχητικής σκιάς στις υψηλές συχνότητες. (Romigh, 2018)



Εικόνα 1.44 Απεικόνιση των διαφορών στην στάθμη εξαιτίας της ηχητικής σκιάς. (Romigh, 2018)



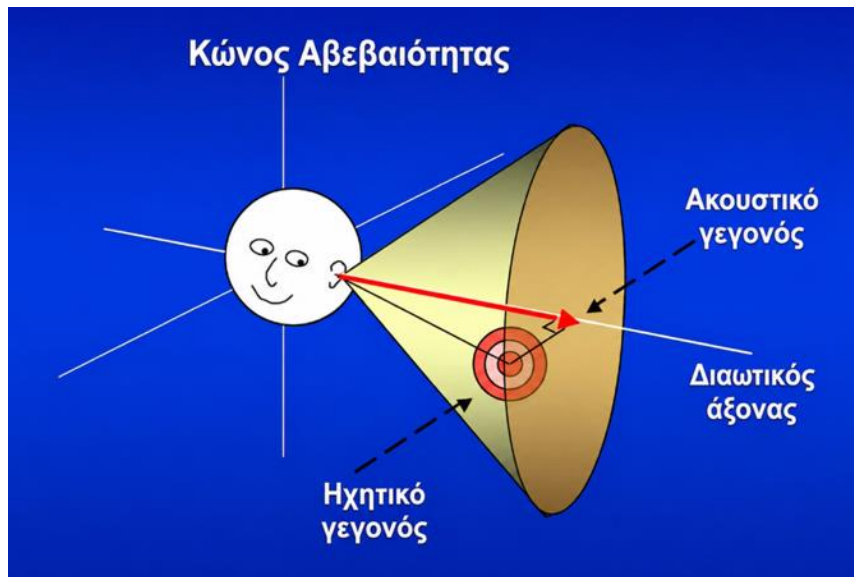
Εικόνα 1.45 Απεικόνιση των τιμών ILD σε σχέση με τη γωνία τοποθέτησης (αξιμούθιο) της πηγής ήχου, της συχνότητας που έχει το ηχητικό σήμα και τριών διαφορετικών αποστάσεων. (Koka, 2011)



Εικόνα 1.46 Σχηματική και γραφική απεικόνιση της διαωτικής διαφοράς χρόνου και της διαωτικής διαφοράς στάθμης του ηχητικού σήματος. (<https://www.researchgate.net>)

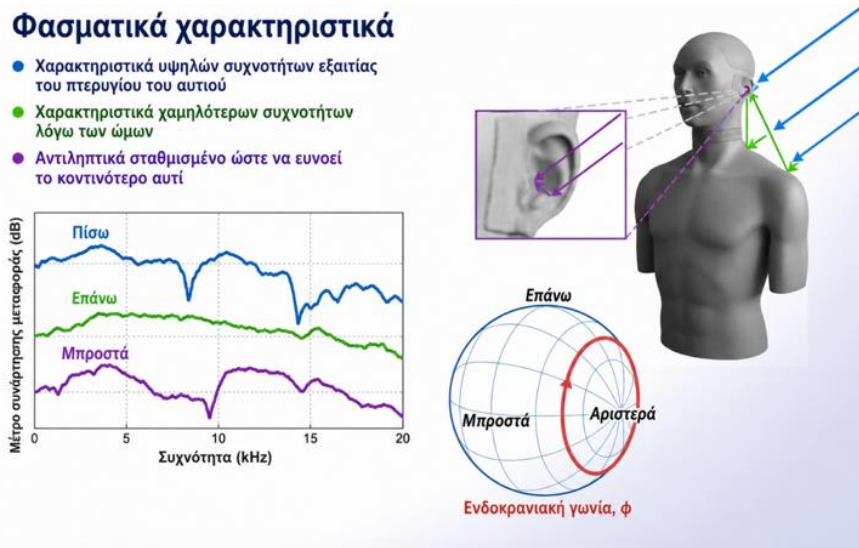
Με την εισαγωγή της δυικής θεωρίας διαωτικής ακρόασης (Duplex Theory) συνεκτιμάται αφενός η διαωτική διαφορά στάθμης (ILD) και αφετέρου η διαωτική διαφορά χρόνου (ITD) για την εκτίμηση του εντοπισμού της πηγής του ήχου. Συγκεκριμένα, η διαωτική διαφορά στάθμης αξιοποιείται για τον εντοπισμό υψηλών συχνοτήτων ηχητικών σημάτων και η διαωτική διαφορά χρόνου για τον εντοπισμό χαμηλών συχνοτήτων ηχητικών σημάτων. (Staab, 2023)

Ο εντοπισμός της πηγής του ήχου, από ανθρώπους που είναι ενήλικες, εμφανίζεται υποβαθμισμένος μεταξύ 2000 Hz και 4000 Hz και αντίστοιχα η διαωτική διαφορά στάθμης και η διαωτική διαφορά χρόνου παρουσιάζουν μειωμένη ευαισθησία σε αυτές τις συχνότητες. Συνεπώς, ένας συγκεκριμένος συνδυασμός τιμών ILD και ITD ενδέχεται να αντιστοιχεί σε πολλαπλές αζιμουθιακές θέσεις στο χώρο, το οποίο μπορεί να οδηγήσει σε εσφαλμένα αποτελέσματα. Το φαινόμενο αυτό είναι ευρέως γνωστό ως Κώνος Αβεβαιότητας (Cone of Confusion). Ενδεικτικά, ένα ηχητικό σήμα που προέρχεται από 45ο μοίρες αριστερά και μπροστά, με ένα ηχητικό σήμα που προέρχεται 45ο μοίρες αριστερά και πίσω δύναται να έχουν την ίδια τιμή για ITD. (Zhong, 2015)



Εικόνα 1.47 Απεικόνιση σχηματικά του κώνου αβεβαιότητας. (<https://blog.naver.com>)

Μια από τις βέλτιστες λύσεις για την αντιμετώπιση του προβλήματος του χωρικού εντοπισμού των ηχητικών σημάτων, όπως αυτό αποτυπώνεται μέσω του κώνου αβεβαιότητας, επιτυγχάνεται βάσει του αμφιωτικού μοντέλο χωρικής επεξεργασίας, με τις συναρτήσεις μεταφοράς σχετικές με το κεφάλι (Head Related Transfer Functions – HRTFs). Οι συναρτήσεις μεταφοράς σχετικές με το κεφάλι συμβάλουν στον ακριβή εντοπισμό της ηχητικής πηγής, λαμβάνοντας υπόψιν και ανατομικά κριτήρια, όπως την φυσιολογία κάθε ανθρώπου, το σχήμα και το μέγεθος του ανθρώπινου κεφαλιού, του προσώπου, του εξωτερικού μέρους του αυτιού αλλά και την θέση του σώματος, τους ώμους. (Andreopoulou & Katz, 2016)



Εικόνα 1.48 Απεικόνιση της επίδρασης των HRTFs από το άνω μέρος του σώματος. (Romigh, 2018)

1.5 Σύνοψη 1^{ου} Κεφαλαίου

Στο σημείο αυτό κρίνεται σκόπιμη μια ανακεφαλαίωση του πρώτου τμήματος της παρούσας εργασίας. Αρχικά, αναλύθηκε η έννοια του ήχου σε θεωρητικό επίπεδο και εξετάστηκαν τα διάφορα φαινόμενα που σχετίζονται με αυτόν. Ο ήχος είναι το κεντρικό αντικείμενο μελέτης όλης της εργασίας, όπως και τα διάφορα ηχητικά φαινόμενα, τα οποία είναι σημαντικά διότι μέσω αυτών μπορούμε να επεξεργαστούμε, να αναλύσουμε και να αλληλοεπιδράσουμε με τις διάφορες ακουστικές εφαρμογές, όπου στην περίπτωση της παρούσας εργασίας, είναι ο εντοπισμός της ηχητικής πηγής.

Στην συνέχεια, πραγματοποιήθηκε αναφορά στο θορύβου και στην επίδραση του στις ανθρώπινες δραστηριότητες. Κρίθηκε σκόπιμο, να μελετηθούν οι περιπτώσεις τόσο του θορύβου όσο και του ήχου, με την παράλληλη παράθεση, συζήτηση και ανάλυση των επιπτώσεων τους στον άνθρωπο. Ειδικότερα, αναφορικά με την παρούσα εργασία, η δομή και η ανάλυση επικεντρώνονται στη σχέση των φαινομένων αυτών με τον άνθρωπο. Τέλος, όσο αφορά τον θόρυβο, παρουσιάζονται τόσο οι αρνητικές επιπτώσεις που επιφέρει στις ανθρώπινες δραστηριότητες όσο και οι τρόποι με τους οποίους αυτός δύναται να γίνει ωφέλιμος για αυτές.

Στην συνέχεια του συγκεκριμένου τμήματος της εργασίας, εξετάστηκαν δύο σημαντικά σημεία λήψης και επεξεργασίας του ηχητικού σήματος. Συγκεκριμένα, αναλύεται το μικρόφωνο, το οποίο διαδραματίζει σημαντικό ρόλο στην συνέχεια της μελέτης, καθώς και το ανθρώπινο αισθητήριο της ακοής. Όσο αφορά το μικρόφωνο, πραγματοποιήθηκε μια εκτενής θεωρητική επισκόπηση, αλλά έγινε και εξέταση πρακτικών ζητημάτων, τα οποία σχετίζονται με αυτό. Μέσα από τη θεωρητική ανάλυση αναδείχθηκε η τεχνολογική εξέλιξη του μικρόφωνου και η μετάβαση από συσκευή μεγάλους μεγέθους σε σύγχρονες πολύ μικρές συσκευές, οι οποίες βρίσκουν εφαρμογή στα κινητά τηλέφωνα και σε άλλα φορητές συσκευές. Επιπλέον, υπογραμμίστηκε η αναγκαιότητα διεξαγωγής της απαραίτητης μελέτης για την επιλογή του κατάλληλου τύπου μικροφώνου ανάλογα με την εκάστοτε εφαρμογή. Στην περίπτωση της παρούσας εργασίας, τα κριτήρια επιλογής του μικροφώνου εστιάστηκαν στην ευχρηστία του για τις ανάγκες του πειραματικού μέρους, καθώς και στην άμεση εμπορική του διαθεσιμότητα. Επιπρόσθετα, βασικό κριτήριο επιλογής αποτέλεσε η αποφυγή υψηλής ευαισθησίας, καθώς η ανάπτυξη του πειραματικού μέρους δεν καθιστούσε αναγκαίο ένα τέτοιο χαρακτηριστικό. Εντούτοις, αναλυτικότερη παρουσίαση του χρησιμοποιούμενου μικροφώνου θα πραγματοποιηθεί σε επόμενο τμήμα της παρούσας εργασίας.

Στην συνέχεια, πραγματοποιήθηκε θεωρητική ανάλυση του ανθρώπινου αισθητήριου της ακοής. Συγκεκριμένα, αναλύθηκαν τα τμήματα του αυτιού και ο τρόπος συνεργασία τους για την επίτευξη της ακουστικής αντίληψης από τον άνθρωπο. Κρίθηκε σκόπιμο, παρότι δεν παρουσιάζει άμεση εφαρμογή στο πειραματικό μέρος της παρούσας εργασίας, να γίνει αναφορά στο ανθρώπινο αισθητήριο της ακοής, ώστε μέσα από το σύνολο της εργασίας να εξαχθούν σφαιρικά συμπεράσματα, και συγκρίσιμα αποτελέσματα μεταξύ του βασικού ηχητικού αισθητηρίου, του μικροφώνου και του αντίστοιχου ανθρώπινου αισθητηρίου, του αυτιού.

Στο καταλυτικό τμήμα αυτής της ενότητας, αναπτύχθηκε σε θεωρητικό επίπεδο το πεδίο της ηχητικής αντίληψης, με έμφαση στη χρονική και τη χωρική αντίληψη του ηχητικού σήματος. Τα φαινόμενα αυτά θα αναλυθούν εις βάθος σε επόμενο τμήμα της παρούσας εργασίας.

Επιπλέον, επισημαίνεται ότι στο πρώτο μέρος της εργασίας παραλείφθηκε η αναφορά στον τρόπο παραγωγής του ηχητικού σήματος, δηλαδή στο ηχείο, καθώς και στο πεδίο των υπερήχων, παρότι πρόκειται να αξιοποιηθούν στην συνέχεια. Οι συγκεκριμένες παραλείψεις κρίθηκαν αναγκαίες, καθώς η εκτενής ανάλυση τους ξεφεύγει από το σκοπό, το πνεύμα και την φιλοσοφία της παρούσας εργασίας.

Κεφάλαιο 2

2. Ανάλυση των τεχνικών ηχητικού εντοπισμού

2.1 Εισαγωγή

Κατά την εξέλιξη της παρούσας εργασίας πραγματοποιείται μια θεωρητική ανασκόπηση για τις μεθόδους επίλυσης του προβλήματος του εντοπισμού της ηχητικής πηγής. Υπάρχουν πολλές μέθοδοι εντοπισμού, οι οποίες επιλέγονται ανάλογα με τα κριτήρια που τίθενται για κάθε εφαρμογή, όπως το κόστος και η πολυπλοκότητα υλοποίησης και κυρίως η επιθυμητή ακρίβεια για τον εντοπισμό της ηχητικής πηγής. Ένα επιπλέον σημαντικό κριτήριο αποτελεί η αποτελεσματικότητα της μεθόδου σε ακουστικές εφαρμογές σε περιβάλλον με θόρυβο.

2.2 Ιστορική αναδρομή

Ακολούθως, πραγματοποιείται μια σύντομη ιστορική αναδρομή, με σκοπό την πληρέστερη, αντικειμενικότερη και ολοκληρωμένη παρουσίαση των τεχνικών εντοπισμού ηχητικής πηγής. Ο Α.Μ.Μayer σχεδίασε την πρώτη συσκευή για ηχητικό εντοπισμό το έτος 1880. Η ηχητική συσκευή ονομαζόταν τοπόφωνο Mayer και χρησίμευε στη βελτίωση της πλοήγησης στην ομίχλη.



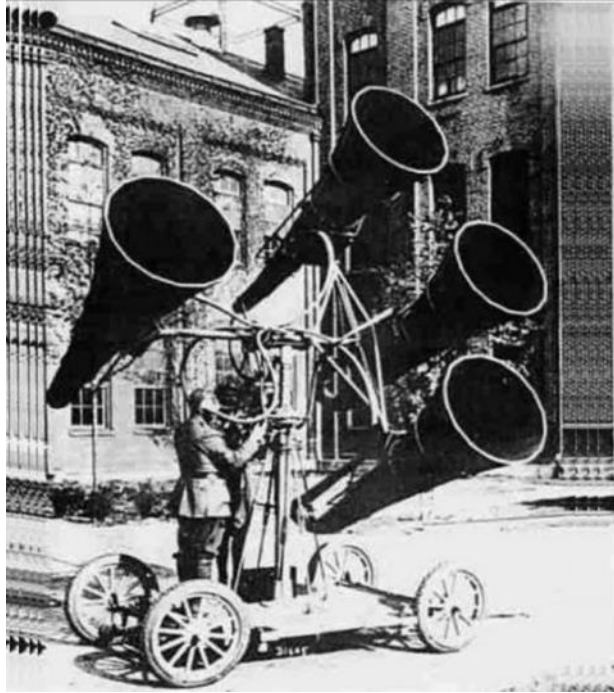
Εικόνα 2.1 Απεικόνιση του τοποφώνου Mayer. (<https://www.alamy.de>)

Στην συνέχεια όμως, εμφανίστηκαν πολλές παρόμοιες συσκευές βασισμένες στο τοπόφωνο Mayer, χωρίς όμως ιδιαίτερη πρακτική χρήση.

Το ενδιαφέρον για σχεδίαση και παραγωγή ηχητικών συστημάτων εντοπισμού, αναζωπυρώθηκε κατά το χρονικό διάστημα μεταξύ του Α΄ και του Β΄ παγκοσμίου πολέμου, με κύριο σκοπό την ανίχνευση και τον εντοπισμό του ήχου που εξέπεμπαν οι κινητήρες των πολεμικών αεροσκαφών. Τα δεδομένα που λαμβάνονταν σχετικά με τη θέση του αεροσκάφους, μεταφέρονταν απευθείας στο πυροβολικό αεράμυνας, το οποίο με την σειρά του μπορούσε να στοχεύσει πριν την οπτική επαφή. Η κατασκευή και οι διαστάσεις αυτών των συστημάτων ηχητικού εντοπισμού ήταν πολύ διαφορετικές, αλλά ωστόσο βασίζονταν στο τοπόφωνο Mayer, φέροντας βελτιώσεις με δύο επιπλέον κórνες σε κατακόρυφο επίπεδο προσανατολισμού. Για τη λειτουργία των συστημάτων αυτών απαιτούνταν τουλάχιστον δύο άτομα, προκειμένου να γίνει ανάλυση του ήχου, ο οποίος προερχόταν από το ηχητικό σύστημα εντοπισμού.

Λόγω της αδυναμίας συνεχούς μεγέθυνσης των διαστάσεων της κórνας για την επίτευξη μεγαλύτερου κέρδους, αναπτύχθηκαν τοίχοι και στατικά κάτοπτρα, τα οποία βασίζονταν σε σφαιρική επιφάνεια ανάκλασης. Αυτά τα στατικά συστήματα ανίχνευσης είχαν τη δυνατότητα ανίχνευσης αεροσκαφών σε μεγαλύτερες αποστάσεις. Ωστόσο, μετά την εφεύρεση του ραδιοεντοπιστή το 1934, οι συσκευές ηχητικού εντοπισμού δεν αναπτύχθηκαν περαιτέρω σε αυτό τον τομέα και αντικαταστάθηκαν πλήρως από τα συστήματα ραντάρ. (Dostalek, 2009)

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.



Εικόνα 2.2 Απεικόνιση συστήματος εντοπισμού ηχητικής πηγής που βασίζεται στο τοπόφωνο Mayer. (<https://uacgr.blogspot.com>)



Εικόνα 2.3 Απεικόνιση διάφορων διατάξεων στατικών συστημάτων ηχητικού εντοπισμού. (<https://weburbanist.com>)

2.3 Σύγχρονες μέθοδοι ηχητικού εντοπισμού

Στη σύγχρονη εποχή, τα συστήματα ηχητικού εντοπισμού έχουν εξελιχθεί σημαντικά, βρίσκοντας εφαρμογή σε ποικίλους τομείς της ανθρώπινης δραστηριότητας. Στο στρατιωτικό τομέα, τα συστήματα εντοπισμού ηχητικών πηγών αποτελούν ένα σημαντικό συμπλήρωμα στο σύστημα εντοπισμού θέσης με ραντάρ. Το πλεονέκτημα αυτών των συστημάτων είναι ότι παρέχουν υψηλότερη ασφάλεια, ως ένα παθητικό σύστημα εντοπισμού θέσης, το οποίο δεν παράγει τα δικά του ακουστικά κύματα. Στην αυτοκινητοβιομηχανία πραγματοποιούνται μελέτες με σκοπό ο υπολογιστής του αυτοκινήτου να προσδιορίζει αυτόματα τη θέση των φωνητικών εντολών, ώστε να γίνονται οι αντίστοιχες ενέργειες ανάλογα με τη θέση του εκάστοτε ομιλητή και τις αντίστοιχες φωνητικές εντολές που λαμβάνει. Στον τομέα της αεροδυναμικής, η χρήση τεχνολογιών ακουστικής απεικόνισης σε συνδυασμό με τον εντοπισμό ηχητικών πηγών παρέχει σημαντικά δεδομένα, υποστηρίζοντας κατάλληλα τις δοκιμές αεροσκαφών σε αεροδυναμικές σήραγγες.

Επιπλέον, οι τεχνικές και εφαρμογές εντοπισμού ηχητικών πηγών παρουσιάζουν αυξημένο ενδιαφέρον στην ανάπτυξη και στην έρευνα στον ακαδημαϊκό χώρο.

Στην συνέχεια, παρουσιάζονται οι βασικές κατηγορίες των τεχνικών ηχητικού εντοπισμού.

Κατηγορίες ηχητικού εντοπισμού

1. Μέθοδος εντοπισμού Ισχύος Λαμβανόμενου Σήματος – Received Signal Strength localization method (RSS)
2. Μέθοδος εντοπισμού Κατεύθυνσης Άφιξης – Direction Of Arrival localization method (DOA)
3. Μέθοδος εντοπισμού Γωνίας Άφιξης – Angle Of Arrival localization method (AOA)
4. Μέθοδος εντοπισμού Χρόνου Άφιξης – Time Of Arrival localization method (TOA)
5. Μέθοδος εντοπισμού Διαφοράς Χρόνου Άφιξης – Time Difference Of Arrival localization method (TDOA)
6. Μέθοδος εντοπισμού Διαφοράς Συχνότητας Άφιξης – Frequency Difference Of Arrival localization method (FDOA)
7. Μέθοδος εντοπισμού Πολλαπλής Ταξινόμησης Σημάτων – Multiple Signal Classification positioning method (MUSIC)
8. Μέθοδος Ενεργειακού Ηχητικού Εντοπισμού – Energy Based Localization method (EBL)

Επιπλέον, υπάρχουν ακόμα πολλές άλλες μέθοδοι που βασίζονται στις τεχνικές διαμόρφωσης δέσμης ή σε προηγμένες μεθόδους εντοπισμού με αξιοποίηση τεχνητών νευρωνικών δικτύων. (Zhao, 2023)

2.4 Μέθοδος εντοπισμού Ισχύος Λαμβανόμενου Σήματος (RSS)

Για τον εντοπισμό της θέσης της ηχητικής πηγής, μπορεί να εφαρμοστεί η μέθοδος ηχητικού εντοπισμού ισχύος του λαμβανόμενου σήματος (RSS). Η ισχύς ενός ηχητικού σήματος είναι ανάλογη με τη ισχύς που έχει το σήμα όταν εκπέμπεται από την ηχητική πηγή. Επιπλέον, η ισχύς μεταβάλλεται αντιστρόφως ανάλογα με το τετράγωνο της μεταξύ τους απόστασης. Σε αυτά τα φαινόμενα βασίζεται η μέθοδος εντοπισμού ισχύος λαμβανόμενου σήματος με σκοπό τον εντοπισμό της ηχητικής πηγής. Η ισχύς του ηχητικού σήματος λόγω της απόστασης, σε συνδυασμό με το άθροισμα των διανυσμάτων των ηχητικών κυμάτων μέσω των διαφορετικών διαδρομών από τις οποίες λαμβάνει το μικρόφωνο, καθορίζουν την εκτιμώμενη απόσταση μεταξύ της ηχητικής πηγής και του εκάστοτε μικροφώνου, καθώς και τη θέση της ηχητικής πηγής. Η εκτίμηση της απόστασης περιγράφεται με την εξίσωση που ακολουθεί

$$P_r(d) = \left(\frac{\lambda}{4\pi d} \right)^2 P_t G_t G_r \quad (2.1)$$

όπου

P_t : η ισχύς του εκπεμπόμενου σήματος

G_t : η ισχύς της ηχητικής πηγής

G_r : η ισχύς του σήματος που λαμβάνει το μικρόφωνο

λ : το μήκος κύματος του εκπεμπόμενου σήματος

Στην πραγματικότητα, η μετάδοση καθώς και η ισχύς του ηχητικού σήματος επηρεάζονται από ποικίλα ηχητικά φαινόμενα, όπως η ανάκλαση και η διάχυση. Για την μαθηματική ανάλυση της συγκεκριμένης μεθόδου, τα ηχητικά φαινόμενα μπορούν να προσεγγιστούν ως τυχαίες μεταβλητές. Στην περίπτωση αυτή και λαμβάνοντας υπόψη την παράμετρο της τυχαιότητας, η ισχύς του ηχητικού σήματος μειώνεται συναρτήσει της απόστασης, όπως παρουσιάζεται από την εξίσωση που ακολουθεί

$$P_r(d) = P_0(d_0) - n 10 \log_{10} \left(\frac{d}{d_0} \right) + X_0 \quad (2.2)$$

όπου

$P_r(d)$: η ισχύς του σήματος σε απόσταση d από την ηχητική πηγή

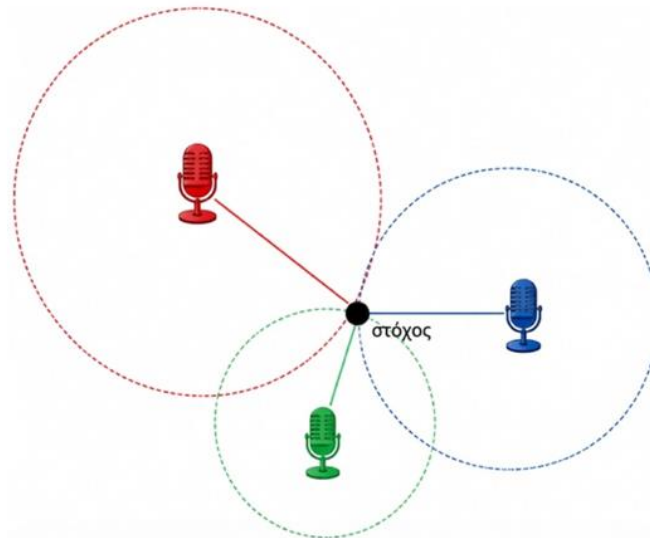
$P_0(d_0)$: η λαμβανόμενη ισχύς του ηχητικού σήματος σε μια απόσταση αναφοράς d_0

X_0 : μια τυχαία μεταβλητή σε λογαριθμική κλίμακα με διακύμανση σ^2

Η μέθοδος εντοπισμού ισχύος λαμβανόμενου σήματος δύναται να προσδιορίσει με ακρίβεια την απόσταση που έχει διανύσει το σήμα, υπό την προϋπόθεση ότι η ισχύς του σήματος της πηγής και οι απώλειες που έχει το σήμα στο μέσο διάδοσης, όπου θεωρείται κυρίως ο αέρας, είναι γνωστές. Ωστόσο, οι περιβαλλοντικές συνθήκες όπως η υγρασία και η θερμοκρασία, καθώς και τα ακουστικά φαινόμενα όπως η σκίαση, επηρεάζουν σημαντικά την ισχύ του λαμβανόμενου σήματος. Τα

φαινόμενα αυτά χαρακτηρίζονται από τυχαιότητα και μπορούν να οριστούν αποκλειστικά μέσω στατιστικών παραμέτρων.

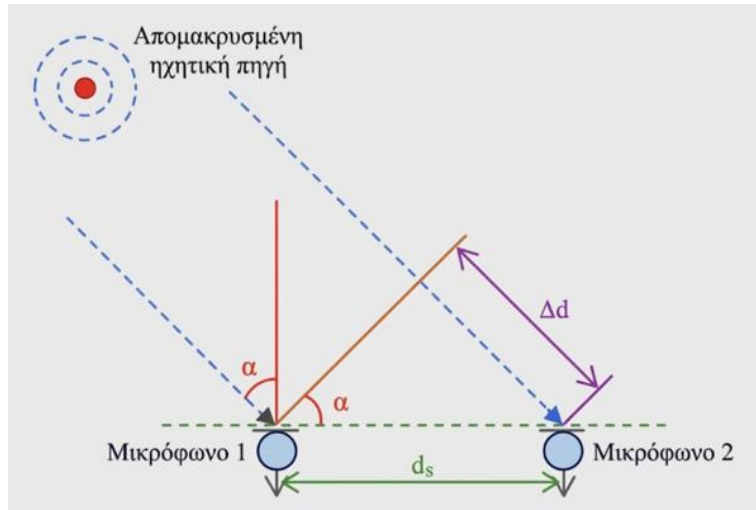
Επιπλέον, για τον εντοπισμό της θέσης της ηχητικής πηγής χρησιμοποιείται η τεχνική διακουστικής διαφοράς στάθμης (ILD), η οποία βασίζεται στη διαφορά της ενέργειας του ηχητικού σήματος στα μικρόφωνα. (Liu, 2010)



Εικόνα 2.4 Απεικόνιση της μεθόδου εντοπισμού ισχύς λαμβανόμενου σήματος (RSS) για τρία μικρόφωνα. (Yun, 2020)

2.5 Μέθοδος εντοπισμού Κατεύθυνσης Άφιξης (DOA)

Η μέθοδος εντοπισμού κατεύθυνσης άφιξης χρησιμοποιείται ευρέως σε συστήματα ακουστικών αισθητήρων, καθώς παρουσιάζει απλή εφαρμογή τόσο σε επίπεδο υλικού όσο και σε λογισμικού. Η μέθοδος αυτή βασίζεται κυρίως στην ανάλυση δειγμάτων ήχου, τα οποία λαμβάνονται από κάθε μικρόφωνο σε μια συστοιχία. Η πραγματοποίηση της ανάλυσης έχει δύο κύρια στάδια, αρχικά πραγματοποιείται ο προσδιορισμός της χρονικής καθυστέρησης για κάθε δείγμα δεδομένων που λαμβάνεται από τα διαφορετικά μικρόφωνα, και στη συνέχεια γίνεται ο υπολογισμός της θέσης της πηγής ήχου σε σχέση με τη γεωμετρική διάταξη της συστοιχίας μικροφώνων. Υποθέτοντας ότι η ηχητική πηγή βρίσκεται σε σημαντικά μεγαλύτερη απόσταση από την απόσταση μεταξύ των αισθητήρων - μικροφώνων d_s , με αποτέλεσμα η σφαιρική επιφάνεια του ηχητικού κύματος να δύναται να θεωρηθεί ως επίπεδη επιφάνεια. Η παραδοχή αυτή απλοποιεί την εκτίμηση της κατεύθυνσης του ηχητικού σήματος, όπως αυτή παρουσιάζεται στην εικόνα που ακολουθεί.



Εικόνα 2.5 Απεικόνιση της λήψης του ηχητικού σήματος από τα μικρόφωνα. (Dostalek, 2009)

Βάσει της προηγούμενης εικόνας, ο προσδιορισμός της γωνίας άφιξης α του ηχητικού σήματος στο ζεύγος των μικροφώνων πραγματοποιείται σύμφωνα με την παρακάτω εξίσωση

$$\alpha = \arcsin\left(\frac{\Delta d}{d_s}\right) \quad (2.3)$$

Επιπλέον, μπορεί να υπολογιστεί η διαφορά του ηχητικού σήματος Δd βάσει της εκτιμώμενης χρονικής καθυστέρησης t και της ταχύτητας του ήχου v . Στην περίπτωση αυτή, η γωνία άφιξης υπολογίζεται με την επόμενη εξίσωση

$$\alpha = \arcsin\left(\frac{tv}{d_s}\right) \quad (2.4)$$

Η εκτίμηση της χρονικής καθυστέρησης t πραγματοποιείται σύμφωνα με την ακόλουθη εξίσωση

$$R_{x_1x_2}[k] = \sum_{n=0}^{N-1} x_1[n]x_2[n+k] \quad (2.5)$$

όπου

x_1 : το σήμα στο μικρόφωνο 1

x_2 : το σήμα στο μικρόφωνο 2

N : ο αριθμός των δειγμάτων των δεδομένων

Για τον υπολογισμό της χρονικής καθυστέρησης, το σήμα x_1 , το οποίο αντιστοιχεί στο πρώτο μικρόφωνο, ορίζεται ως σήμα αναφοράς, ενώ το σήμα x_2 αντιστοιχεί στο δεύτερο μικρόφωνο και συγκρίνεται με το σήμα αναφοράς. Η ετεροσυσχέτιση των δύο μετατοπισμένων σημάτων δημιουργεί το συντελεστή συσχέτισης των δύο λαμβανόμενων σημάτων. Συνεπώς, η μέγιστη τιμή του

συντελεστή συσχέτισης και η μετατόπιση κατά k δείγματα δύναται να προσδιορίσουν τη σχετική μετατόπιση των σημάτων. Η χρονική μετατόπιση t υπολογίζεται βάσει του τύπου που ακολουθεί

$$t_{12} = T_s \arg_k \max(R_{x_1 x_2}[k]) \quad (2.6)$$

όπου

T_s : η περίοδος δειγματοληψίας

Για μεγάλο πλήθος δειγμάτων N , ο υπολογισμός της ετεροσυσχέτισης απαιτεί αυξημένη διάρκεια υπολογισμού, καθιστώντας τη συγκεκριμένη διαδικασία μη εφαρμόσιμη σε εφαρμογές που απαιτούν ταχεία απόκριση. Στις περιπτώσεις αυτές, γίνεται χρήση του μετασχηματισμού Fourier και πραγματοποιείται ο πολλαπλασιασμός των σημάτων στο πεδίο της συχνότητας, όπως παρουσιάζεται από τον τύπο που ακολουθεί

$$F\{x_1 * x_2\} = (F\{x_1\})^* F\{x_2\} \quad (2.7)$$

όπου

$*$: δηλώνει το μιγαδικό συζυγές

Στην συνέχεια εφαρμόζεται ο αντίστροφος μετασχηματισμός Fourier για την εξαγωγή του αποτελέσματος της ετεροσυσχέτισης, σύμφωνα με τον τύπο που ακολουθεί

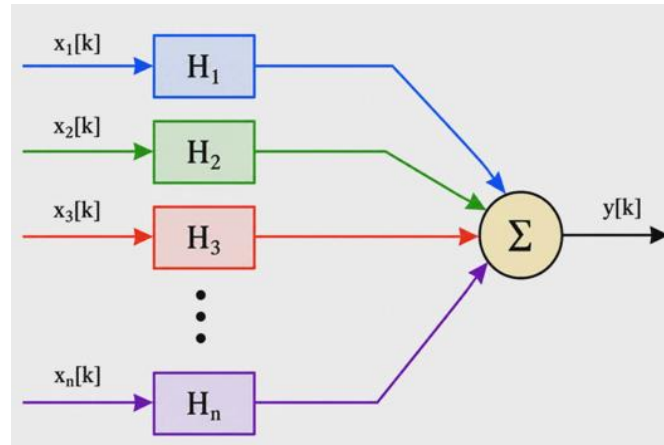
$$R_{x_1 x_2} = F^{-1}\{(F\{x_1\})^* F\{x_2\}\} \quad (2.8)$$

Η μέθοδος υπολογισμού κατεύθυνσης άφιξης δύναται να εφαρμοστεί με τον ίδιο τρόπο και για διατάξεις με περισσότερα από δύο μικρόφωνα. (Dostalek, 2009)

Ακολούθως, πραγματοποιείται μια συνοπτική παρουσίαση της μεθόδου διαμόρφωσης δέσμης, καθώς στην πλειονότητα των εφαρμογών πραγματοποιείται συνδυασμός των δύο αυτών μεθόδων.

Μέθοδος διαμόρφωσης δέσμης

Στην μέθοδο εντοπισμού μέσω διαμόρφωσης δέσμης, η οποία παρουσιάζεται στην εικόνα που ακολουθεί, τα σήματα $x[k]$ από την διάταξη μικροφώνων παρουσιάζουν χρονική καθυστέρηση στην είσοδο του συστήματος. Η χρονική αυτή καθυστέρηση εξαρτάται από τη γωνία κατεύθυνσης του κυρίου λοβού και τη γεωμετρική διαμόρφωση της διάταξης των μικροφώνων. Με την προϋπόθεση ότι τα σήματα από τα μικρόφωνα εμφανίζουν μόνο χρονική μετατόπιση, μπορούν να προσδιοριστούν οι κατάλληλες οριζόντιες καθυστερήσεις, οι οποίες συμβολίζονται με το γράμμα H σε κάθε κανάλι ήχου. Στην συνέχεια, πραγματοποιείται άθροιση για την επίτευξη του μέγιστου επιπέδου του ωφέλιμου σήματος, με παράλληλη βελτίωση του λόγου σήματος προς θόρυβο.



Εικόνα 2.6 Απεικόνιση της μεθόδου διαμόρφωσης δέσμης, με την λήψη των σημάτων, τις χρονικές καθυστερήσεις και έπειτα την άθροιση τους. (Dostalek, 2009)

Ακολούθως, για τον υπολογισμό της εξόδου $y[k]$ του διαμορφωτή δέσμης, εφαρμόζεται ο τύπος, ο οποίος παρουσιάζεται στην συνέχεια

$$y[k] = \sum_{m=1}^M x_m[k - H_m] \quad (2.9)$$

όπου

$x_m[k]$: σήμα εισόδου από το μικρόφωνο m

H : η εισαγόμενη καθυστέρηση από το διαμορφωτή σήματος

M : ο συνολικός αριθμός των μικροφώνων

Υπό την προϋπόθεση, ότι η ηχητική πηγή βρίσκεται σε σημαντικά μεγαλύτερη απόσταση σε σχέση με την απόσταση μεταξύ των μικροφώνων d_s , και ότι η διάταξη των μικροφώνων παρουσιάζει σχετική ομοιομορφία. Τότε η χρονική καθυστέρηση για την κατεύθυνση άφιξης α του ηχητικού σήματος υπολογίζεται από την εξίσωση που παρουσιάζεται στη συνέχεια

$$t_k = \frac{\sin \alpha d_s}{v} (k - 1) \quad (2.10)$$

όπου

k : ο δείκτης μικροφώνου

v : η ταχύτητα του ήχου στον αέρα

Το μικρόφωνο H_1 , ορίζεται ως μικρόφωνο αναφοράς, θεωρώντας ότι παρουσιάζει μηδενική χρονική μετατόπιση.

Επιπλέον, για τον εντοπισμό της ηχητικής πηγής εφαρμόζεται η μέθοδος καθυστέρησης και αθροίσματος του διαμορφωτή δέσμης. Ο υπολογισμός του αθροίσματος βασίζεται στη στάθμη του σήματος εξόδου, το οποίο παράγεται από τον διαμορφωτή δέσμη για κάθε πιθανή γωνία της κατεύθυνσης άφιξης.

Για τον υπολογισμό της ενεργούς τιμής RMS του σήματος εξόδου, εφαρμόζεται η εξίσωση που παρουσιάζεται στην συνέχεια

$$V_{RMS}[j] = \sqrt{\frac{1}{n} \sum_{k=1}^n y[k, j]^2} \quad (2.11)$$

όπου

n: αριθμός δειγμάτων

j: η γωνία αζιμούθιου

Ακολουθως, για τον υπολογισμό της μέγιστης γωνίας της ενεργούς τιμής RMS του σήματος εξόδου, εφαρμόζεται η εξίσωση που παρουσιάζεται στην συνέχεια. (Dostalek, 2009)

$$a = \arg_j \max(V_{RMS}[j]) \quad (2.12)$$

2.6 Μέθοδος εντοπισμού Γωνίας άφιξης (AOA)

Ο εντοπισμός της γωνίας άφιξης (AoA) αποτελεί μια τεχνική προσδιορισμού της θέσης μιας ακουστικής πηγής μέσω του υπολογισμού της κατεύθυνσης, από την οποία λαμβάνονται τα ηχητικά κύματα από μια προκαθορισμένη συστοιχία μικροφώνων. Η μέθοδος αυτή απαιτεί τη χρήση περισσότερων του ενός μικροφώνων, προκειμένου να καταστεί εφικτός ο προσδιορισμός της γωνίας άφιξης. Η επεξεργασία για τις γωνιακές πληροφορίες μπορεί να πραγματοποιηθεί στον ίδιο τον κόμβο, ενώ δεν απαιτείται ο ακριβής συγχρονισμός μεταξύ των κόμβων, καθώς η μέθοδος γωνίας άφιξης βασίζεται σε πληροφορίες κατεύθυνσης.

Η εκτίμηση της γωνίας άφιξης (AOA) επιτυγχάνεται μέσω μιας από τις τρεις μεθόδους που αναφέρονται στην συνέχεια

1. Η μέθοδος βασιζόμενη στην καθυστέρηση, στην οποία οι χρονικές διαφορές μεταξύ των ηχητικών σημάτων που λαμβάνονται από τα μικρόφωνα υπολογίζονται με τη χρήση της διασταυρούμενης συσχέτισης.
2. Η μέθοδος κατευθυνόμενης απόκρισης ισχύος (SRP), η οποία υλοποιείται με την προσθήκη καθυστέρησης μορφοποίησης δέσμης για μια διάταξη πάνω από τα καθορισμένα χωρικά σημεία ενδιαφέροντος, με σκοπό τη λήψη της κατανομής του λαμβανόμενου σήματος.
3. Η μέθοδος παραμετρικών αλγορίθμων, η οποία βασίζεται στη χρήση εξελιγμένων αλγορίθμων, όπως ο αλγόριθμος ταξινόμησης πολλαπλών σημάτων (MUSIC) ή ο αλγόριθμος Generalized EigenValue Decomposition MUSIC (GEVD-MUSIC), οι οποίοι αλγόριθμοι βελτιώνουν την ανθεκτικότητα ως προς το θόρυβο.

Ο αλγόριθμος της μεθόδου εντοπισμού της γωνίας άφιξης (AOA) βασίζεται στη χρονική καθυστέρηση των ηχητικών σημάτων, υπολογίζοντας τη χρονική διαφορά μεταξύ των σημάτων που λαμβάνονται

από τα μικρόφωνα. Η συγκεκριμένη μέθοδος χρησιμοποιεί τη διασταυρούμενη συσχέτιση με ή χωρίς χρήση συναρτήσεων στάθμισης. Η κορυφή της συνάρτησης συσχέτισης που προκύπτει αποτελεί το μέτρο της χρονικής καθυστέρησης Δt_{ij} μεταξύ των μικροφώνων i και j . Ο εντοπισμός της γωνίας άφιξης a_{ij} υπολογίζεται με βάση την εξίσωση που παρουσιάζεται στην συνέχεια

$$\sin(a_{ij}) = \frac{\Delta t_{ij} v_s}{d_{ij}} \quad (2.13)$$

όπου

d_{ij} : δηλώνει την απόσταση μεταξύ των μικροφώνων

v_s : δηλώνει την ταχύτητα του ήχου

Αρχικά, επισημαίνεται ότι η συνάρτηση διασταυρούμενης συσχέτισης R_{freq} υπολογίζεται στο πεδίο της συχνότητας, με βάση την εξίσωση που παρατίθεται στη συνέχεια.

$$R_{\text{freq}} = F\{s_i\} * F\{s_j\} \quad (2.14)$$

Επιπλέον, η συνάρτηση r_{freq} προκύπτει από τον αντίστροφο μετασχηματισμό Fourier του διασταυρούμενου φάσματος R_{freq} , το οποίο είναι η μιγαδική συζυγή συνάρτηση του μετασχηματισμού Fourier του σήματος s_i και η αντίστοιχη εξίσωση παρουσιάζεται στη συνέχεια.

$$r_{\text{freq}} = F^{-1}\{R_{\text{freq}}\} \quad (2.15)$$

Η κορυφή της συνάρτησης συσχέτισης προσδιορίζει τη χρονική καθυστέρηση Δt_{ij} .

Η διασταυρούμενη συσχέτιση δύναται να υπολογιστεί στο πεδίο της συχνότητας για το σύνολο των σημείων συσχέτισης των σημάτων εισόδου s_i και s_j . Η συνάρτηση συσχέτισης που προκύπτει με αυτή τη μέθοδο αποτελείται από $2N-1$ σημεία, όπου N ορίζεται ο αριθμός των δειγμάτων σε κάθε κανάλι μικροφώνου. Τα ωφέλιμα σημεία στη συνάρτηση διασταυρούμενης συσχέτισης n_s εξαρτώνται σε μεγάλο βαθμό από τη συχνότητα δειγματοληψίας f_s . Ένας επιπλέον καθοριστικός παράγοντας που επηρεάζει τη συγκεκριμένη τιμή είναι η απόσταση μεταξύ των μικροφώνων d_{ij} . Στην συνέχεια, παρατίθεται η εξίσωση για αυτά τα χρήσιμα σημεία της συσχέτισης

$$n_s = 2 \left\lceil d_{ij} \frac{f_s}{v_s} \right\rceil + 1 \quad (2.16)$$

Επιπλέον, επισημαίνεται ότι για χαμηλές συχνότητες δειγματοληψίας και μικρές αποστάσεις μεταξύ των μικροφώνων, καθίσταται δύσκολο να επιτευχθεί υψηλή τιμή για την παράμετρο n_s .

Για την επίλυση του συγκεκριμένου προβλήματος χρησιμοποιείται η διασταυρούμενη συσχέτιση στο πεδίο του χρόνου, καθώς επιτρέπει αποτελεσματικό υπολογισμό με ένα μικρό αριθμό σημείων συσχέτισης. Συγκεκριμένα, η μέθοδος αυτή δύναται να υπολογιστεί ως το τελικό γινόμενο των δύο σημάτων εισόδου s_i και s_j . Δεδομένου, ότι η συνάρτηση πιθανότητας πρέπει να είναι θετική, πραγματοποιείται διαχωρισμός σε θετικά και σε αρνητικά σήματα εισόδου s_i και s_j , όπως

αποτυπώνεται στις εξισώσεις που ακολουθούν. Στη συνέχεια ο μέσος όρος των αποτελεσμάτων των δύο γινομένων αποτελεί μια αποκλειστικά θετική συνάρτηση. Σε κάθε σημείο της διασταυρούμενης συσχέτισης m , πραγματοποιείται υπολογισμός μετατοπίζοντας το s_j πάνω από το s_i για m φορές. Ακολούθως, πραγματοποιείται ο υπολογισμός της εν δυνάμει παραγόμενης συσχέτισης $r_{\text{dotprod}}(m)$, όπου κάθε τιμή της διασταυρούμενης συσχέτισης διαιρείται με τον αριθμό των επικαλυπτόμενων δειγμάτων, τα οποία έχουν συμπληρωθεί με μηδενικά.

$$r_{\text{dotprod,max}}(m) = \sum_{n=0}^{N-1} s_i^+[n] * s_j^+[n+m] \quad (2.17)$$

$$r_{\text{dotprod,min}}(m) = \sum_{n=0}^{N-1} s_i^-[n] * s_j^-[n+m] \quad (2.18)$$

$$r_{\text{dotprod}}(m) = \frac{r_{\text{dotprod,max}}(m) + r_{\text{dotprod,min}}(m)}{2} \quad (2.19)$$

Για τον υπολογισμό του διασταυρούμενου φάσματος μετασχηματισμένης φάσης, το διασταυρούμενο φάσμα της τυπικής διασταυρούμενης συσχέτισης διαιρείται με την απόλυτη τιμή του, όπως αποτυπώνεται με τον παρακάτω τύπο

$$R_{\text{freq,PHAT}} = \frac{R_{\text{freq}}}{|R_{\text{freq}}|} \quad (2.20)$$

Στην περίπτωση που απαιτείται η εφαρμογή της ανωτέρω μεθόδου στο πεδίο του χρόνου, δύναται να πραγματοποιηθεί «λεύκανση» του ήχου εισόδου στο επίπεδο GCC-PHAT, διαδικασία η οποία επιτυγχάνεται με τη χρήση φίλτρου γραμμικής πρόβλεψης (LP). Ο κύριος στόχος είναι όλα τα ζεύγη συχνοτήτων να έχουν κέρδος ίσο με τη μονάδα, χωρίς να προκαλείται αλλοίωση στη φάση. Ως αποτέλεσμα, προκύπτει η GCC-PHAT στο πεδίο του χρόνου για τη μέθοδο της διασταυρούμενης συσχέτισης. (Theon, 2019)

Στην συνέχεια, παρουσιάζονται συνοπτικά ορισμένοι αλγόριθμοι της μεθόδου εντοπισμού της γωνίας άφιξης, με κύρια αναφορά στους αλγόριθμους τριγωνισμού και αντιστοίχισης.

2.6.1 Αλγόριθμος δισδιάστατου εντοπισμού θέσης με τριγωνισμό

Η εκτίμηση της θέσης της ηχητικής πηγής δύναται να πραγματοποιηθεί μέσω τριγωνισμού, με τη εφαρμογή της μεθόδου των ελαχίστων τετραγώνων στο σύνολο των λαμβανόμενων σημάτων $A_0 A_s(a_1, \dots, Q)$ για έναν συγκεκριμένο αριθμό ηχητικών αισθητήρων Q . Στη συνέχεια, η εκτιμώμενη θέση $\hat{e}$ ως προς όλους τους αισθητήρες, με συντεταγμένες $x_{1, \dots, Q}$ και $y_{1, \dots, Q}$, υπολογίζεται μέσω της εξίσωσης που παρουσιάζεται στην συνέχεια

$$\hat{e} = (H'H)^{-1}H'c \quad (2.21)$$

όπου

$$H = \begin{pmatrix} \sin(a_1) & -\cos(a_1) \\ \sin(a_2) & -\cos(a_2) \\ \vdots & \vdots \\ \sin(a_Q) & -\cos(a_Q) \end{pmatrix}$$

και

$$c = \begin{pmatrix} x_1 \sin(a_1) - y_1 \cos(a_1) \\ x_2 \sin(a_2) - y_2 \cos(a_2) \\ \vdots \\ x_Q \sin(a_Q) - y_Q \cos(a_Q) \end{pmatrix}$$

Ωστόσο, με την χρήση γραμμικής διάταξης αισθητήρων, ο διαχωρισμός της γωνίας πρόσπτωσης μεταξύ του πρόσθιου και του οπίσθιου μέρους της διάταξης είναι ιδιαίτερα δύσκολη. Η λύση του εν λόγω προβλήματος επιτυγχάνεται με περιορισμό στη λήψη των ηχητικών σημάτων από το πίσω μέρος.

Εντούτοις, η εφαρμογή του αλγορίθμου παρουσιάζει ένα σημαντικό μειονέκτημα, το οποίο είναι ότι μικρές αποκλίσεις στην εκτίμηση της γωνίας άφιξης δύναται να επιφέρουν μεγάλο σφάλμα στην εκτίμηση της θέσης.

2.6.2 Αλγόριθμος δισδιάστατου εντοπισμού θέσης με αντιστοίχιση

Με τον αλγόριθμο αυτόν, τα λαμβανόμενα δεδομένα σε κάθε κόμβο αισθητήρα αντιστοιχίζονται αρχικά σε ένα σύνολο δεδομένων αναφοράς. Ειδικότερα, η διαδικασία αντιστοίχισης αρχίζει με την καταγραφή των χωρικών διαστάσεων και ακολούθως, δημιουργεί ένα προϋπολογισμένο σύνολο δεδομένων αναφοράς T . Το σύνολο αυτό αποτελείται από τις συναρτήσεις πυκνότητας πιθανότητας PDFs ($f_{LOS,i}$) των πραγματικών γραμμών ακουστότητας (LOS) για ένα προκαθορισμένο σημείο

αναφοράς N_f . Στην περίπτωση αυτή, πραγματοποιείται κατανομή των θέσεων κατά μήκος ενός χωρικού πλέγματος, ενώ παράλληλα ορίζεται ένα σύνολο δεδομένων για κάθε κόμβο αισθητήρα στον χώρο. Κάθε LOS-PDF αποτελείται από 181 τιμές, οι οποίες αντιστοιχούν στο εύρος μιας διάταξης από -90° έως 90° με ανάλυση 1° . Η θέση της πραγματικής γωνίας λαμβάνει την τιμή ένα, ενώ τα υπόλοιπα σημεία μηδενίζονται. Προκειμένου να πραγματοποιηθεί η τελική αντιστοίχιση μιας μετρημένης PDF με κάθε σημείο αναφοράς $f_{LOS,i}$, απαιτείται η παρεμβολή της μετρούμενης συνάρτησης συσχέτισης. Η παρεμβολή των δεδομένων πραγματοποιείται με τη χρήση προσαρμοσμένης κυβικής παρεμβολής στην ίδια γωνιακή ανάλυση με τις PDFs στο σύνολο των δεδομένων αναφοράς, με την παραγόμενο PDF να συμβολίζεται εφεξής με m . Τέλος, η πραγματική αντιστοίχιση μεταξύ των $f_{LOS,i}$ και m υπολογίζεται με βάση τον συντελεστή συσχέτισης Pearson $r_{corr}(i)$ για τα σημεία αναφοράς N_f .

$$r_{corr}(i) = \frac{\text{cov}(f_{LOS,i}, m)}{\sqrt{\text{var}(f_{LOS,i})\text{var}(m)}} \quad (2.22)$$

Οι συντελεστές που προκύπτουν από την ανωτέρω διαδικασία είναι θετικοί, γραμμικά κλιμακούμενοι SPDF και υπολογίζονται με βάση την ακόλουθη εξίσωση

$$r_{SPDF,corr}(i) = \frac{r_{corr}(i)+1}{N_f + \sum_{i=1}^{N_f} r_{corr}(i)} \quad (2.23)$$

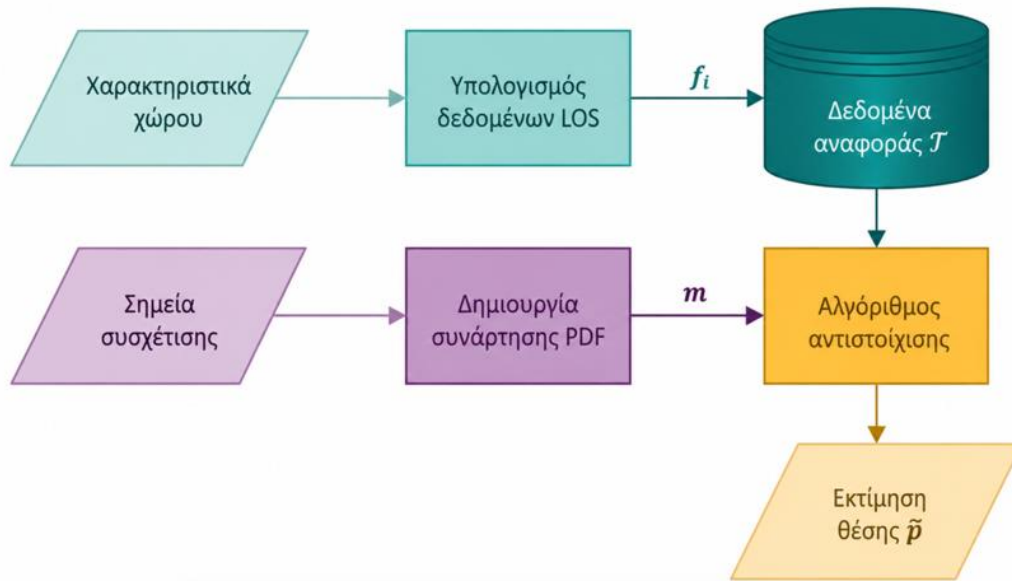
Ο αλγόριθμος SPDF έχει την δυνατότητα να καθορίσει τη συσχέτιση μεταξύ του m και κάθε σημείου από το σύνολο των δεδομένων αναφοράς T . Στην περίπτωση όμως που υπάρχουν πολλοί κόμβοι αισθητήρων, οι SPDF συνδυάζονται γραμμικά σε ίσα βάρη, με βάση την εξίσωση που παρουσιάζεται στην συνέχεια

$$r_{SPDF}(i) = \frac{1}{Q} \sum_{q=1}^Q r_{SPDF,q}(i) \quad (2.24)$$

όπου

Q: αριθμός κόμβων των αισθητήρων

Και p η εκτιμώμενη θέση με την υψηλότερη τιμή. Το διάγραμμα ροής στην συνέχεια παρουσιάζει τη λύση εντοπισμού. (Thoen, 2019)



Εικόνα 2.7 Απεικόνιση του διαγράμματος ροής του αλγόριθμου εντοπισμού PDF. (Thoen, 2019)

2.7 Μέθοδος εντοπισμού Χρόνου άφιξης (TOA)

Στο παρόν τμήμα της εργασίας εξετάζεται η μέθοδος εντοπισμού ηχητικής πηγής με βάση το χρόνο άφιξης του σήματος (TOA), η οποία χαρακτηρίζεται από αξιοπιστία και δύναται να παρέχει αποτελέσματα υψηλής ακρίβειας. Αρχικά επιλέγεται η αξιοποίηση των ακουστικών πληροφοριών κυρίως λόγω της ιδιαίτερα πολύπλοκης διαδικασίας εντοπισμού της ηχητικής πηγής.

Η συγκεκριμένη μέθοδος αποτελείται από διαδοχικά στάδια. Η ακρίβεια κάθε επιμέρους βήματος σε κάθε στάδιο καθορίζει τη μείωση των σφαλμάτων και τη βέλτιστη επίλυση του προβλήματος ανίχνευσης της ηχητικής πηγής.

Για τον εντοπισμό της ηχητικής πηγής, το πρώτο στάδιο περιλαμβάνει της αναγνώριση του εκπεμπόμενου ηχητικού σήματος από μια καθορισμένη πηγή, και στην συνέχεια, την καταγραφή του τουλάχιστον από τρεις διαφορετικές χωρικές θέσεις. Οι θέσεις αυτές πρέπει να είναι καθορισμένες με ακρίβεια, ώστε να ελαχιστοποιείται το σφάλμα του αλγόριθμου εντοπισμού. Στο επόμενο στάδιο πραγματοποιείται η ανάλυση της ανίχνευσης, κατά την οποία η καταγραφή του σήματος ακολουθείται από την αναγνώριση του. Για τη συγκεκριμένη διαδικασία εφαρμόζονται διάφορες τεχνικές φιλτραρίσματος και επεξεργασίας σήματος. Η εφαρμογή των τεχνικών αυτών εξαρτάται από την ισχύ του λαμβανόμενου σήματος, τον θόρυβο υποβάθρου και την απόκριση του ηχητικού δεκτή στο καθορισμένο φάσμα συχνοτήτων. Κατά το στάδιο της ανίχνευσης, εντοπίζεται το καταγεγραμμένο ηχητικό σήμα, και ακολούθως πραγματοποιείται προσδιορισμός του χρόνου άφιξης του (TOA), ο οποίος ορίζεται ως ο χρόνος που απαιτείται για την ανίχνευση του σήματος από το δέκτη. Στη συνέχεια, πραγματοποιείται ο εντοπισμός της ηχητικής πηγής, η οποία εξέπεμψε το ηχητικό σήμα, προκειμένου να ακολουθήσει μια λεπτομερής μελέτη. Στην περίπτωση, στην οποία η ηχητική πηγή είναι άγνωστη, δηλαδή δεν ελέγχεται από το σύστημα λήψης, η διαδικασία καθίσταται ιδιαίτερα πιο περίπλοκη.

Για τον εντοπισμό της ηχητικής πηγής, το αρχικό βήμα αφορά τον καθορισμό του χρησιμοποιούμενου ηχητικού σήματος και εν συνεχεία την εκπομπή του από την πηγή. Το σήμα διαδίδεται στο μέσο μετάδοσης, που συνήθως θεωρείται ότι είναι ο αέρας, καταλήγοντας στους δέκτες. Η ταχύτητα διάδοσης των ηχητικών κυμάτων c θεωρείται σταθερή, ενώ ο χρόνος που απαιτείται για τη λήψη του

ηχητικού σήματος από το δέκτη εξαρτάται πρωτίστως από την απόσταση μεταξύ του δέκτη της ηχητικής πηγής.

Για τον εντοπισμό της ηχητικής πηγής, καθοριστικό ρόλο επιτελεί ο χρόνος πτήσης (TOF), ο οποίος ορίζεται ως το χρονικό διάστημα που είναι απαραίτητο προκειμένου το ηχητικό σήμα να διαδοθεί από την πηγή στους δέκτες. Επιπλέον μια παράμετρο συνιστά ο χρόνος εκπομπής (TOE), ο οποίος προσδιορίζει τη χρονική στιγμή της εκπομπής του σήματος. Η συγκεκριμένη παράμετρος είναι συνήθως άγνωστη και ανεξάρτητη από το συνολικό σύστημα λήψης. Μια επιπλέον παράμετρος είναι ο χρόνος άφιξης (TOA), η οποία αποτελεί τη χρονική στιγμή ανίχνευσης του ηχητικού σήματος από το δέκτη.

Ο υπολογισμός του χρόνου πτήσης υπολογίζεται από την εξίσωση που παρουσιάζεται στην συνέχεια

$$\text{ToE} = \text{ToA} - \text{ToF} \quad (2.25)$$

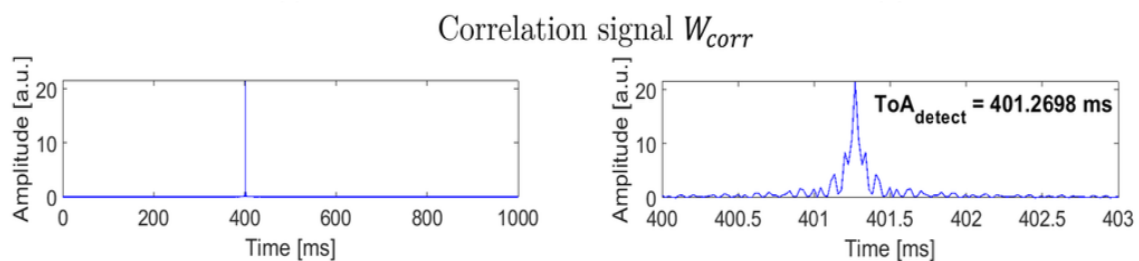
Στη συνέχεια, ο υπολογισμός της απόστασης μεταξύ της ηχητικής πηγής και του εκάστοτε δέκτη πραγματοποιείται με την ακόλουθη εξίσωσης

$$d_{\text{ER}} = \text{ToF} \cdot c \quad (2.26)$$

όπου

c: η ταχύτητα του ήχου

Στην περίπτωση σημάτων σάρωσης, η αποτελεσματική ανίχνευση του ηχητικού σήματος, πραγματοποιείται με τη μέθοδο GCC. Η μέθοδος αυτή βασίζεται στη συσχέτιση μεταξύ εκπεμπόμενου και λαμβανόμενου σήματος. Στην συνέχεια προσδιορίζεται ο χρόνος άφιξης TOA από την κορυφή της κυματομορφής του λαμβανόμενου σήματος, όπως παρουσιάζεται στην ακόλουθη εικόνα.



Εικόνα 2.8 Απεικόνιση του σήματος που προκύπτει από τη συσχέτιση των εκπεμπόμενων και των λαμβανόμενων σημάτων (αριστερά) και του αντίστοιχου ανιχνεύσιμου σήματος TOA (δεξιά). (Tortosa, 2021)

Ως χρονική διαφορά ανίχνευσης του ηχητικού σήματος ορίζεται η καθυστέρηση που προκύπτει από την χρονική διαφορά του λαμβανόμενου ηχητικού σήματος και του προσδιορισμένου χρόνου άφιξης TOA.

Για το σήμα ετεροσυσχέτισης W_{corr} , η συνάρτηση GCC υπολογίζεται για το εκπεμπόμενο σήματος W_s και το λαμβανόμενο σήμα W_r , και εξαρτάται από τη συνάρτηση της ετεροφασματικής πυκνότητας ισχύος $G_{W_s W_r}$, όπως παρουσιάζεται από την ακόλουθη εξίσωση

$$W_{\text{corr}}^{\text{GCC}}(t) = \int_{-\infty}^{+\infty} W_s(f)W_r^*(f)G_{W_sW_r}(f)e^{i2\pi ft} df = \varphi^{\text{GCC}}(f)G_{W_sW_r}(f)e^{i2\pi ft} df \quad (2.27)$$

όπου

εκθέτης *: υποδηλώνει το μιγαδικό συζυγές

$\varphi^{\text{GCC}}(f)$: είναι συνάρτηση βάρους εξαρτημένη από τη συχνότητα

Δεδομένου, ωστόσο, ότι ο αριθμός των παρατηρήσεων είναι πεπερασμένος, δύναται να γίνει μόνο η εκτίμησης της $G_{W_sW_r}$

Ακολουθως, για κάθε ζεύγος δεικτών, ο χρόνος άφιξης TOA προσδιορίζεται ως η χρονική καθυστέρηση, η οποία μεγιστοποιεί την ετεροσυσχέτιση των φιλτραρισμένων σημάτων από τα δύο μικρόφωνα, όπως παρουσιάζεται από τη εξίσωση που ακολουθεί

$$\hat{t}_{ij}^{\text{GCC}} = \arg(\max_t \{\hat{W}_{\text{corr}}^{\text{GCC1}}(t)\}) \quad (2.28)$$

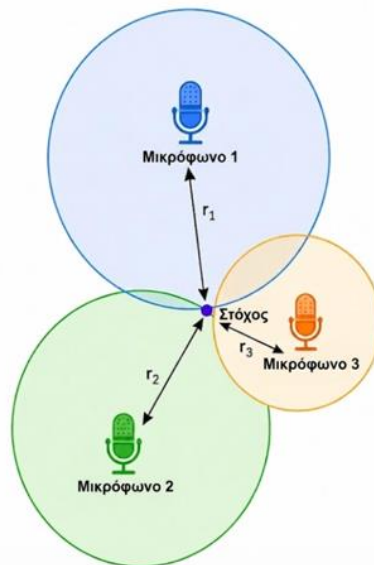
Επιπλέον, ο προσδιορισμός της θέσης της ηχητικής πηγής, πραγματοποιείται μέσω ενός γενικού μοντέλου i δεκτών. Ως προϋπόθεση για το προσδιορισμό αυτό, αποτελεί η γνώση της χωρικής θέσης των δεκτών (x_i, y_i, z_i) σε καρτεσιανό σύστημα συντεταγμένων. Εάν η προς εντοπισμό θέση της ηχητικής πηγής οριστεί με (x_E, y_E, z_E), τότε η απόσταση μεταξύ της ηχητικής πηγής και του i -στου μικροφώνου είναι d_{ER} και υπολογίζεται από την εξίσωση που παρουσιάζεται στην συνέχεια

$$d_{ERi} = \sqrt{(x_i - x_E)^2 + (y_i - x_E)^2 + (z_i - x_E)^2} \quad (2.29)$$

Ακολουθως, με βάση την παραπάνω εξίσωση δύναται ο σχηματισμός ενός μη γραμμικό συστήματος i εξισώσεων με τρεις αγνώστους, οι οποίοι είναι οι συντεταγμένες της ηχητικής πηγής (x_E, y_E, z_E). Για να πραγματοποιηθεί ο προσδιορισμός τους είναι απαραίτητη η χρήση τουλάχιστον τριών μικροφώνων. Έστω το σύστημα των εξισώσεων να επιλύεται με βάση τη διαφορά μεταξύ της απόστασης του i -στου μικρόφωνου (d_i) και το μικροφώνου αναφοράς (d_1), η οποία διαφορά δίνεται από την παρακάτω εξίσωση.

$$d_{ERi1} = d_{ERi} - d_{ER1} = \sqrt{(x_i - x_E)^2 + (y_i - x_E)^2 + (z_i - x_E)^2} - \sqrt{(x_1 - x_E)^2 + (y_1 - x_E)^2 + (z_1 - x_E)^2} \quad (2.30)$$

Για τον εντοπισμός της θέσης της ηχητικής πηγής απαιτείται η επίλυση του συστήματος των m εξισώσεων με τους n αγνώστους. (Tortosa, 2021)



Εικόνα 2.9 Απεικόνιση του σχεδιαγράμματος για τον εντοπισμό της ηχητικής πηγής μέσω της μεθόδου χρόνου άφιξης (TOA) (Aditya, 2018)

2.8 Μέθοδος εντοπισμού Διαφοράς Χρόνου άφιξης (TDOA)

Η μέθοδος ηχητικού εντοπισμού μέσω διαφοράς χρόνου άφιξης (TDOA) εκτιμά αρχικά τη χρονική διαφορά μεταξύ των ηχητικών σημάτων που λαμβάνονται από δύο ανεξάρτητα μικρόφωνα που βρίσκονται σε διαφορετικές χωρικές θέσεις. Ακολούθως, η εκτίμηση της χρονικής καθυστέρησης (TDE) συνδυάζεται με τη χωρική θέση των μικροφώνων για τον υπολογισμό του αζιμούθιου της πηγής. Για τον αλγόριθμο TDOA, βασική προϋπόθεση για την ορθή χωρική εκτίμηση της ηχητικής πηγής είναι ο ακριβής προσδιορισμός της χρονικής διαφοράς των ηχητικών σημάτων. Με βάση τις φυσικές παραμέτρους, οι τεχνικές εκτίμησης της χρονικής καθυστέρησης ταξινομούνται σε δύο κατηγορίες. Η μια κατηγορία είναι σε εκείνες που χρησιμοποιούν τη συνάρτησης γενικευμένης ετεροσυσχέτισης (GCC) και τη φάση του ετεροφασματικού πυκνώματος ισχύος. Η άλλη κατηγορία είναι σε όσες βασίζονται στον υπολογισμό της κρουστικής απόκρισης της διαδρομής διάδοσης του ηχητικού σήματος.

Έχουν σχεδιαστεί πολυάριθμα συστήματα ηχητικού εντοπισμού που αξιοποιούν συστοιχίες μικροφώνων και εφαρμόζουν τη μέθοδο εντοπισμού με διαφορά χρόνου άφιξης. Τα συστήματα αυτά κατηγοριοποιούνται σε μοντέλα εγγύς ή μακρινού πεδίου, ανάλογα με την εκάστοτε εφαρμογή. Για τη μεγιστοποίηση της ακρίβειας της εκτίμησης της θέσης της ηχητικής πηγής, η χωρική τοποθέτηση, η γεωμετρία και οι διαστάσεις της συστοιχίας των μικροφώνων διαφοροποιούνται αναλόγως. Οι διαφορές αυτές καθιστούν κάθε σύστημα ηχητικού εντοπισμού μοναδικό για τη συγκεκριμένη εφαρμογή που υλοποιεί. (Zhao, 2023)

Μέθοδος εντοπισμού διαφοράς χρόνου άφιξης με χρήση διασταυρούμενης συσχέτισης (GCC)

Στην συνέχεια της εργασίας, παρουσιάζεται η μέθοδος εντοπισμού της διαφοράς του χρόνου άφιξης με χρήση της γενικευμένης διασταυρούμενης συσχέτισης (GCC).

Αρχικά, γίνεται η υπόθεση ύπαρξης ενός ιδανικού περιβάλλοντος, στο οποίο υφίσταται ένα μοναδικό ακουστικό σήμα, το οποίο προέρχεται από μια ηχητική πηγή. Εάν η πηγή παράγει ένα σήμα το οποίο λαμβάνεται από το i -στο μικρόφωνο, αυτό περιγράφεται από την παρακάτω εξίσωση

$$x_i(t) = a_i s(t - \tau_i) + v_i(t) \quad (2.31)$$

όπου

a_i : εξασθένηση ηχητικού σήματος στο i -στο μικρόφωνο

$s(t)$: πηγή ήχου

τ_i : χρονική διαφορά μεταξύ πηγής και στο i -στο μικρόφωνο

$v_i(t)$: ο θόρυβος που λαμβάνεται στο i -στο μικρόφωνο

Η συνάρτηση συσχέτισης του ηχητικού σήματος που λαμβάνεται από το i -στο και το j -στο μικρόφωνο ορίζεται από την ακόλουθη εξίσωση

$$R_{x_i x_j} = E[x_i(t)x_j(t - \tau)] \quad (2.32)$$

Η προσδοκώμενη τιμή $E[A(t)B(t)]$ εκφράζει τη συνάρτηση διασταυρούμενης συσχέτισης μεταξύ των ηχητικών σημάτων $A(t)$ και $B(t)$. Αντικαθιστώντας την εξίσωση 2.31 στην εξίσωση 2.32, λαμβάνεται η εξίσωση που παρουσιάζεται στη συνέχεια

$$R_{x_i x_j}(\tau) = a_i a_j R_{ss}[\tau - (\tau_i - \tau_j)] + a_i R_{sv_j}(\tau - \tau_i) + a_j R_{sv_i}(\tau - \tau_j) + R_{v_i v_j}(\tau) \quad (2.33)$$

όπου

$R_{xy}(\tau)$: ορίζεται ως η συνάρτηση ετεροσυσχέτισης μεταξύ των σημάτων $x(t)$ και $y(t-\tau)$ σε χρονικό διάστημα τ , η οποία ορίζεται από τη εξίσωση

$$R_{xy}(\tau) = E[x(t)y(t - \tau)] \quad (2.34)$$

με την προϋπόθεση ιδανικού περιβάλλοντος, χωρίς θόρυβο, η εξίσωση διαμορφώνεται όπως παρουσιάζεται παρακάτω

$$R_{x_i x_j}(\tau) = a_i a_j R_{ss}(\tau - \tau_{ij}) \quad (2.35)$$

όπου

$\tau_{ij} = \tau_i - \tau_j$: ορίζεται ως η εκτιμώμενη τιμή, η οποία είναι η διαφορά χρόνου άφιξης του ηχητικού κύματος μεταξύ των δύο μικροφώνων.

Όταν $\tau = \tau_{ij}$ και $Rx_i x_j(\tau) = a_i a_j R_{ss}(0)$, η συνάρτηση συσχέτισης μεταξύ των δύο μικροφώνων λαμβάνει τη μέγιστη τιμή. Μέσω του εντοπισμού της μέγιστης τιμής, καθίσταται εφικτός ο υπολογισμός της χρονικής διαφοράς άφιξης του ηχητικού σήματος. Η τεχνική αυτή είναι η βασική μέθοδος διασταυρούμενης συσχέτισης (BCC). Ωστόσο ο αλγόριθμος αυτός για τον υπολογισμό της διαφοράς χρόνου άφιξης του ηχητικού σήματος, δεν είναι ευρέως χρησιμοποιούμενος, καθώς είναι ιδιαίτερα ευαίσθητος στον θόρυβο, με αποτέλεσμα τη μείωση της ακρίβειας προσδιορισμού της θέσης της ηχητικής πηγής. Για το λόγο αυτό, με τη εφαρμογή του μετασχηματισμού Fourier στην εξίσωση, η οποία παρουσιάστηκε προηγουμένως, προκύπτει η παρακάτω εξίσωση.

$$\Phi x_i x_j(\omega) = a_i a_j \Phi_{ss}(\omega) e^{-i\omega \tau_{ij}} \quad (2.36)$$

Ακολούθως, για τη λήψη των δεδομένων στο πεδίο του χρόνου και ταυτόχρονη καταστολή του θορύβου, καθώς και άλλων ακουστικών φαινομένων, όπως οι ανακλάσεις, εφαρμόζεται ο αντιστροφος μετασχηματισμός Fourier, ο οποίος ορίζεται από την εξίσωση που παρουσιάζεται στη συνέχεια

$$R^G x_i x_j = \int_{-\infty}^{+\infty} \Phi x_i x_j(\omega) \Psi_{ij}(\omega) e^{i\omega \tau} d\omega \quad (2.37)$$

όπου

$\Psi_{ij}(\omega)$: η γενική συνάρτηση στάθμισης, η οποία εισάγεται ανάλογα με τα χαρακτηριστικά του περιβάλλοντος εφαρμογής της μεθόδου.

Οι συνηθέστερες μέθοδοι στάθμισης παρουσιάζονται στο πίνακα που ακολουθεί. (Zhao, 2023)

Μέθοδος στάθμισης	Στάθμιση PHAT	Στάθμιση ROTH	Στάθμιση SCOT
Συναρτήσεις βάρους $\Psi_{ij}(\omega)$	$\frac{1}{ \Phi x_i x_j(\omega) }$	$\frac{1}{\Phi x_i x_i(\omega)}$	$\frac{1}{\sqrt{\Phi x_i x_j(\omega) \Phi x_i x_i(\omega)}}$
Χαρακτηριστικά στάθμισης	Ικανοποιητική ακρίβεια εκτίμησης σε περιβάλλον θορύβου	Υψηλή ακρίβεια εκτίμησης σε περιβάλλον με χαμηλά επίπεδα θορύβου	Βελτιωμένη ακρίβεια εκτίμησης σε συνθήκες χαμηλού λόγου σήματος προς θόρυβο

Πίνακας 2.1 Απεικόνιση συνηθέστερων μεθόδων στάθμισης GCC.

2.9 Μέθοδος εντοπισμού Διαφοράς Συχνότητας άφιξης (FDOA)

Η μέθοδος εντοπισμού διαφοράς συχνότητας άφιξης (FDOA), γνωστή και ως διαφορά Doppler, αποτελεί τεχνική ηχητικού εντοπισμού για τον προσδιορισμό της θέσης αλλά και της απόστασης της ηχητικής πηγής. Η ηχητική πηγή αυτή δύναται σε αρκετές περιπτώσεις, να είναι κινούμενη, όπως ένα όχημα. Η μέθοδος αυτή αναλύει τις διαφορές στη συχνότητα του εκπεμπόμενου σήματος, οι οποίες παρουσιάζουν σχετική μετατόπιση Doppler, όπως αυτή λαμβάνεται από πολλά ανεξάρτητα μικρόφωνα.

Η μέθοδος βασίζεται στη σχετική κίνηση μεταξύ της ηχητικής πηγής και των αντίστοιχων μικροφώνων, η οποία παράγει μετρήσιμες μετατοπίσεις συχνότητας. Οι μετατοπίσεις αυτές είναι ανάλογες της αντίστοιχης διαφοράς των ακτινικών ταχυτήτων κατά μήκος των γραμμών διάδοσης του ακουστικού σήματος από την ηχητική πηγή.

Στην περίπτωση όμως, που χρησιμοποιούνται δύο μικρόφωνα i και j , η μέθοδος εντοπισμού διαφοράς συχνότητας άφιξης εκφράζεται από την ακόλουθη εξίσωση

$$\Delta f_{ij} = f_i - f_j = \frac{f_0}{c}(r_i - r_j) \quad (2.38)$$

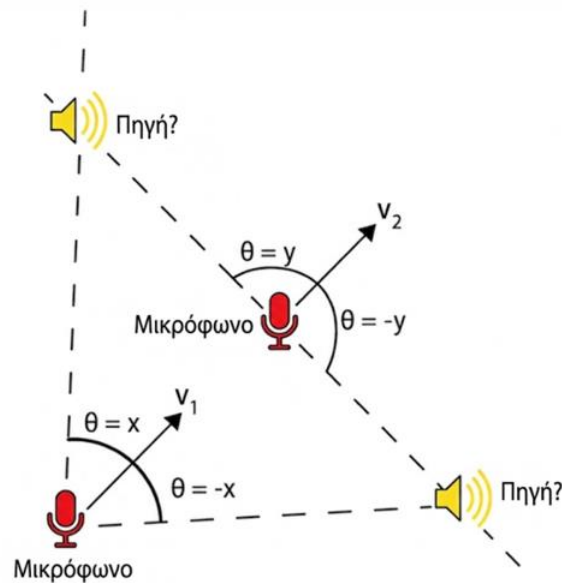
όπου

f_0 : συχνότητα ηχητικής πηγής

c : ταχύτητα του ήχου

r_i, r_j : ρυθμοί εμβέλειας κάθε μικροφώνου

Ακολούθως, με δεδομένη την αρχική συχνότητα του σήματος, προσδιορίζεται η διαφορική μετατόπιση Doppler (Δf) μεταξύ των δύο μικροφώνων. Με τον τρόπο αυτό, εξάγονται πληροφορίες αναφορικά με τη θέση και την ταχύτητα της πηγής, βασιζόμενες στη διαφορά των ακτινικών ταχυτήτων ως προς τη σχετική κίνηση μεταξύ της ηχητικής πηγής και των μικροφώνων.



Εικόνα 2.10 Απεικόνιση των δύο πιθανών θέσεων που δύναται να έχει η ηχητική πηγή με βάση των μετρήσεων από δύο μικρόφωνα. (Edge, 2019)

Η μέθοδος εντοπισμού διαφοράς συχνότητας άφιξης βασίζεται στην σύγκριση των συχνοτικών διαφορών μεταξύ των μικροφώνων ή των διαδοχικών παρατηρήσεων από ένα κινούμενο μικρόφωνο, προκειμένου να δημιουργηθούν ίσο-Doppler περιγράμματα. Πρόκειται για καμπύλες σταθερής διαφοράς συχνότητας, με τις οποίες προσδιορίζεται η αντίστοιχη θέση της ηχητικής πηγής, στην περίπτωση που χρησιμοποιούνται δεδομένα από πολλούς ακουστικούς αισθητήρες. Για παράδειγμα, η κίνηση ενός μόνο ακουστικού αισθητήρα, δύναται να προκαλεί μετατόπιση Doppler, η οποία υποδηλώνει τη γωνία της κίνησης ως προς την ηχητική πηγή. Επισημαίνεται όμως, ότι στην περίπτωση αυτή απαιτούνται περισσότερες από δύο μετρήσεις, ώστε να επιτευχθεί συμμετρία και να μη υφίστανται ασάφειες για τον ακριβή εντοπισμό της θέσης της ηχητικής πηγής.

Τα συστήματα εντοπισμού διαφοράς συχνότητας άφιξης βασίζονται κυρίως στην διαφορική μετατόπιση Doppler, όπως παρουσιάζεται με την παρακάτω εξίσωση

$$\Delta f = fr_1 - fr_2 \quad (2.39)$$

όπου

fr_1, fr_2 : είναι οι συχνότητες που λαμβάνονται από δύο χωρικά ανεξάρτητα μικρόφωνα, και οι οποίες προκύπτουν από τη διαφορά των ακτινικών ταχυτήτων κατά μήκος των αντίστοιχων γραμμών διάδοσης των ηχητικών σημάτων ως προς την ηχητική πηγή.

Επιπλέον, οι σταθερές τιμές της διαφορικής μετατόπισης Δf ορίζουν τις γεωμετρικές επιφάνειες στις οποίες η διαφορά των σχετικών ταχυτήτων παραμένει σταθερή. Σε αυτή την περίπτωση απαιτείται η εφαρμογή αριθμητικών μεθόδων για τον προσδιορισμό της θέσης της ηχητικής πηγής, με βάση τον καθορισμό της γραμμής βάσης του ακουστικού δέκτη και το διάνυσμα της σχετικής ταχύτητας.

Η μέθοδος εντοπισμού διαφοράς συχνότητας άφιξης επηρεάζεται από πολλούς παράγοντες, οι οποίοι επιδρούν κυρίως στις μετατοπίσεις Doppler, οι οποίες αποτελούν τη βάση των μετρήσεων της συγκεκριμένης μεθόδου. Επιπλέον, καθοριστικός παράγοντας είναι η κατεύθυνση καθώς και το μέγεθος της ταχύτητας της ηχητικής πηγής, τα οποία ενισχύουν αναλογικά τη μετατόπιση λόγω της

κίνησης των μικροφώνων, συμβάλλοντας στις συνιστώσες της σχετικής ταχύτητας. Επιπλέον, ένας άλλος παρόντας που επιδρά είναι οι διαδρομές διάδοσης του ακουστικού σήματος, οι οποίες επηρεάζονται από διάφορα ηχητικά φαινόμενα όπως η διάθλαση. Ωστόσο σε πολλές περιπτώσεις, τα φαινόμενα αυτά αναιρούνται μέσω της διαδικασίας της διαφοροποίησης.

Στην συνέχεια, παρουσιάζεται μια αναφορά στην παραγωγή των εξισώσεων για τη μέθοδο εντοπισμού διαφοράς συχνότητας άφιξης. Η οποία βασίζεται στο μη σχετικιστικό φαινόμενο Doppler για ακουστικό σήμα, που εκπέμπεται από κινούμενη ηχητική πηγή και λαμβάνεται από σταθερό μικρόφωνο. Θεωρώντας ότι η ηχητική πηγή εκπέμπει σήμα συχνότητα f_0 με ταχύτητα v_e κατά μήκος της γραμμής διασύνδεσης, η τιμή της ταχύτητας είναι θετική στην περίπτωση που ηχητική πηγή απομακρύνεται από το μικρόφωνο. Αντίστοιχα, εάν το μικρόφωνο κινείται με σταθερή ταχύτητα v_r , η τιμή αυτή θεωρείται θετική όταν κινείται προς την ηχητική πηγή. Με βάση αυτά, η λαμβανόμενη συχνότητα f_r περιγράφεται από την εξίσωση που ακολουθεί

$$f_r = f_0 \frac{c + v_r}{c + v_e} \quad (2.40)$$

όπου

c : ταχύτητα διάδοσης του ήχου

Όσο αφορά τη μέθοδο εντοπισμού της διαφοράς συχνότητας άφιξης, θεωρούνται δύο σταθερά μικρόφωνα στις θέσεις $\vec{r}_1$ και $\vec{r}_2$ και μια κινούμενη ηχητική πηγή στο $\vec{r}_e$ της οποίας η ταχύτητα είναι $\vec{v}_e$. Παρατηρώντας το ηχητικό σήμα της πηγής, δηλαδή τα μοναδιαία διανύσματα με κατεύθυνση από την ηχητική πηγή προς τα μικρόφωνα, τα οποία παρουσιάζονται με τις ακόλουθες εξισώσεις

$$\hat{u}_1 = \frac{(\vec{r}_1 - \vec{r}_e)}{\|\vec{r}_1 - \vec{r}_e\|} \quad (2.41)$$

$$\hat{u}_2 = \frac{(\vec{r}_2 - \vec{r}_e)}{\|\vec{r}_2 - \vec{r}_e\|} \quad (2.42)$$

και οι αντίστοιχες συχνότητες των ηχητικών σημάτων που λαμβάνονται από τα μικρόφωνα παρουσιάζονται στη συνέχεια

$$f_1 = f_0 \left(1 + \frac{\vec{v}_e \cdot \hat{u}_1}{c}\right) \quad (2.43)$$

$$f_2 = f_0 \left(1 + \frac{\vec{v}_e \cdot \hat{u}_2}{c}\right) \quad (2.44)$$

στην περίπτωση όπου θεωρηθεί ότι $v_r = 0$, η μέθοδος εντοπισμού της διαφοράς συχνότητας άφιξης εκφράζεται από την παρακάτω εξίσωση

$$\Delta f = f_1 - f_2 \approx -\frac{f_0}{c} \vec{v}_e (\widehat{u}_2 - \widehat{u}_1) \quad (2.45)$$

Για την ηχητική πηγή σε αντιστοιχία με τις θέσεις των μικροφώνων, και με βάση στους σχετικούς ρυθμούς μεταβολής της απόστασης, μπορεί να δημιουργηθεί η διαφορική μορφή της εξίσωσης. Η απόσταση μέχρι το πρώτο μικρόφωνο ορίζεται ως $d_1 = \|\vec{r}_e - \vec{r}_1\|$, ενώ ο αντίστοιχος ρυθμός μεταβολής της απόστασης παρουσιάζεται από την ακόλουθη εξίσωση

$$\dot{d}_1 = \vec{v}_e \nabla_{\vec{r}_e} d_1 = \frac{\vec{v}_e (\vec{r}_e - \vec{r}_1)}{d_1} = -\vec{v}_e \widehat{u}_1 \quad (2.46)$$

Αντιστοίχως, η απόσταση μέχρι το δεύτερο μικρόφωνο ορίζεται ως $d_2 = \|\vec{r}_e - \vec{r}_2\|$, ενώ ο αντίστοιχος ρυθμός μεταβολής της απόστασης παρουσιάζεται από την ακόλουθη εξίσωση

$$\dot{d}_2 = \vec{v}_e \nabla_{\vec{r}_e} d_2 = \frac{\vec{v}_e (\vec{r}_e - \vec{r}_2)}{d_2} = -\vec{v}_e \widehat{u}_2 \quad (2.47)$$

Συνεπώς, η διαφορά μεταξύ των ρυθμών μεταβολής της απόστασης παρουσιάζονται με την εξίσωση που ακολουθεί

$$\dot{\Delta d} = \dot{d}_1 - \dot{d}_2 = \vec{v}_e (\widehat{v}_2 - \widehat{v}_1) \quad (2.48)$$

Ως εκ τούτου, η συνάρτηση της μεθόδου εντοπισμού διαφοράς συχνότητας άφιξης προσεγγίζεται από την εξίσωση η οποία παρουσιάζεται στη συνέχεια

$$\Delta f = -\frac{f_0}{c} \dot{\Delta d} = -\frac{f_0}{c} (\Delta v_r - \Delta v_e) \quad (2.49)$$

όπου

$\Delta v_r, \Delta v_e$: οι διαφορικές συνιστώσες της ταχύτητας κατά μήκος της γραμμής διάδοσης του ηχητικού σήματος για $\Delta v_r = 0$

Η εξίσωση της μεθόδου εντοπισμού διαφοράς συχνότητας άφιξης, τόσο για την περίπτωση του δισδιάστατου όσο και του τρισδιάστατου εντοπισμού, εκφράζεται με βάση τα διανύσματα θέσης. Στην περίπτωση αυτή, η κλίση της συνάρτησης της απόστασης ως προς τη θέση της ηχητικής πηγής εκφράζεται από την εξίσωση που παρουσιάζεται στη συνέχεια

$$\Delta_{\vec{r}_e} \|\vec{r}_e - \vec{r}_1\| = \frac{(\vec{r}_e - \vec{r}_1)}{\|\vec{r}_e - \vec{r}_1\|} = -\widehat{u}_1 \quad (2.50)$$

Συνεπώς, η παραπάνω εξίσωση ορίζει μια υπερβολική δισδιάστατη ή τρισδιάστατη επιφάνεια με διαφορά συχνότητας Δf , όπου οι τιμές από τις γραμμές βάσης προσδιορίζουν τη θέση της ηχητικής πηγής.

Σε πραγματικές συνθήκες, ωστόσο, οι λαμβανόμενες μετρήσεις περιέχουν και σήματα θόρυβο. Στην περίπτωση αυτή, ο θόρυβος δύναται να εκφραστεί ως προσθετική Γκαουσιανή διαταραχή στη διαφορά συχνότητας Δf .

Αναφορικά με τη διαφορά της εκτίμησης για τη μέθοδο εντοπισμού διαφοράς συχνότητας άφιξης $\sigma_{\Delta f}^2$, καθορίζεται μέσω του λόγου σήματος προς θόρυβο (SNR) και παρουσιάζεται από την παρακάτω εξίσωση

$$\sigma_{\Delta f}^2 = \frac{1}{\rho T_i B T_f} \quad (2.51)$$

όπου

ρ : ο λόγος σήματος προς θόρυβο

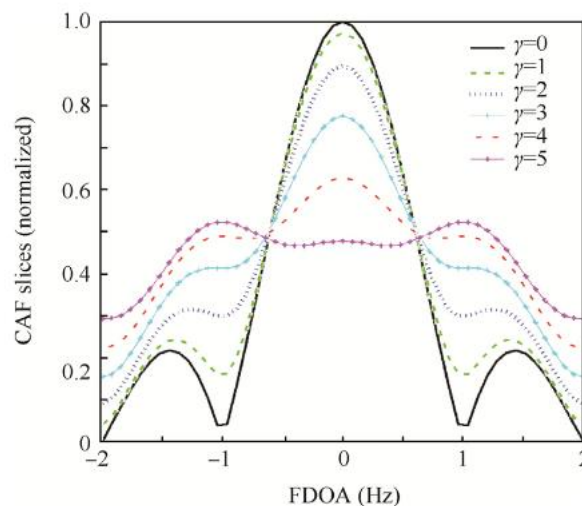
T_i : ο χρόνος ολοκλήρωσης

B : το εύρος ζώνης του σήματος

T_f : ο συντελεστής εύρους ζώνης του θορύβου

Η επίδραση του θορύβου στην εκτίμηση Doppler δύναται να υπολογιστεί με στατιστικές μεθόδους, μέσω της συνάρτησης αμοιβαίας ασάφειας σε περιβάλλον λευκού θορύβου Gauss.

Στην πραγματικότητα, η μέθοδος εντοπισμού διαφοράς συχνότητας άφιξης περιλαμβάνει την επεξεργασία των λαμβανόμενων ηχητικών σημάτων από τους συγχρονισμένους ακουστικούς αισθητήρες, με σκοπό την εξαγωγή της διαφορικής μετατόπισης Doppler. Η μετατόπιση αυτή δημιουργείται από τη σχετική κίνηση μεταξύ της ηχητικής πηγής και της γραμμής βάσης των μικροφώνων. Η κύρια τεχνική στην οποία βασίζεται η συγκεκριμένη μέθοδος ηχητικού εντοπισμού χρησιμοποιεί την ετεροσυσχέτιση των ηχητικών σημάτων για τον προσδιορισμό των διαφορών φάσης, οι οποίες στην συνέχεια παραγωγίζονται ως προς τον χρόνο, ώστε να προκύψει η διαφορά συχνότητας. Κυρίως η συνάρτηση αμοιβαίας ασάφειας (CAF) υπολογίζεται μέσω της συσχέτισης των σημάτων τόσο στο πεδίο του χρόνου όσο και στο πεδίο της συχνότητας. (<https://grokipedia.com>)



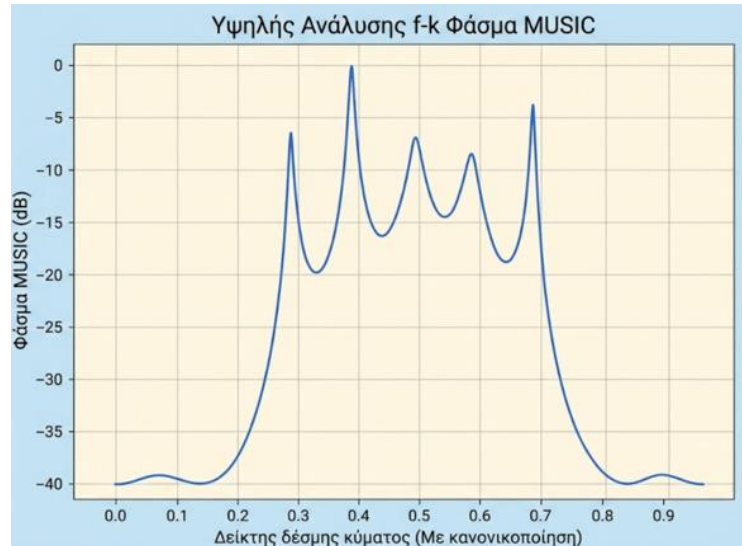
Εικόνα 2.11 Απεικόνιση της παραμόρφωσης FDOA στη συνάρτηση CAF σε σχέση με το διαφορικό Doppler. (Hu, 2018)

2.10 Μέθοδος εντοπισμού Πολλαπλής Ταξινόμησης Σημάτων (MUSIC)

Η μέθοδος εντοπισμού Πολλαπλής Ταξινόμησης Σημάτων (Multiple Signal Classification positioning method, MUSIC) αποτελεί μια μέθοδος υψηλής διακριτικής ικανότητας, η οποία βασίζεται στην ανάλυση υποδιαστημάτων και χρησιμοποιείται για την εκτίμηση παραμέτρων του σήματος όπως η ανάκτηση συχνότητας. Ως βασική αρχή, η μέθοδος πολλαπλής ταξινόμησης σημάτων (MUSIC) πραγματοποιεί αποσύνθεση των μετρούμενων δεδομένων, προκειμένου να αξιοποιήσει την ορθογωνιότητα μεταξύ του υποδιαστήματος θορύβου και των διανυσμάτων πολλαπλής διεύθυνσης. Η μέθοδος αυτή έχει δυνατότητα υψηλής διακριτικής ικανότητας, ενώ παρουσιάζει ικανοποιητική σταθερότητα σε συνθήκες χαμηλού επιπέδου θορύβου σε συνδυασμό με μεγάλο αριθμό δειγμάτων. Επιπλέον, δύναται η μέθοδος αυτή να εφαρμοστεί σε πεδία όπως η φασματική εκτίμηση και η επεξεργασία πολυδιάστατων πινάκων.

Στην συνέχεια παρουσιάζονται συνοπτικά τα κύρια χαρακτηριστικά της μεθόδου

- α) Αποτελεί μέθοδο υψηλής διακριτικής ικανότητας, η οποία βασίζεται στον διαχωρισμό του υποδιαστήματος σήματος από το υποδιάστημα θορύβου, αξιοποιώντας την ορθογωνιότητα μεταξύ του υποδιαστήματος θορύβου και των διαστημάτων διεύθυνσης.
- β) Δύναται να αξιοποιήσει ειδικές δομές πινάκων δεδομένων, όπως Hankel και Vandermonde επιτυγχάνοντας ακριβή ανάκτηση παραμέτρων, χωρίς να επηρεάζεται αρνητικά από περιβάλλον με χαμηλό θόρυβο ή με περιορισμένο αριθμό μετρήσεων.
- γ) Παρουσιάζει υψηλή υπολογιστική αποδοτικότητα, καθώς απαιτεί τον υπολογισμό μιας μόνο αποσύνθεσης μοναδιαίων τιμών SVD και υποστηρίζεται από αυστηρή ανάλυση σταθερότητας με τη χρήση διακριτών ανισοτήτων `Ingham`.



Εικόνα 2.12 Απεικόνιση ενός δείγματος της μεθόδου MUSIC. (<https://zhuanlan.zhihu.com>)

Α. Θεμελίωση και μοντέλο σήματος

Στην συνέχεια αναφέρεται ότι η μέθοδος εντοπισμού πολλαπλής ταξινόμησης σημάτων λειτουργεί με την αξιοποίηση δομημένων μοντέλων μέτρησης με προσθήκη θορύβου. Στην περίπτωση αυτή, το παρατηρούμενο σήμα αποτελεί το άθροισμα των σύνθετων διαμορφωμένων εκθετικών σημάτων.

Στην περίπτωση της μονοδιάστατης φασματικής εκτίμησης, η δειγματοληψία των δεδομένων πραγματοποιείται βάσει της εξίσωσης που παρουσιάζεται στη συνέχεια

$$y_k = \sum_{j=1}^s x_j e^{-2\pi i \omega_j k} + \varepsilon_k, \quad k = 0, \dots, M, \quad (2.52)$$

όπου

ω_j : οι άγνωστες συχνότητες

x_j : τα σύνθετα πλάτη

Στην περίπτωση, όμως, που απαιτείται επεξεργασία σε χωρικούς πίνακες, πραγματοποιείται γενίκευση σε γραμμικούς συνδυασμούς διανυσμάτων διεύθυνσης με παραμετροποίηση ως προς τη θέση, τη γωνία ή τη συχνότητα. Μια παραλλαγή της μεθόδου MUSIC αποτελεί ο μετασχηματισμός του μετρούμενου διανύσματος σε έναν δομημένο πίνακα τύπου Hankel ή Toeplitz, γεγονός που επιτρέπει μια αποσύνθεση Vandermonde ή πολλαπλού πίνακα στην περίπτωση απουσίας θορύβου. Ενδεικτικά αναφέρεται ότι ο πίνακας Hankel παρουσιάζεται στη συνέχεια, ως στιγμιότυπο μιας φασματικής εκτίμησης γραμμής.

$$H = \text{Hankel}(y) = \begin{bmatrix} y_0 & y_1 & \dots & y_{M-L} & y_2 & \dots & y_{M-L+1} & \vdots & \vdots & y_L & y_{L+1} & \dots & y_M \end{bmatrix}$$

όπου

L : η παράμετρος σχεδιασμού

M : ο μέγιστος δείκτης του δείγματος

Β. Αποσύνθεση υποδιαστήματος και συνάρτηση απεικόνισης

Οι δομημένοι πίνακες, σε περίπτωση της απουσίας θορύβου, παρουσιάζουν χαμηλή βαθμίδα, καθώς οι στήλες τους παράγονται από το σύνολο των διανυσμάτων απεικόνισης, όπως για την εκτίμηση φάσματος, από το σύνολο $\{\Phi^L(\omega_j)\}$. Το μοντέλο μέτρησης εκφράζεται μέσω της διαδικασίας της παραγοντοποίησης, όπως παρουσιάζεται με την εξίσωση στη συνέχεια

$$H = \Phi^L X (\Phi^{M-L})^T \quad (2.53)$$

όπου

Φ^L : ο πίνακας Vandermonde με στήλες $\Phi^L(\omega_j)$

X : ο διαγώνιος πίνακας με στοιχεία τα πλάτη x_j

Αξίζει να σημειωθεί ότι στην περίπτωση όπου το εύρος στηλών του πίνακα (H) ισούται με το εύρος στηλών του Φ^L , η εξίσωση, η οποία ισχύει για οποιοδήποτε ω , παρουσιάζεται στη συνέχεια

$$\omega \in \{\omega_j\} \Leftrightarrow \Phi^L(\omega) \in \text{Range}(H) \quad (2.54)$$

Η μέθοδος MUSIC βασίζεται στην ανωτέρω ιδιότητα και υπολογίζει την ορθογώνια προβολή στο υποδιάστημα θορύβου, το οποίο αποτελεί το ορθογώνιο συμπλήρωμα του εύρους στηλών του πίνακα (H). Με μια βάση U_2 του υποδιαστήματος θορύβου, η σχετική εξίσωση παρουσιάζεται στη συνέχεια

$$R(\omega) = \frac{\|U_2^T \Phi^L(\omega)\|_2}{\|\Phi^L(\omega)\|_2} \quad (2.55)$$

Με τη συνάρτηση απεικόνισης $J(\omega) = \frac{1}{R(\omega)}$, τα μηδενικά της συνάρτησης $R(\omega)$ δηλώνουν τις πραγματικές συχνότητες, δηλαδή τις πραγματικές κατευθύνσεις.

Σημειώνεται ότι στην περίπτωση των πολλαπλών πινάκων πραγματοποιείται εφαρμογή των αντίστοιχων διαδικασιών.

Γ. Ακριβής Ανάκτηση και Υπέρανάλυση

Ένα βασικό πλεονέκτημα της μεθόδου MUSIC είναι ότι επιτυγχάνει ακριβή ανάκτηση των παραμέτρων, ανεξάρτητα από τις συχνότητες, υπό την προϋπόθεση ότι ο αριθμός των μετρήσεων είναι διπλάσιος από το πλήθος των παραμέτρων και με ταυτόχρονη απουσία θορύβου. Τυπικά ισχύει αν

$$L \geq s \text{ και } M - L + 1 \geq s \text{ τότε ισχύει } \omega \in \{\omega_j\} \Leftrightarrow R(\omega) = 0 \quad (2.56)$$

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Η δυνατότητα υπερανάλυσης αποτελεί ένα μεγάλο πλεονέκτημα της μεθόδου, ακόμα και στην περίπτωση όπου η απόσταση μεταξύ των συχνοτήτων είναι εξαιρετικά μικρή, προσεγγίζοντας το όριο Rayleigh. Συνεπώς, η μέθοδος εντοπισμού πολλαπλής ταξινόμησης σημάτων δύναται να εντοπίζει ηχητικές πηγές με υψηλή ακρίβεια, σε συνθήκες με χαμηλό λόγο σήματος προς θόρυβο.

Δ. Ανάλυση σταθερότητας και χρήση των διακριτών ανισοτήτων Ingham

Παρέχεται η δυνατότητα ποσοτικής εξαγωγής των ορίων διαταραχής για τη χωρική συνάρτηση συσχέτισης του θορύβου, με χρήση ενός πρόσθετου σταθερού θορύβου. Εάν ο πίνακας Handel ισούται με $H^e = H + E$, και θεωρώντας ότι $\|E\|_2 < \sigma_s$, τότε η απόκλιση της $R(\omega)$ από το $R^e(\omega)$ δίνεται από την εξίσωση που παρουσιάζεται στη συνέχεια

$$|R^e(\omega) - R(\omega)| \leq \frac{4\sigma_1 + 2\|E\|_2}{(\sigma_s - \|E\|_2)^2} \|E\|_2 \quad (2.57)$$

Επισημαίνεται ότι στην περίπτωση όπου ο θόρυβος είναι χαμηλότερος από το όριο που σχετίζεται με την ελάχιστη μοναδιαία τιμή του δομημένου πίνακα, τότε η ανάκτηση των παραμέτρων θεωρείται δεδομένη.

Επιπλέον, κρίνεται σκόπιμη η εφαρμογή των διακριτών ανισοτήτων Ingham, με σκοπό τον καθορισμό καθορισμένων κατώτατων ορίων για τις μοναδιαίες τιμές του πίνακα Vandermonde, σε συνάρτηση με την ελάχιστη απόσταση q των συχνοτήτων.

Για κάθε τιμή που λαμβάνει η παράμετρος L , ισχύει η εξίσωση που παρουσιάζεται στη συνέχεια

$$\frac{1}{L} \sigma_s^2(\Phi^L) \geq \frac{2}{\pi} - \frac{2}{\pi L^2 q^2} - \frac{4}{L} \quad (2.58)$$

Βάσει της προαναφερθείσας εξίσωσης, παράγεται το συμπέρασμα, ότι εάν η απόσταση μεταξύ των συχνοτήτων δεν είναι υπερβολικά μικρή, το υποδιάστημα θα παραμείνει σε καλή κατάσταση. (Liao, 2015).

2.11 Μέθοδος ενεργειακού ηχητικού εντοπισμού (EBL)

Η μέθοδος ενεργειακού ηχητικού εντοπισμού (Energy Based Localization, EBL) αποτελεί μια από τις μεθόδους εντοπισμού της ηχητικής πηγής, η οποία βασίζεται στο φυσικό φαινόμενο της μείωσης της στάθμης του ήχου αναλογικά με την αύξηση της απόστασης μεταξύ της ηχητικής πηγής και του ακροατή. Μέσω του καθορισμού της σχέσης μεταξύ της απόστασης από την ηχητική πηγή και της

αντίστοιχης ακουστικής στάθμης, παρέχεται η δυνατότητα εκτίμησης της θέσης της ηχητικής πηγής, με τη χρήση των επιπέδων της ηχητικής ενέργειας σε γνωστές, διακριτές θέσεις των ακουστικών αισθητήρων. Σημειώνεται ότι η διάδοση του ήχου σε συνάρτηση με την απόσταση από την ηχητική πηγή εξαρτάται από τα χαρακτηριστικά της όπως το σχήμα και το μέγεθος. Επιπλέον, εξαρτάται από της συχνότητες του διαδιδόμενου ήχου, την κατεύθυνση του ανέμου, καθώς και από άλλα περιβαλλοντικά χαρακτηριστικά. (Hu, 2003)

Στην συνέχεια, παρουσιάζεται ένα από τα θεωρητικά μοντέλα της μεθόδου ενεργειακού ηχητικού εντοπισμού.

Ο εντοπισμός της ενδεικτικής θέσης μιας ηχητικής πηγής δύναται να πραγματοποιηθεί με την αξιοποίηση των μετρήσεων της ενέργειας που λαμβάνονται από τους διάφορους ηχητικούς αισθητήρες. Στην περίπτωση του παρόντος μοντέλου, το οποίο εξετάζει M μετρήσεις με παρουσία θορύβου μέσα σε ένα χρονικό παράθυρο $T=M/f_s$, όπου η f_s δηλώνει τη συχνότητα της ακουστικής δειγματοληψίας, υπολογίζεται ο μέσος όρος των ενεργειακών χαρακτηριστικών στο δεδομένο χρονικό παράθυρο $[t-T/2, t+T/2]$. Η λαμβανόμενη μέτρηση από τον i -στο αισθητήρα μοντελοποιείται σύμφωνα με την εξίσωση που παρουσιάζεται στη συνέχεια

$$y_i = \frac{g_i P}{\|x - s_i\|^\beta} + v_i \quad i = 1, \dots, N, \quad (2.59)$$

όπου

g_i : το κέρδος του αισθητήρα i

P : η μεταδιδόμενη ισχύς

$x = (x_x, x_y)^T$: η άγνωστη θέση

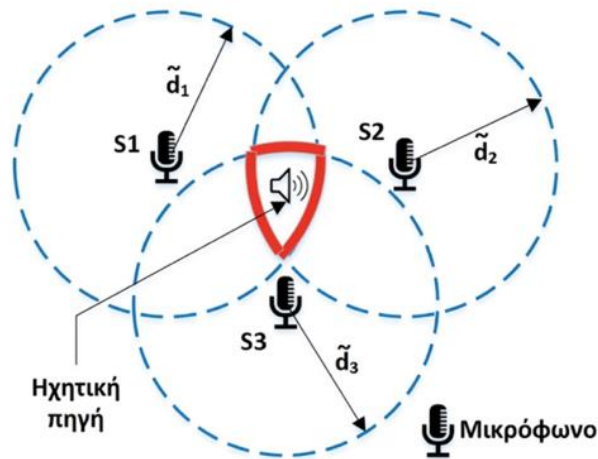
$s_i = (s_{ix}, s_{iy})$: προσδιορίζει τη γνωστή θέση του αισθητήρα i

v_i : ο θόρυβος μέτρησης, ο οποίος καθορίζεται σαν Γκαουσιανός

N : ο αριθμός των αισθητήρων

β : ο παράγοντας αποσύνθεσης, εξαρτώμενος από τις περιβαλλοντικές συνθήκες, ο οποίος σε απλοποιημένη μορφή ισοδυναμεί με 2 για το εξωτερικό περιβάλλον

Στο αναφερόμενο παράδειγμα, οι παρατηρούμενες μετρήσεις αντιστοιχούν στη λαμβανόμενη ισχύ ανά μονάδα επιφάνειας στην οποία εκπέμπεται η ενέργεια. Στην περίπτωση που λαμβάνουν χώρα αρκετές μετρήσεις, όπως περισσότερες από δύο μετρήσεις για διδιάστατο χώρο, καθορίζεται η άγνωστη θέση της ηχητικής πηγής ως το σημείο στο οποίο εφάπτονται οι υποθετικοί κύκλοι που δημιουργούνται με κέντρο τη θέση του αντίστοιχου ηχητικού αισθητήρα και ακτίνα λαμβανόμενη από το παράδειγμα. Επιπλέον, δύναται μέσω της αλλοίωσης του θορύβου των μετρήσεων, οι αισθητήρες να δημιουργήσουν μια κυκλική περιοχή αντί για να δημιουργήσουν μόνο ένα σημείο, όπως παρουσιάζεται στο παρακάτω σχήμα. (Correia, 2021)



Εικόνα 2.13 Απεικόνιση της εκτιμώμενης απόστασης των αισθητήρων, στην περίπτωση που λαμβάνεται υπόψη ο θόρυβος. (Correia, 2021)

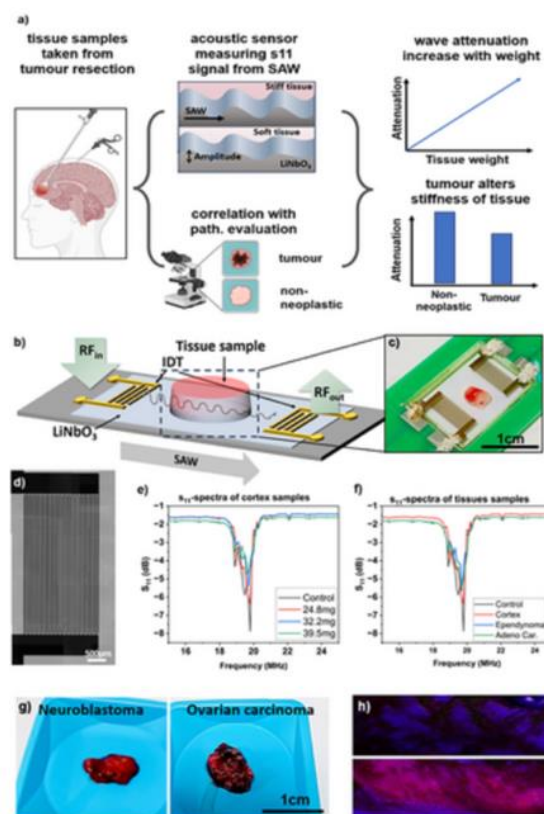
Οι μετρήσεις από τους αισθητήρες συνδυάζονται για τη διαμόρφωση μιας εκτίμησης του x , όπου ως λύση του προβλήματος εντοπισμού προκύπτει το όρισμα που ελαχιστοποιεί τη σχετική συνάρτηση, όπως παρουσιάζεται από την εξίσωση στη συνέχεια

$$\hat{x} = \arg_x \min \sum_{i=1}^N \frac{1}{\sigma_{v_i}^2} \left(y_i - \frac{g_i P}{\|x - s_i\|^2} \right)^2 \quad (2.59)$$

2.12 Σύνοψη 2^{ου} κεφαλαίου

Ο εντοπισμός ηχητικών πηγών παρουσιάζει αυξανόμενο ενδιαφέρον τα τελευταία χρόνια, και συγκεκριμένα από τα τέλη του 20^{ου} αιώνα έως και σήμερα.

Ο εντοπισμός των ηχητικών πηγών βρίσκει ευρεία εφαρμογή σε πολυάριθμους τομείς της ανθρώπινης δραστηριότητας. Αρχικά, χρησιμοποιείται για στρατιωτικούς σκοπούς, λειτουργώντας ως παθητικό στοιχείο σε συστήματα ραντάρ και σόναρ. Σε διαστημικές εφαρμογές, αξιοποιείται για την εύρεση της χωρικής θέσης πομπών εκτός της γήινης ατμόσφαιρας, καθώς και για τον εντοπισμό αστροσωματιδίων. Επιπλέον, παρουσιάζει εφαρμογή σε περιβαλλοντικά ζητήματα και σε ανθρώπινες δραστηριότητες όπως στην οδική κυκλοφορία, στη διαχείριση του θορύβου σε αεροδρόμια, καθώς και στον εντοπισμό ζώων σε δασικές εκτάσεις. Τέλος, χρησιμοποιείται σε θαλάσσιες εφαρμογές αλλά και για ιατρικούς σκοπούς, όπως για παράδειγμα στον εντοπισμό όγκων στον εγκέφαλο.



Εικόνα 2.14 Αναλυτική απεικόνιση της πειραματικής ροής για τον εντοπισμό καρκινικού όγκου στον εγκέφαλο. Στο τμήμα α παρουσιάζεται η λήψη τμήματος ιστού από τον εγκέφαλο και η ανίχνευση επιφανειακών ακουστικών σημάτων (SAW), προκειμένου να καταγραφεί η ακουστική εξασθένιση του ηχητικού σήματος. Στο τμήμα β πραγματοποιείται η παρουσίαση των ηλεκτρονικών ακουστικών αισθητήρων, καθώς και η απορρόφηση της ενέργειας του σήματος SAW από τον καρκινικό ιστό κατά μήκος της ηχητικής διαδρομής, γεγονός που έχει ως αποτέλεσμα την εμφανή εξασθένιση του σήματος. (Stamp, 2026)

Για τον λόγο αυτό, δηλαδή για τον εντοπισμό της ηχητικής πηγής σε ένα τόσο μεγάλο εύρος εφαρμογών έχουν αναπτυχθεί και μελετηθεί πολλές μέθοδοι εντοπισμού.

Στο πλαίσιο του παρόντος κεφαλαίου παρουσιάστηκαν ορισμένες από τις πιο γνωστές μεθόδους εντοπισμού ηχητικής πηγής, ωστόσο υφίσταται ένα ευρύ φάσμα εναλλακτικών εφαρμογών που δεν εξετάστηκαν στην παρούσα φάση. Οι μέθοδοι αυτές, ανάλογα με τα χαρακτηριστικά τους, δύναται να εφαρμοστούν σε διαφορετικές εργασίες. Ορισμένες μέθοδοι είναι κατάλληλες για σταθερές ηχητικές πηγές, ενώ άλλες μέθοδοι εφαρμόζονται σε κινούμενες ηχητικές πηγές. Σε αυτές τις εφαρμογές καθοριστικό ρόλο παίζει η απόσταση από την ηχητική πηγή, η παρουσία θορύβου, καθώς και το μέσο διάδοσης όπως το νερό, ο αέρας ή κάποιο άλλο υλικό.

Επισημαίνεται ότι, βάσει των διαθέσιμων βιβλιογραφικών δεδομένων, η πλειονότητα των εφαρμογών που αφορά τις μεθόδους εντοπισμού ηχητικής πηγής εστιάζει στις εφαρμογές στη θάλασσα ή σε υποθαλάσσια συστήματα. Το δεύτερο σημαντικό πεδίο εφαρμογών αποτελούν τα ρομποτικά συστήματα και τα συστήματα αυτόματης οδήγησης.

Παράλληλα το οικονομικό κόστος υλοποίησης συνιστά βασικό κριτήριο εφαρμογής κάθε μεθόδου, καθώς ορισμένες προσεγγίσεις απαιτούν τη χρήση μόνο δύο ακουστικών αισθητήρων, σε αντίθεση με άλλες που προϋποθέτουν πολύ περισσότερους. Επιπρόσθετα, η ανάλυση των δεδομένων μετά τη φάση της δειγματοληψίας παίζει σημαντικό ρόλο. Τέλος, σε κάθε περίπτωση, θεμελιώδους σημασίας είναι η ακρίβεια της εφαρμοζόμενης μεθόδου, καθώς και η ταχύτητα ολοκλήρωσης της διαδικασίας εντοπισμού της ηχητικής πηγής.

Στη πράξη, στις περισσότερες εφαρμογές πραγματοποιείται ο συνδυασμός περισσότερων από δύο μεθόδων, ώστε να επιτευχθούν ακριβέστερα και ταχύτερα αποτελέσματα.

Ένα τελευταίο σημείο που χρήζει αναφοράς αφορά το κατά πόσο η ηχητική πηγή είναι γνωστή ή άγνωστη, δηλαδή αν τα ακουστικά σήματα είναι γνωστής δομής. Μια χαρακτηριστική μέθοδος η οποία εφαρμόζεται σε άγνωστες ηχητικές πηγές είναι η μέθοδος εντοπισμού με εκτίμηση του χρόνου άφιξης (TOA).

Κεφάλαιο 3

3. Σχεδίαση και παρουσίαση τεχνικών ηχητικού εντοπισμού στη ρομποτική

3.1 Εισαγωγή

Στη συνέχεια της εργασίας, πραγματοποιείται μια γενική αλλά παράλληλα εστιασμένη αναφορά, σχετικά με τον εντοπισμό ηχητικής πηγής στο πλαίσιο ρομποτικών εφαρμογών. Βασικός στόχος της συγκεκριμένης διαδικασίας είναι ο εντοπισμός της ηχητικής πηγής ως προς δύο βασικά κριτήρια. Από την μια ως προς την θέση της ηχητικής πηγής και από την άλλη ως προς την απόσταση της ηχητικής πηγής από το σύστημα εντοπισμού της.

Επιπλέον, στον τομέα της ρομποτικής, η διαδικασία του εντοπισμού της ηχητικής πηγής κρίνεται ιδιαίτερα χρήσιμη, για διάφορες καταστάσεις, όπως ο εντοπισμός ενός ομιλητή σε μια παρουσίαση, ο εντοπισμός εγκλωβισμένων ατόμων σε μια επιχείρηση διάσωσης, όπου η δυνατότητα οπτικής επαφής είναι αδύνατη.

Η ικανότητα του εντοπισμού της ηχητικής πηγής, μπορεί να ενισχύσει τη λειτουργία και άλλων, διαφορετικών ρομποτικών συστημάτων ακρόασης. Αυτό δύναται να συμβεί διότι τα αρχικά αποτελέσματα μπορούν να χρησιμοποιηθούν σε επόμενα στάδια της επεξεργασίας, όπως για παράδειγμα, η αυτόματη αναγνώριση ομιλίας ή ο διαχωρισμός πολλαπλών ηχητικών πηγών.

Τα δεδομένα που δύναται να ληφθούν υπόψη για τον προσδιορισμό της θέσης της ηχητικής πηγής σε ένα σύστημα ηχητικού εντοπισμού είναι

- α) Η εκτιμώμενη κατεύθυνση άφιξης του ήχου
- β) Η εκτιμώμενη απόσταση της ηχητικής πηγής από το σύστημα εντοπισμού

Στην πραγματικότητα, ωστόσο, υφίσταται η πιθανότητα παρουσίας πολλαπλών ηχητικών πηγών στον χώρο, γεγονός που συνεπάγεται την ανάγκη εκτίμησης των θέσεων όλων αυτών των ηχητικών πηγών. Μια επιπλέον περίπτωση παρουσιάζεται όταν το ρομποτικό όχημα ή η ηχητική πηγή βρίσκονται σε κίνηση, και παρουσιάζεται η ανάγκη παρακολούθησης και καθορισμού της θέσης τους στον χρόνο.

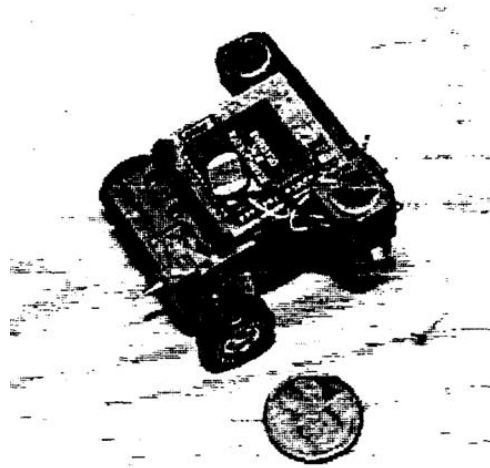
Οι μέθοδοι εντοπισμού σε ρομποτικές εφαρμογές αξιοποιούν τις τεχνικές ακουστικού εντοπισμού, οι οποίες αναλύθηκαν στο προηγούμενο κεφάλαιο, όπως η τεχνική της κατεύθυνσης άφιξης (DOA). Παράλληλα, στις ρομποτικές εφαρμογές δύναται να χρησιμοποιηθούν τεχνικές που βασίζονται στη μηχανική μάθηση όπως τα τεχνητά νευρωνικά δίκτυα. Επιπλέον, δύναται να χρησιμοποιηθούν μέθοδοι όπως η χρονική ομαδοποίηση των ηχητικών πηγών. Ωστόσο, από τις σημαντικότερες και πλέον διαδεδομένες τεχνικές εντοπισμού ηχητικών πηγών θεωρούνται τα φίλτρα Kalman και τα φίλτρα σωματιδίων (particle filtering).

Αναφορικά με τη διαδικασία του εντοπισμού της ηχητικής πηγής, παρουσιάζονται διάφοροι παράγοντες που διαδραματίζουν καθοριστικό ρόλο στην αποτελεσματικότητα της διαδικασίας του

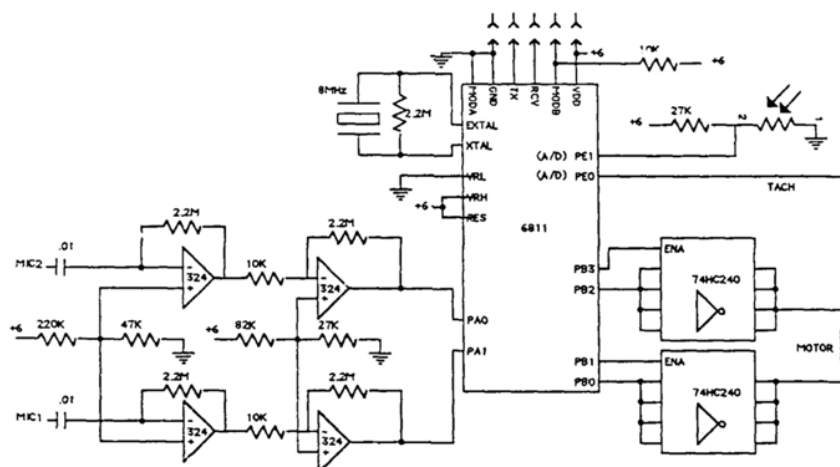
εντοπισμού. Οι οποίοι περιλαμβάνουν ποικίλα τεχνικά χαρακτηριστικά όπως ο τύπος, το είδος και ο συνολικός αριθμός των χρησιμοποιούμενων μικροφώνων. Επιπλέον, η ακρίβεια της εκτίμησης εντοπισμού εξαρτάται και από εξωγενείς παράγοντες όπως η κίνηση και ο αριθμός των ηχητικών πηγών, η ύπαρξη περιβαλλοντικού θορύβου και ακουστικών φαινομένων όπως η αντήχηση. (Rascon, 2016)

3.2 Ιστορική αναδρομή

Η ραγδαία εξέλιξη των υπολογιστικών συστημάτων και γενικότερα η τεχνολογική ανάπτυξη κατά τον 20ό αιώνα, αποτέλεσαν το έναυσμα για την ανάπτυξη της ρομποτικής. Εντούτοις, η ανάπτυξη και η πρακτική εφαρμογή μεθόδων ηχητικού εντοπισμού πηγής αναπτύχθηκαν σημαντικά αργότερα. Η πρώτη αναφορά σε ρομποτικό σύστημα με δυνατότητα εντοπισμού της θέσης ηχητικής πηγής χρονολογείται το 1989, με την παρουσίαση του ρομπότ Squirrt.



Εικόνα 3.1 Απεικόνιση του αυτόνομου ρομποτικού οχήματος Squirrt με δυνατότητα προσδιορισμού της θέσης ηχητικής πηγής. (Flynn, 1989)



Εικόνα 3.2 Απεικόνιση του σχηματικού διαγράμματος του αυτόνομου ρομποτικού οχήματος Squirrt με δυνατότητα εντοπισμού ηχητικής πηγής. (Flynn, 1989)

Στην συνέχεια, αρκετές ερευνητικές ομάδες επικεντρώθηκαν στην ανάπτυξη ρομποτικών συστημάτων με δυνατότητα ηχητικού εντοπισμού πηγής, όπως η ερευνητική ομάδα Brook με το ρομπότ Cog. Επιπλέον, ο Takanashi πραγματοποίησε έρευνα το 1993 σχετικά με ένα ανθρωπομορφικό ακουστικό σύστημα για δυναμική εφαρμογή σε ρομποτικά συστήματα, όπως για παράδειγμα το ρομπότ Robit. Παράλληλα, ο Irie από το MIT το 1995 ανέπτυξε ένα απλό ακουστικό σύστημα για ρομπότ, το οποίο προσέγγιζε τις ανθρώπινες ακουστικές ικανότητες, αν και με περιορισμένες αρχικά δυνατότητες. Συμπερασματικά, κατά το πρώτο στάδιο ανάπτυξης του ηχητικού εντοπισμού πηγής υλοποιήθηκαν αρκετές αξιόλογες ρεαλιστικές εφαρμογές για ρομπότ. Επιπλέον, από τις διάφορες ερευνητικές ομάδες αντιμετωπίστηκαν διάφορα σύνθετα σενάρια όπως η αλληλεπίδραση ανθρώπου-ρομπότ, ο σύγχρονος εντοπισμός ηχητικής πηγής, ο εντοπισμός κινούμενων ηχητικών πηγών καθώς και η ενσωμάτωσης ενός ολοκληρωμένου ακουστικού συστήματος σε ρομποτικές εφαρμογές.

Κατά τη δεκαετία του 2000, σημειώθηκε μια συστηματική προσπάθεια για κατασκευή ρομποτικών συστημάτων για ηχητικό εντοπισμό όπως το ρομπότ SIG του οποίου η βασική ιδιότητα επικεντρωνόταν στη ακρόαση. Κατά την ίδια περίοδο, ξεκίνησε η ερευνητική δραστηριότητα με σκοπό την εφαρμογή περισσότερων από δύο μικρόφωνα στα ρομποτικά συστήματα. Παράλληλα, άρχισαν να εφαρμόζονται τεχνικές εντοπισμού ηχητικής πηγής όπως η διαμόρφωση δέσμης Beamforming και η μέθοδος MUSIC. Ενδεικτικά, το ερευνητικό ινστιτούτο της Honda ανέπτυξε το 2006 ένα σύστημα παρακολούθησης πολλαπλών ηχητικών πηγών σε πραγματικό χρόνο, χρησιμοποιώντας μια ρομποτική κεφαλή με μια διάταξη οκτώ μικροφώνων. Αντιστοίχως, το 2007, μια ερευνητική ομάδα από τον Καναδά ανέπτυξε ένα ρομπότ με διάταξη οκτώ μικροφώνων, με σκοπό την αποφυγή εμποδίων μέσω της τεχνολογίας διαμόρφωσης δέσμης Beamforming.

Ο βασικός σκοπός της ανάπτυξης εφαρμογών εντοπισμού ηχητικής πηγής στη ρομποτική υπήρξε, από την αρχή, η αλληλεπίδραση ανθρώπου-ρομπότ. Το πρώτο ρομπότ που αλληλοεπίδρασε τυπικά με τον άνθρωπο, βασιζόμενο στο μηχανισμό της προσοχής, ήταν το ρομπότ Chiye. Έκτοτε, έχουν αναπτυχθεί πολυάριθμα ρομποτικά συστήματα με περισσότερες δυνατότητες αλληλεπίδρασης και επικοινωνίας με τον άνθρωπο. (Rascon, 2016)

3.3 Παρουσίαση των προβλημάτων εντοπισμού ηχητικής πηγής

Το πρόβλημα του εντοπισμού ηχητικής πηγής (SSL) ορίζεται ως ο προσδιορισμός της θέσης της ηχητικής πηγής με βάση το εκπεμπόμενο ηχητικό σήμα. Μόλις το ηχητικό σήμα ληφθεί από τη διάταξη των μικροφώνων, δημιουργείται μια αρχική εκτίμηση της θέσης της ηχητικής πηγής, εξετάζοντας τη μεθοδολογία τελικής καθοδηγούμενης επεξεργασίας. Σε πρώτο στάδιο, εξάγονται τα διάφορα χαρακτηριστικά από τα σήματα εισόδου, ενώ σε επόμενο στάδιο υλοποιείται η αντιστοιχία των χαρακτηριστικών αυτών σε σχέση με τη χωρική θέση της πηγής. (Rascon, 2016)



Εικόνα 3.3 Απεικόνιση της συνολικής διαδικασίας αναγνώρισης ηχητικής πηγής με τη μεθοδολογία τελικής καθοδηγούμενης επεξεργασίας. (Rascon, 2016)

Στη συνέχεια της παρούσας εργασίας, θα πραγματοποιηθεί η παρουσίαση των διάφορων μοντέλων.

3.3.1 Παρουσίαση του μοντέλου διάδοσης

Για τον καθορισμό του μοντέλου διάδοσης που χρησιμοποιείται σε κάθε εφαρμογή, λαμβάνονται υπόψιν διάφοροι παράγοντες, όπως η απόσταση των μικροφώνων, τα χαρακτηριστικά του χώρου, η φύση της ρομποτικής εφαρμογής, καθώς και η ύπαρξη τυχόν εμποδίων για το ηχητικό σήμα.

Στη συνέχεια πραγματοποιείται αναφορά στα βασικά μοντέλα διάδοσης

1. Μακρινό πεδίο ή πεδίο διάχυσης. Όταν η απόσταση μεταξύ του μικροφώνου και της ηχητικής πηγής είναι επαρκώς μεγάλη, ώστε το ηχητικό σήμα να προσεγγίζει τα χαρακτηριστικά επίπεδου κύματος.
2. Ελεύθερο πεδίο. Πρόκειται για τον χώρο, στον οποίο το ηχητικό σήμα διαδίδεται ανεμπόδιστα προς όλες τις κατευθύνσεις. Στην περίπτωση αυτή, η στάθμη της ακουστικής πίεσης του ήχου είναι αντιστρόφως ανάλογη της απόστασης από την πηγή του ήχου. (Φλώρος, 2013)

Όσο αφορά τα ρομποτικά συστήματα εντοπισμού ηχητικής πηγής, εκτός από τα προαναφερθέντα μοντέλα διάδοσης, αξιοποιούνται και εναλλακτικά μοντέλα όπως το σφαιρικό μοντέλο κεφαλής Woodworth-Schlosberg, το οποίο χρησιμοποιείται ευρέως σε δυαδικές διατάξεις στο πλαίσιο εφαρμογών με ρομποτικές κεφαλές. Επιπλέον, υφίσταται το μοντέλο διάδοσης εγγύς πεδίου, καθώς και ένα τροποποιημένο μοντέλο μακρινού πεδίου, το οποίο ωστόσο, δεν τυγχάνει ευρείας χρήσης, καθώς μειώνει σημαντικά την ακρίβεια εκτίμησης για τον εντοπισμό της ηχητικής πηγής. (Rascon, 2016)

3.3.2 Παρουσίαση των ακουστικών χαρακτηριστικών

Στο σημείο αυτό της εργασίας πραγματοποιείται μια συνοπτική παρουσίαση των ακουστικών χαρακτηριστικών που αξιοποιούνται σε ρομποτικά συστήματα για τον εντοπισμό ηχητικής πηγής. Επισημαίνεται ότι, αν και τα εν λόγω χαρακτηριστικά έχουν ήδη αναλυθεί στο πρώτο κεφάλαιο, στην παρούσα ενότητα εξετάζονται αποκλειστικά σε συνάρτηση με τις ρομποτικές εφαρμογές.

Φασματικές διαφορές: Στην περίπτωση που στο ρομπότ χρησιμοποιείται εξωτερικός ακροδέκτης ή επέκταση του ακουστικού αισθητήρα, παρατηρείται μια μικρή απόκλιση στη συμμετρία μεταξύ των ηχητικών σημάτων στον αισθητήρα. Η αφαίρεση αυτής της διαφοράς προκαλεί αντίστοιχα ενίσχυση ή μείωση σε ορισμένες συχνότητες, η οποία εξαρτάται από την κατεύθυνση της ηχητικής πηγής. Με βάση αυτές τις φασματικές διαφορές, καθίσταται δυνατός ο εντοπισμός της ηχητικής πηγής, κυρίως μέσω επαναλαμβανόμενων πειραματικών διαδικασιών.

Διωτικών ή φασματικών ενδείξεων: Πρόκειται για ένα επιπλέον χαρακτηριστικό, το οποίο χρησιμοποιείται για να συνοψίσει συνδυαστικά το σύνολο των χαρακτηριστικών των IPD και ILD. Το συγκεκριμένο χαρακτηριστικό χρησιμοποιείται κυρίως σε μεθόδους χαρτογράφησης που βασίζονται στη μηχανική μάθηση, καθώς και για τον περιορισμό των επιπτώσεων της αντήχησης.

Διαφορά της στάθμης της ακουστική πίεσης του μικροφώνου (IID): Η διαφορά της στάθμης της ακουστικής πίεσης αντιπροσωπεύει ουσιαστικά τη διαφορά ενέργειας μεταξύ δύο ηχητικών σημάτων που λαμβάνονται σε καθορισμένο χρονικό διάστημα. Το εν λόγω χαρακτηριστικό είναι ιδιαίτερα χρήσιμο στο πεδίο του χρόνου, καθώς δύναται να προσδιορίσει την θέση της ηχητικής πηγής.

Διωτική χρονική άφιξη (ITD) ή διαφορά χρόνου άφιξης (TDOA): Το συγκεκριμένο χαρακτηριστικό προσδιορίζει τη χρονική διαφορά μεταξύ των ηχητικών σημάτων. Μπορεί να υπολογιστεί είτε με βάση τη χρονική διαφορά μεταξύ των στιγμών μηδενικής διέλευσης των σημάτων, είτε εναλλακτικά με βάση τη χρονική έναρξη των σημάτων ή τη διαφορά φάσης μεταξύ τους. (Youssef, 2013)

3.3.3 Παρουσίαση των διαδικασιών χαρτογράφησης

Σε ορισμένες εφαρμογές απαιτείται η χαρτογράφηση της τοποθεσίας, ιδιαίτερα όταν πραγματοποιείται εκτίμηση της θέσης πολλαπλών πηγών. Μια βασική μέθοδος προσέγγισης είναι η αναζήτηση πλέγματος, στην οποία περίπτωση, η συνάρτηση της χαρτογράφησης εφαρμόζεται στο σύνολο του χώρου ενδιαφέροντος για τον εντοπισμό της ηχητικής πηγής. Στη συνέχεια, καταγράφεται η έξοδος της συνάρτησης για κάθε εκτιμώμενη θέση της ηχητικής πηγής. Η διαδικασία αυτή παράγει ως λύση ένα φάσμα, του οποίου τα τοπικά μέγιστα είναι οι λύσεις για τον εντοπισμό των ηχητικών πηγών. Χαρακτηριστικό παράδειγμα αυτής της προσέγγισης αποτελεί η μέθοδος MUSIC. (Rascon, 2016)

3.4 Παρουσίαση μεθόδων εντοπισμού ηχητικής πηγής από άκρο σε άκρο

Για τον προσδιορισμό της θέσης μιας ηχητικής πηγής απαιτούνται συνήθως δύο παράμετροι: η κατεύθυνση άφιξης (DOA) της ηχητικής πηγής και η απόσταση της από τον ακουστικό αισθητήρα. Στην πραγματικότητα, οι δύο αυτοί παράμετροι δύναται να υπολογιστούν μέσω διαφορετικών μεθοδολογιών.

Στην παρούσα ενότητα παρουσιάζεται μια αναφορά των μεθόδων εκτίμησης της κατεύθυνσης άφιξης της ηχητικής πηγής

1. **Εκτίμηση κατεύθυνσης άφιξης από μια μόνο ηχητική πηγή.** Με τη μέθοδο αυτή πραγματοποιείται μια γενική περιγραφή τεχνικών που δύναται να υπολογίσουν την εκτίμηση της κατεύθυνσης άφιξης στο επίπεδο του αζιμούθιου, για μία μόνο ηχητική πηγή.
2. **Δισδιάστατη εκτίμηση κατεύθυνσης άφιξης από μια μόνο ηχητική πηγή.** Με τη μέθοδο αυτή πραγματοποιείται μια γενική περιγραφή τεχνικών που δύναται να υπολογίσουν την εκτίμηση της κατεύθυνσης άφιξης στο επίπεδο του αζιμούθιου όσο και στο επίπεδο ανύψωσης για μια μόνο ηχητική πηγή.
3. **Εκτίμηση κατεύθυνσης άφιξης πολλαπλών ηχητικών πηγών.** Με τη μέθοδο αυτή πραγματοποιείται μια γενική περιγραφή τεχνικών που δύναται να υπολογίσουν την εκτίμηση της κατεύθυνσης άφιξης πολλαπλών ηχητικών πηγών. Η μέθοδος αυτή εφαρμόζεται κυρίως στο αζιμουθιακό επίπεδο, με δυνατότητα επέκτασης και στα δύο επίπεδα. Για την αρτιότερη μελέτη τους, η μέθοδος αυτή υποδιαιρείται στις ακόλουθες κατηγορίες, ανάλογα με την εφαρμοζόμενη τεχνική.
 - **Η μέθοδος βασιζόμενη σε διαμόρφωση δέσμης.** Η τεχνική υλοποιεί ένα χωρικό φίλτρο για τις εκτιμώμενες ηχητικές πηγές.
 - **Η μέθοδος υποχώρου.** Η τεχνική αυτή δύναται να εκμεταλλευτεί τον διαχωρισμό μεταξύ του υποχώρου σήματος και του υποχώρου θορύβου στο χώρο των μετρήσεων.
 - **Η μέθοδος χρονικής ομαδοποίησης ηχητικών πηγών.** Η τεχνική αυτή δύναται να πραγματοποιήσει εκτίμηση εντοπισμού ηχητικών πηγών σε διακριτά χρονικά τμήματα. Στη συνέχεια, συνδυάζουν τα επιμέρους αποτελέσματα μέσω ομαδοποίησης για τη συνολική αποτύπωση.
4. **Εκτίμηση απόστασης ηχητικής πηγής.** Με τη μέθοδο αυτή πραγματοποιείται εκτίμηση της απόστασης της ηχητικής πηγής από τη διάταξη των ακουστικών αισθητήρων. (Rascon, 2016)

3.4.1 Εκτίμηση κατεύθυνσης άφιξης από μια ηχητική πηγή

Μια παράμετρος για τον εντοπισμό μιας ηχητικής πηγής αποτελεί η χρονική διαφοράς άφιξης (TDOA) του σήματος μεταξύ των ακουστικών αισθητήρων. Η εκτίμηση της παραμέτρου αυτής βασίζεται στον υπολογισμό του διανύσματος διασταυρούμενης συσχέτισης (CCV) μεταξύ δύο λαμβανόμενων σημάτων. Για τον υπολογισμό του εν λόγω διανύσματος αξιοποιείται ο συντελεστής συσχέτισης Pearson, όπως παρουσιάζεται από τον παρακάτω εξίσωση

$$CCV[\tau] = \frac{\sum_t (x_1[t] - \bar{x}_1)(x_2[t - \tau] - \bar{x}_2)}{\sqrt{(\sum_t (x_1[t] - \bar{x}_1)^2) \sum_t (x_2[t - \tau] - \bar{x}_2)^2}} \quad (3.1)$$

όπου

x_1, x_2 : τα δύο λαμβανόμενα σήματα

τ : η χρονική μετατόπιση του σήματος x_2

$\bar{x}_1, \bar{x}_2$: οι μέσες τιμές των δύο σημάτων

Η μέγιστη τιμή του τ για το διάνυσμα διασταυρούμενης συσχέτισης (CCV) ορίζει τη χρονική διαφορά άφιξης τ_0 της ηχητικής πηγής.

Για το μοντέλο διάδοσης ελεύθερου πεδίου δημιουργείται μια αντιστοίχιση μεταξύ της χρονικής διαφοράς άφιξης (τ_0) και της κατεύθυνσης άφιξης (θ_0) του ηχητικού σήματος της πηγής, όπως παρουσιάζεται από την εξίσωση που ακολουθεί

$$\theta_0 = \arcsin\left(\frac{V_{\text{sound}} \tau_0}{f_{\text{sample}} d}\right) \quad (3.2)$$

όπου

V_{sound} : ταχύτητα του ήχου (340 m/s)

f_{sample} : η συχνότητα δειγματοληψίας (Hz)

d : η απόσταση μεταξύ των μικροφώνων (m)

τ_0 : η χρονική διαφορά άφιξης από την ηχητική πηγή, εκφρασμένη σε αριθμό δειγμάτων

Στην πλειονότητα των εφαρμογών, η διάταξη των μικροφώνων πραγματοποιείται με τέτοιο τρόπο, ώστε η νοητή γραμμή που τα ενώνει να είναι παράλληλη προς το αζιμουθιακό επίπεδο. Η συγκεκριμένη διάταξη πραγματοποιείται, ώστε η εκτιμώμενη κατεύθυνση άφιξης του ηχητικού σήματος της πηγής να περιορίζεται στο επίπεδο αυτό. Στην περίπτωση ωστόσο που απαιτείται η μετακίνηση των μικροφώνων σε διαφορετικό ύψος, δύναται αυτά να τοποθετηθούν έτσι ώστε να τέμνουν και το επίπεδο της ανύψωσης. Αξίζει να σημειωθεί ότι στην περίπτωση αυτή, βάσει του επιπέδου των μικροφώνων δύναται να εκτιμηθεί η κατεύθυνση άφιξης του ηχητικού σήματος μόνο ως προς μια διάσταση.

Η εκτίμηση της χρονικής διαφοράς άφιξης του ηχητικού σήματος μέσω της μεθόδου του διανύσματος διασταυρούμενης συσχέτισης παρουσιάζει ένα σημαντικό μειονέκτημα, ότι εμφανίζει υψηλή ευαισθησία στην παρουσία θορύβου, καθώς και σε άλλα ακουστικά φαινόμενα όπως η αντήχηση. Για

τον περιορισμό της ευαισθησίας αυτής, ο υπολογισμός του διανύσματος διασταυρούμενης συσχέτισης πραγματοποιείται στο πεδίο της συχνότητας (CC_F), όπως παρουσιάζεται από την παρακάτω εξίσωση

$$CC_F[f] = X_1[f]X_2[f]^* \quad (3.3)$$

όπου

X_1, X_2 : ο μετασχηματισμός Fourier των σημάτων x_1, x_2

$\{.\}^*$: τελεστής του μιγαδικού συζυγούς

Μια ιδιαίτερα σημαντική παρατήρηση είναι ότι ο αντίστροφος μετασχηματισμός Fourier της διασταυρούμενης συσχέτισης στο πεδίο της συχνότητας $F^{-1}(CC_F)$ παρουσιάζει την πληροφορία της συσχέτισης με διαφορετικό τρόπο σε σύγκριση με το αντίστοιχο διάνυσμα στο πεδίο του χρόνου. Ωστόσο, οι κορυφές της συνάρτησης $F^{-1}(CC_F)$ παρουσιάζονται στα ίδια ακριβώς σημεία υψηλής συσχέτισης με τις αντίστοιχες κορυφές του διαστήματος στο πεδίο του χρόνου, εντός του ίδιου χρονικού διαστήματος $-\tau_{\max} \leq \tau \leq \tau_{\max}$. Όπου η παράμετρος $\tau_{\max}$ είναι η μέγιστη τιμή της χρονικής υστέρησης που μπορεί να προκύψει βάση των φυσικών περιορισμών του συστήματος.

Με την μεταφορά της παραπάνω διαδικασίας στο πεδίο της συχνότητας, δύναται να εισαχθεί μια συνάρτηση στάθμισης $\psi[f]$ στην παρακάτω εξίσωση

$$GCC_F[f] = \psi[f]X_1[f]X_2[f]^* \quad (3.4)$$

Όπου η συνάρτηση στάθμισης $\psi[f]$ είναι ανάλογη με τον επιδιωκόμενο στόχο του διανύσματος συσχέτισης. Στην ειδική περίπτωση όμως, όπου η συνάρτηση $\psi[f]=1$, υλοποιείται απλοποίηση του διανύσματος $F^{-1}(GCC_F) = F^{-1}(CC_F)$. Στην περίπτωση αυτή όμως, παρουσιάζεται εκ νέου το πρόβλημα της υψηλής ευαισθησίας στο φαινόμενο της αντήρησης. Η παραπάνω εξίσωση ορίζει τη Γενικευμένη Διασταυρούμενη συσχέτιση (GCC). (Knapp, 1976)

Οι συναρτήσεις Dirac Delta εντοπίζονται στις θέσεις υψηλής συσχέτισης εντός του διανύσματος. Δεδομένου ότι ο μετασχηματισμός Fourier τοποθετεί τα πλάτη των αντίστοιχων συχνοτήτων στη μονάδα, διευκολύνεται η εμφάνιση αυτών των συναρτήσεων Dirac Delta στο πεδίο του χρόνου. Προκειμένου να επιτευχθεί μια πιο ομαλή απόκριση της γενικευμένης διασταυρούμενης συσχέτισης στο πεδίο της συχνότητας, εφαρμόζεται μια προσέγγιση εξομάλυνσης στις συναρτήσεις αυτές για τις θέσεις υψηλής συσχέτισης. Για την υλοποίηση της ομαλοποίησης αυτής, η συνάρτηση $\psi[f]$ ορίζεται ως το αντίστροφο του μέτρου του σήματος, το οποίο προκύπτει από τον πολλαπλασιασμό στο πεδίο της συχνότητας, όπως παρουσιάζεται από την παρακάτω εξίσωση

$$\psi[f] = \frac{1}{X_1[f]X_2[f]^*} \quad (3.5)$$

Ωστόσο, με την εφαρμογή της συνάρτησης στάθμισης $\psi[f]$ στη γενικευμένη διασταυρούμενη συσχέτιση, οι πληροφορίες που αφορούν τη φάση του σήματος δεν επηρεάζονται στο πεδίο της συχνότητας. Λόγω της διατήρησης αυτής της πληροφορίας, η προηγούμενη εξίσωση ορίζεται ως μετασχηματισμός φάσης (PHAT).

Στην συνέχεια, παρουσιάζεται η εξίσωση της γενικευμένης διασταυρούμενης συσχέτισης με μετασχηματισμό φάσης (GCC-PHAT)

$$\text{PHAT}_F[f] = \frac{X_1[f]X_2[f]^*}{|X_1[f]X_2[f]^*|} \quad (3.6)$$

Αξίζει να σημειωθεί, ότι η προσπάθεια σταθεροποίησης του σήματος δημιουργεί μια ομαλοποίηση σε όλα τα σήματα των εισόδων, καθώς το μέτρο των συχνοτήτων κανονικοποιείται στη μονάδα. Με τη διαδικασία αυτή, ωστόσο, παράγεται ένα έντονο διασταυρούμενο φάσμα ισχύος, το οποίο παράγει ένα βασικό μειονέκτημα, ότι όλες οι πηγές ήχου, ακόμα και οι ανακλάσεις, εμφανίζονται ως διακριτές κορυφές στο διάγραμμα συσχέτισης. Ωστόσο, με τη συγκεκριμένη μέθοδο αντισταθμίζεται η επίδραση των δευτερευουσών χρονικών διαφορών άφιξης κατά τον υπολογισμό συσχέτισης. Συνεπώς, η διαδικασία της μεθόδου της γενικευμένης διασταυρούμενης συσχέτισης με μετασχηματισμό φάσης επιδεικνύει υψηλή ανθεκτικότητα τόσο στην αντήχηση όσο και στον θόρυβο. Ωστόσο, η ομοιόμορφη εφαρμογή της συνάρτησης στάθμισης του μετασχηματισμού φάσης σε όλα τα ζεύγη των συχνοτήτων επιφέρει ευαισθησία σε πηγές θορύβου ευρείας ζώνης. Για την αντιμετώπιση του προβλήματος αυτού, δύναται να εισαχθούν πρόσθετοι συντελεστές βάρους, ανάλογα με τα επίπεδα θορύβου των αντίστοιχων ζευγών συχνοτήτων. Η επιλογή αυτή, ωστόσο, με την σειρά της δημιουργεί ένα δευτερεύον ζήτημα, καθώς η εφαρμογή μη μηδενικών βαρών σε ζεύγος συχνοτήτων, ενδέχεται να προκαλέσει θόρυβο σε άλλες συχνότητες. Για τη λύση του εν λόγω προβλήματος εφαρμόζεται μια σταθερή δυαδική μάσκα. Αυτό σημαίνει ότι στη συνάρτηση στάθμισης μετασχηματισμού φάσης εισάγονται αποκλειστικά δυαδικά βάρη. Συγκεκριμένα, αποδίδεται η τιμή ένα στην περίπτωση που ο λόγος σήματος προς θόρυβος (SNR) υπερβαίνει ένα ορισμένο όριο, ενώ σε αντίθετη περίπτωση αποδίδεται η τιμή μηδέν.

Η χρήση, ωστόσο, των δυαδικών масκών δημιουργεί επιπρόσθετο πρόβλημα των ανεπιθύμητων κυρίαρχων κορυφών στην εκτίμηση της κατεύθυνσης άφιξης. Το θέμα αυτό μπορεί να επιλυθεί με τη χρήση μιας μάσκας μετάβασης μεταξύ του σήματος του θορύβου και των παραθύρων των υπόλοιπων ηχητικών σημάτων όπως ένα σήμα ομιλίας. Μια εναλλακτική λύση υλοποιείται με χρήση δυαδικής μάσκας, η οποία επιφέρει ακύρωση αποκλειστικά στις ομάδες συχνοτήτων που βρίσκονται εκτός του εύρους συχνοτήτων της κάθε ηχητικής πηγής. Στη περίπτωση αυτή, η μάσκα δεν επηρεάζει τις συχνότητες που είναι χρήσιμες σε μια εφαρμογή.

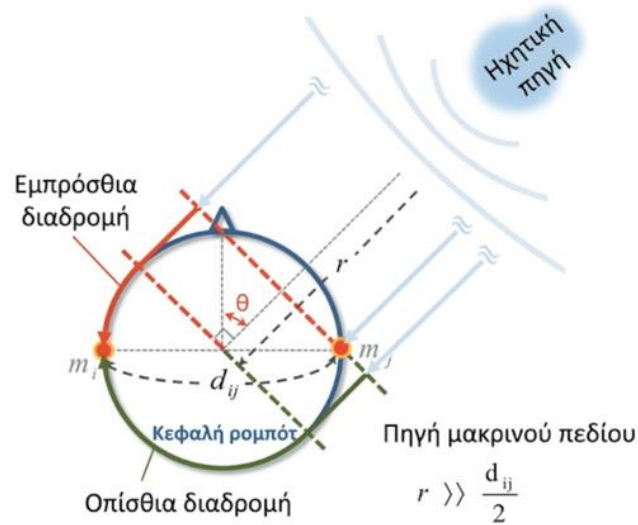
Η συνάρτηση διασταυρούμενης συσχέτισης με μετασχηματισμό φάσης είναι μια ευρέως χρησιμοποιούμενη μέθοδος για την τεχνική της χρονικής διαφοράς άφιξης. Κύριος σκοπός της είναι η εκτίμηση της κατεύθυνσης άφιξης του ηχητικού σήματος σε ένα ρομπότ, επιλογή η οποία βασίζεται κυρίως στην υπολογιστική ευκολία και στην ανθεκτικότητα της συγκεκριμένης τεχνικής.

Στην περίπτωση, ωστόσο, όπου υφίστανται ηχητικά σήματα από πολλαπλές ηχητικές πηγές, δύναται να δημιουργηθεί εμφάνιση πολλαπλών κορυφών στην συνάρτηση διασταυρούμενης συσχέτισης με μετασχηματισμό φάσης, όπου αυτές οι κορυφές διαφέρουν από την κορυφή του χρήσιμου σήματος. Ωστόσο, λόγω του συγκεκριμένου φαινομένου, η συνάρτηση διασταυρούμενης συσχέτισης με μετασχηματισμό φάσης δύναται να εφαρμοστεί για την εκτίμηση πολλαπλών κατευθύνσεων άφιξης. Επιπλέον, αξίζει να αναφερθεί ότι διάφορα μοντέλα διάδοσης του ήχου παρουσιάζουν την δυνατότητα να χρησιμοποιηθούν για τη μονοδιάστατη εκτίμηση της κατεύθυνσης άφιξης, όπως η εφαρμογή σε ένα ρομποτικό σύστημα με σφαιρική κεφαλή. Το οποίο μοντέλο παρουσιάζεται με την παρακάτω εξίσωση

$$\tau(\theta) = \frac{d}{2V_{\text{sound}}} (\theta + \sin(\theta)) \quad (3.7)$$

όπου

$d/2$: ακτίνα της κεφαλής



Εικόνα 3.4 Απεικόνιση των διαδρομών του ηχητικού σήματος σε σφαιρική ρομποτική κεφαλή λόγω του φαινομένου της περίθλασης. (Kim, 2011)

Σύμφωνα με την προηγούμενη εικόνα, το ηχητικό σήμα ακολουθεί δύο διαδρομές, οι οποίες αναπτύσσονται στο εμπρόσθιο και στο οπίσθιο τμήμα της σφαιρικής ρομποτικής κεφαλής αντίστοιχα. Η διαδρομή μέσω του εμπρόσθιου τμήματος χρησιμοποιείται για την εκτίμηση της χρονικής διαφοράς άφιξης. Ωστόσο, προκύπτει το πρόβλημα ότι η διαδρομή του ηχητικού σήματος από την οπίσθια πλευρά της κεφαλής επηρεάζει αρνητικά την εκτίμηση αυτή. (Kim, 2011)

Για την αντιμετώπιση του προβλήματος των παρεμβολών από τη δεύτερη διαδρομή του ηχητικού σήματος, εισάγεται ένας παράγοντας στάθμισης, ο οποίος αντισταθμίζει τις παρεμβολές λόγω των πολλαπλών διαδρομών. Το μοντέλο διάδοσης αυτό με το παράγοντα στάθμισης εκφράζεται με την εξίσωση που ακολουθεί

$$\tau(\theta) = \frac{d}{2V_{\text{sound}}} (\theta + \sin(\theta)) + \frac{d}{2V_{\text{sound}}} (\text{sign}(\theta)\pi - 2\theta)|\sin(\theta)| \quad (3.8)$$

όπου

$$\text{sign}(\theta) = \begin{cases} 1, & \theta \geq 0 \\ -1, & \theta < 0 \end{cases} \quad (3.9)$$

Σε ορισμένες περιπτώσεις, στο μοντέλο διάδοσης που εκφράζεται από την προηγούμενη εξίσωση, εφαρμόζεται ένας παράγοντας εξασθένησης β_m , ο οποίος λαμβάνει τυπικά την τιμή 0.1. Υπό αυτή την περίπτωση, το μοντέλο διάδοσης με την ενσωμάτωση του παράγοντα εκφράζεται με την εξίσωση που ακολουθεί.

$$\tau(\theta) = \frac{d}{2V_{\text{sound}}} (\theta + \sin(\theta)) + \frac{d}{2V_{\text{sound}}} (\text{sign}(\theta)\pi - 2\theta)|\beta_m \sin(\theta)| \quad (3.10)$$

Άλλα μοντέλα για τον εντοπισμό της ηχητικής πηγής μέσω της μονοδιάστατης εκτίμησης της κατεύθυνσης άφιξης βασίζονται στην ακουστική επιπολική γεωμετρία (AEG). Επιπλέον, χρησιμοποιούνται μέθοδοι μηχανικής μάθησης, οι οποίες δεν βασίζονται σε κάποιο μοντέλο διάδοσης του ήχου, αντίθετα εκπαιδεύονται μέσω συναρτήσεων χαρτογράφησης του χώρου με δεδομένα των χαρακτηριστικών του ηχητικού σήματος στις θέσεις της πηγής. (Murray, 2009)

3.4.2 Εκτίμηση της δισδιάστατης κατεύθυνσης άφιξης από μια ηχητική πηγή

Για την εκτίμηση της δισδιάστατης κατεύθυνσης σε ένα τρισδιάστατο περιβάλλον, απαιτείται η εκτίμηση της κατεύθυνσης άφιξης από την ηχητική πηγή τόσο στο αζιμουθιακό όσο και στο υψομετρικό επίπεδο. Στην περίπτωση που χρησιμοποιείται συστοιχία μικροφώνων με τρισδιάστατη χωρική τοποθέτηση, η γωνία ανύψωσης δύναται να υπολογιστεί μέσω των τεχνικών εκτίμησης της χρονικής διαφοράς άφιξης. Ωστόσο, για τον υπολογισμό της γωνίας χρειάζονται επιπλέον μικρόφωνα.

Στα ακουστικά συστήματα που βασίζονται στην αμφιωτική ακοή, τα μικρόφωνα τοποθετούνται στο εσωτερικό του ομοιώματος κεφαλής, τα οποία συνοδεύονται είτε από εξωτερικά πτερύγια αυτιών είτε από εσωτερικούς ακουστικούς πόρους. Η ανωτέρω ακουστική διάταξη εξασφαλίζει ότι το εισερχόμενο ηχητικό σήμα φιλτράρεται πριν από την καταγραφή του από το εκάστοτε μικρόφωνο, βάση της γωνίας πρόσπτωσης και στα δύο επίπεδα. Η μέθοδος αυτή φιλτραρίσματος δύναται να χρησιμοποιηθεί συνδυαστικά με μοντέλα διάδοσης μη ελεύθερου πεδίου, ώστε να καταστεί εφικτή η δισδιάστατη εκτίμηση της κατεύθυνσης άφιξης.

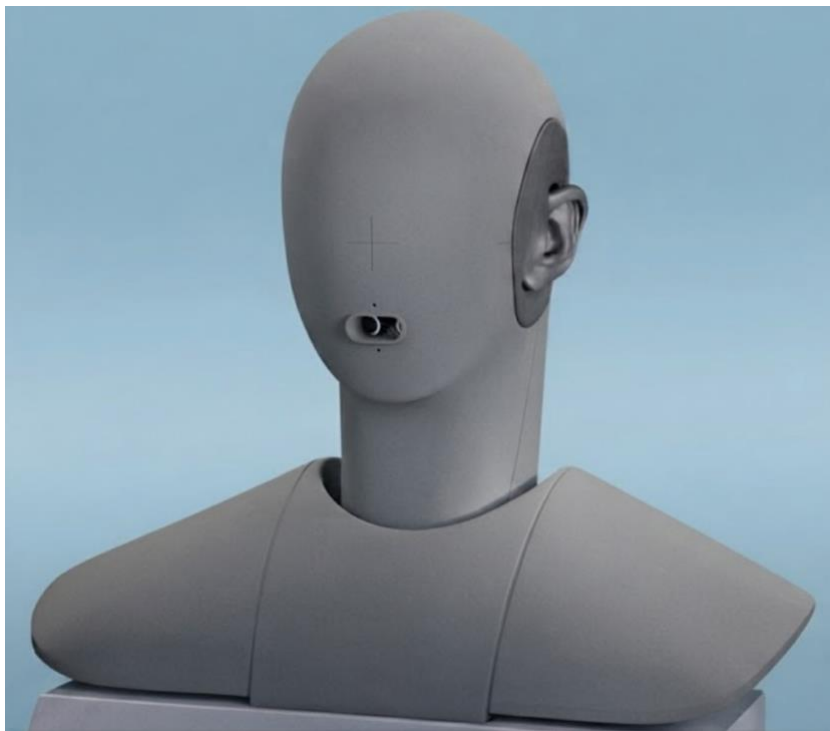
Στην πραγματικότητα, ο άνθρωπος δύναται να εκτιμήσει τη θέση μιας ηχητικής πηγής τόσο στο αζιμουθιακό επίπεδο όσο και στο υψομετρικό επίπεδο, μόνο με το βασικό σύστημα ακοής του, δηλαδή μέσω των δύο ακουστικών του αισθητήριων. Ενδεικτικά, όταν μια πηγή εκπέμπει ένα ηχητικό σήμα στη μια πλευρά του ανθρώπινου κεφαλιού, το αυτί που βρίσκεται σε μεγαλύτερη απόσταση από αυτή λαμβάνει τροποποιημένο το σήμα σε σχέση με το κοντινότερο αυτί. Αυτή η τροποποίηση καθώς και η διαφορά του ηχητικού σήματος μεταξύ των δύο αυτιών, είναι αποτέλεσμα διάφορων φυσικών φαινομένων. Τα οποία περιλαμβάνουν την ανάκλαση του ηχητικού σήματος από το ανθρώπινο κεφάλι, από το σώμα και το εξωτερικό πτερύγιο του αυτιού.

Επιπλέον, τα φυσικά φαινόμενα, τα οποία σχετίζονται με τη διαφορά της ακουστικής απόκρισης μεταξύ των δύο αυτιών, μπορούν να προσδιοριστούν μέσω πειραματικών διαδικασιών. Η μέτρηση τους πραγματοποιείται με την χρήση ενός ομοιώματος κεφαλής, στο οποίο ενσωματώνονται δύο μικρόφωνα στις θέσεις των ακουστικών πόρων, καταγράφοντας τις κρουστικές αποκρίσεις μέσα σε ένα ελεγχόμενο περιβάλλον. Με βάση τα αποτελέσματα που λαμβάνονται, δύναται να γίνει ο υπολογισμός του τρόπου με τον οποίο τα γεωμετρικά γνωρίσματα του ομοιώματος κεφαλής

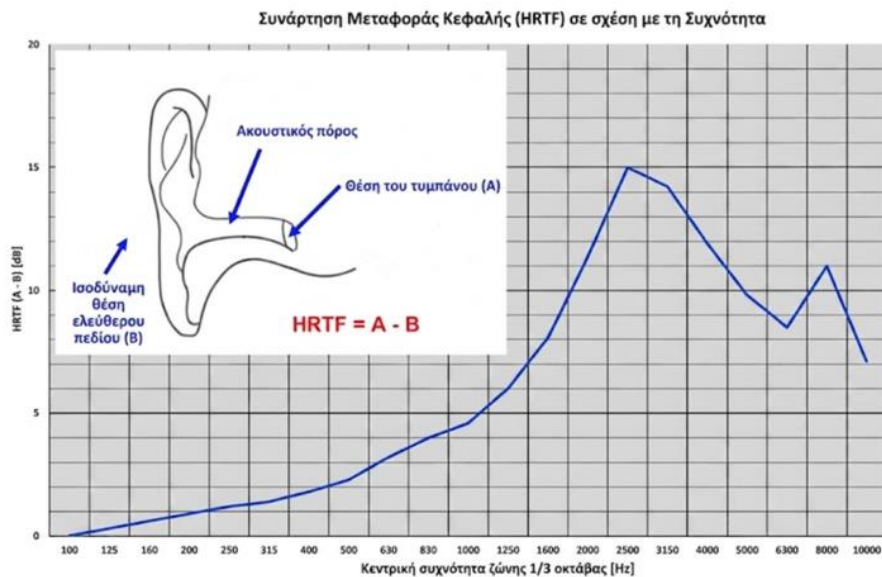
μεταβάλλουν τις ιδιότητες του λαμβανόμενου σήματος, αναφορικά με τη χωρική θέση της ακουστικής πηγής. (Hwang, 2011)

Το σύνολο των ιδιοτήτων αυτών χρησιμοποιείται για τη διαμόρφωση της συνάρτησης μεταφοράς που σχετίζεται με την κεφαλή (HRTF). Η συνάρτηση αυτή αποτελεί έναν τύπο φίλτρου, ο οποίος έχει ως σκοπό την προσομοίωση των φυσικών φαινομένων που σχετίζονται με την τροποποίηση του ηχητικού σήματος από την πηγή, με βάση μια δεδομένη κατεύθυνση άφιξης. Η συνάρτηση μεταφοράς που σχετίζεται με την κεφαλή δύναται να χρησιμοποιηθεί για τη χωρική παραγωγή ήχου, ιδιαίτερα στην περίπτωση που η ακουστική αναπαραγωγή πραγματοποιείται μέσω ακουστικών.

Για την εκτίμηση της κατεύθυνσης άφιξης, χρησιμοποιείται μια βάση δεδομένων, η οποία περιλαμβάνει φίλτρα που πραγματοποιούν την αντίστροφη συνάρτηση μεταφοράς που σχετίζεται με την κεφαλή. Επιπλέον, τα φίλτρα αυτά προσανατολίζονται στην εκάστοτε κατεύθυνση άφιξης και δύνανται να εφαρμοστούν στα εισερχόμενα σήματα ενός προκαθορισμένου πλέγματος ελέγχου, με αποτέλεσμα τον ακριβή προσδιορισμό της κατεύθυνσης. Παράλληλα, αξίζει να σημειωθεί ότι η γεωμετρία του ακουστικού πόρου και η προσομοίωση του πτερυγίου του αυτιού στο ρομποτικό σύστημα διαδραματίζουν σημαντικό ρόλο στον ακριβή υπολογισμό της συνάρτησης μεταφοράς που σχετίζεται με την κεφαλή. Αντίστοιχα, η ελικοειδής δομή του πτερυγίου συμβάλλει αποτελεσματικά στην εκτίμηση της κατεύθυνσης τόσο στο αξιμουθιακό επίπεδο όσο και στο υψομετρικό επίπεδο.



Εικόνα 3.5. Απεικόνιση προσομοιωτή κεφαλής και κορμού (HATS), που διαθέτει τεχνητό αυτί με πτερύγιο, κόγχη, ακουστικό πόρο και μικρόφωνο στο άκρο του. (<https://esselacoustics.com>)



Εικόνα 3.6 Απεικόνιση της συνάρτησης μεταφοράς που σχετίζεται με τη κεφαλή (HTRF) για τον προσομοιωτή κεφαλής και κορμού HATS. (<https://esselacoustics.com>)

Συνοψίζοντας, η διαδικασία προσέγγισης απαιτεί την αποτίμηση της επίδρασης της αντίστροφης συνάρτησης μεταφοράς, αξιοποιώντας τα δεδομένα της συνάρτησης μεταφοράς που σχετίζεται με την κεφαλή υπό ιδανικές συνθήκες, δηλαδή σε περιβάλλοντα απουσίας θορύβου και ακουστικών φαινομένων, όπως η αντήχηση. Ακολουθώντας, δύναται να πραγματοποιηθεί η καταγραφή των ηχητικών σημάτων σε πραγματικές ακουστικές συνθήκες.

Μια απλή υλοποίηση της συγκεκριμένης μεθόδου συνίσταται στην ακουστική μέτρηση του χώρου και στη συνέχεια στην αξιοποίηση των δεδομένων αυτών για τη σχεδίαση των φίλτρων, με αποτέλεσμα ότι τα φίλτρα θα είναι κατάλληλα για πραγματικές συνθήκες. Ωστόσο, η μέθοδος αυτή δεν κρίνεται αποτελεσματική σε περιπτώσεις ρομποτικών εφαρμογών. Αυτό οφείλεται στο γεγονός ότι τα ακουστικά χαρακτηριστικά του περιβάλλοντος δεν είναι εκ των προτέρων γνωστά, ενώ παράλληλα η απόκριση των χώρων εξαρτάται από την χωρική θέση των μικροφώνων. Κατά συνέπεια, καθώς το ρομποτικό όχημα κινείται, απαιτείται να γίνει εκ νέου μέτρηση της απόκρισης του χώρου και επανασχεδιασμός των φίλτρων. Εντούτοις, σε πραγματικές εφαρμογές, αποφεύγεται η μελέτη για επανασχεδιασμό των φίλτρων, λόγω της υψηλής πολυπλοκότητας της διαδικασίας και της δυσκολίας υλοποίησης της.

Για την επίλυση των διαφορών μεταξύ ιδανικών και πραγματικών συνθηκών για την εκτίμηση κατεύθυνσης άφιξης σε ρομποτικά συστήματα, είναι εφικτή η χρήση της ακουστικής επιτολικής γεωμετρίας (RAEG). Για τον εντοπισμό της ηχητικής πηγής, με βάση την συγκεκριμένη μέθοδο προϋποθέτει την εκτίμηση της αντίστροφης συνάρτησης μεταφοράς που σχετίζεται με την κεφαλή, καθώς λαμβάνει υπόψη το σχήμα και τα γεωμετρικά χαρακτηριστικά της ρομποτικής κεφαλής. Ωστόσο, η μέθοδος αυτή παρουσιάζει το μειονέκτημα ότι εξετάζει το σήμα μόνο στην κατεύθυνση του αζιμουθιακού επιπέδου.

Για μια εκτενέστερη μελέτη του χωρικού εντοπισμού δύναται να χρησιμοποιηθεί η θεωρία σκέδασης. Σύμφωνα με τη θεωρία αυτή, η πορεία ενός ακουστικού σήματος διαταράσσεται όταν συναντήσει ένα αντικείμενο, το οποίο λειτουργεί ως εμπόδιο. Αυτή η φυσική ιδιότητα χρησιμοποιείται για τον εντοπισμό των ηχητικών πηγών. Συγκεκριμένα, με βάση τη θεωρία της σκέδασης και τα δεδομένα IPD και IID, δύναται να γίνει ο υπολογισμός της κατεύθυνσης άφιξης τόσο στο αζιμουθιακό επίπεδο όσο και σε στο κατακόρυφο επίπεδο, με την προϋπόθεση ότι τα μικρόφωνα είναι τοποθετημένα σε μια σφαίρα.

Μια επιπλέον προσεγγιστική μέθοδος για την εκτίμηση της κατεύθυνσης άφιξης βασίζεται στο χάρτη αναζήτησης σε πλέγμα. Η διαδικασία της μεθόδου αυτής θα παρουσιαστεί στην συνέχεια.

1. Αρχικά υπολογίζονται τα ζεύγη τιμών (IPDf, IIDf) από κάθε πιθανό συνδυασμό για την κατεύθυνση άφιξης, τόσο σε επίπεδο αζιμούθιου όσο και σε υψομετρικό επίπεδο για μια συγκεκριμένη συχνότητα f , με τη χρήση των εξισώσεων από τη θεωρία σκέδασης.
2. Επιπλέον πραγματοποιείται η εκτίμηση των τιμών IPDf και IIDf για τα καταγεγραμμένα ηχητικά σήματα στις επιλεγμένες συχνότητες f .
3. Στο τέλος πραγματοποιείται η εκτίμηση της κατεύθυνσης άφιξης στις δύο διαστάσεις, η οποία εκτίμηση σχετίζεται με τα τρία προαναφερθέντα χαρακτηριστικά. Συγκεκριμένα, δύναται να επιλεγεί ο συνδυασμός εκείνος που προσεγγίζει περισσότερο την εκτιμώμενη τριάδα δεδομένων.

Μια εναλλακτική μέθοδος για την εκτίμηση της κατεύθυνσης άφιξης σε δύο διαστάσεις αποτελεί η μέθοδος χωρικής απόστασης (SDD). Η μέθοδος αυτή βασίζεται στον υπολογισμό της μετρικής απόστασης στο πεδίο του χρόνου-συχνότητας (TF). Συγκεκριμένα, ο υπολογισμός αυτός βασίζεται στα καταγεγραμμένα χαρακτηριστικά χρόνου και συχνότητας με τις εκτιμώμενες τιμές που θα παράγονταν στην περίπτωση που το ηχητικό σήμα προερχόταν από μια δεδομένη κατεύθυνση. Η μέθοδος αυτή βασίζεται στο σφαιρικό μετασχηματισμό Fourier (SFT). Ο μετασχηματισμός αυτός στοχεύει στον εντοπισμό αντικειμένων με περιστροφική συμμετρία σε σχέση με τις σφαιρικές συντεταγμένες, όπως οι διαταραχές που προκαλούνται στα ηχητικά κύματα λόγω της γεωμετρίας της σφαιρικής ρομποτικής κεφαλής.

Οι συντελεστές του σφαιρικού μετασχηματισμού Fourier προσδιορίζονται είτε από πειραματικές μετρήσεις είτε από προσομοιώσεις, και δύνανται να χρησιμοποιηθούν με βάση τα επιλεγμένα σημεία στο πεδίο χρόνου-συχνότητας. Τα σημεία αυτά φέρουν πληροφορίες με την άμεση διαδρομή του ηχητικού σήματος, η οποία προσδιορίζεται από μια ήδη καθορισμένη κατεύθυνση άφιξης. Με βάση αυτές τις μετρήσεις, πραγματοποιείται αναζήτηση εντός πλέγματος για τον προσδιορισμό της κατεύθυνσης, η οποία ελαχιστοποιεί τη χωρική απόσταση. Επιπλέον, το πλεονέκτημα της μεθόδου χωρικής απόστασης είναι η ανθεκτικότητα της σε σχέση με την αντήχηση του χώρου, ενώ παράλληλα δεν μεταβάλλεται ως προς το συχνοτικό εύρος της ηχητικής πηγής. Η μέθοδος αυτή χρησιμοποιείται σε πειραματικές διαδικασίες με σκοπό την ανάδειξη της επίδρασης που ασκεί μια σφαιρική περιστρεφόμενη ρομποτική κεφαλή στην εκτίμηση της κατεύθυνσης άφιξης.

Για την εκτίμηση του υψομετρικού επιπέδου, όπως προαναφέρθηκε, χρησιμοποιείται κυρίως η χαρτογράφηση μέσω αναζήτησης σε πλέγμα. Μια επιπλέον μέθοδος για την εκτίμηση του υψομετρικού επιπέδου, είναι η μέθοδος που βασίζεται στην έμμεση κωδικοποίηση της συνάρτησης μεταφοράς που σχετίζεται με την κεφαλή, με εφαρμογή μιας τεχνικής αντίστροφης παλινδρόμησης. Η μέθοδος αυτή ορίζεται ως πιθανοτική τμηματικά αφανής χαρτογράφηση (PPAM).

Με την μέθοδο πιθανοτικής τμηματικά αφανής χαρτογράφησης επιτρέπεται η αναπαράσταση των μεθόδων ILD και IPD σε όλο το εύρος των συχνοτήτων. Παράλληλα, η εκτίμηση του θορύβου δύναται να πραγματοποιηθεί μέσω του χρονικού μέσου όρου, βάσει του οποίου υπολογίζεται η μέτρηση της δραστηριότητας. Η μέτρηση αυτή εντάσσεται στη διαδικασία της χαρτογράφησης, ώστε να διασφαλιστεί η ανθεκτικότητα έναντι του ίδιου του θορύβου, του θορύβου υποβάθρου και της περιορισμένης αντήχησης.

Μια εναλλακτική προσέγγιση για τον εντοπισμού της ηχητικής πηγής βασίζεται στη χρήση νευρωνικών δικτύων και μοντέλων μηχανικής μάθησης. Αν και η εν λόγω μέθοδος προϋποθέτει την ύπαρξη ενός διευρυμένου συνόλου δεδομένων, παρουσιάζει το πλεονέκτημα της προσαρμοστικότητας σε μεταβολές των περιβαλλοντικών συνθηκών. (Deleforge, 2015)

3.4.3 Εκτίμηση της κατεύθυνσης άφιξης από πολλές ηχητικές πηγές

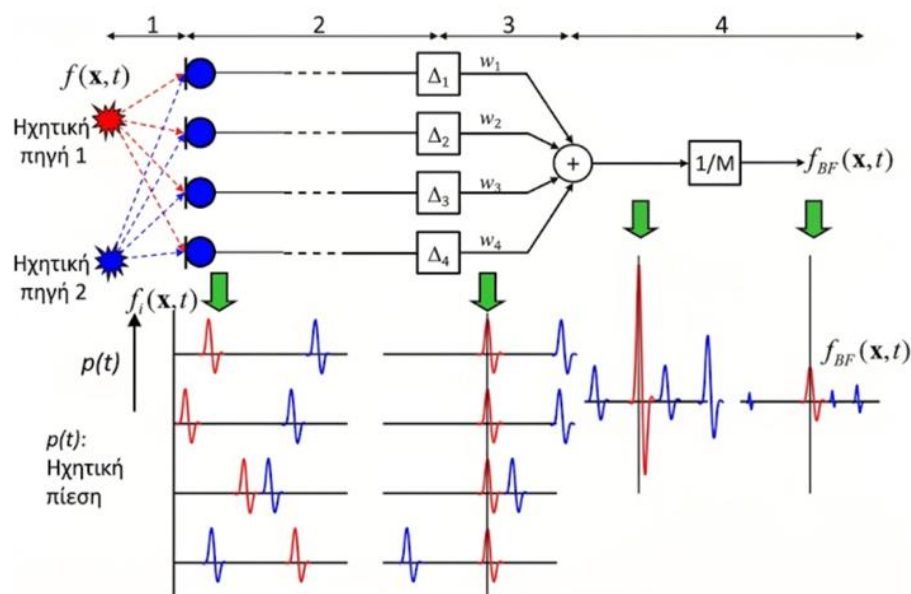
Για τις περιπτώσεις πολλαπλών ηχητικών πηγών, υφίσταται η δυνατότητα χρήσης πληθώρας διαφορετικών τεχνικών για την εκτίμηση της κατεύθυνσης άφιξης από πολλαπλές κατευθύνσεις. Οι τεχνικές αυτές ταξινομούνται σε τρεις βασικές κατηγορίες

1. Η μέθοδος μορφοποίησης δέσμης
2. Η μέθοδος υποχώρου
3. Η μέθοδος χρονικής ομαδοποίησης των ηχητικών πηγών

Για απλοποίηση της ανάλυσης των μεθόδων αυτών, η μελέτη εστιάζεται κυρίως σε μονοδιάστατες περιπτώσεις, όπου η εκτίμηση της κατεύθυνσης άφιξης περιορίζεται στο αζιμουθιακό επίπεδο. Ωστόσο επισημαίνεται ότι οι προσεγγίσεις αυτές δύνανται να επεκταθούν σε δύο διαστάσεις με χρήση γεωμετρικών μεθόδων όπως οι τρισδιάστατοι πίνακες.

3.4.3.1 Παρουσίαση της μεθόδου μορφοποίησης δέσμης

Η μορφοποίηση δέσμης αποτελεί μια μέθοδο φιλτραρίσματος, η οποία βασίζεται στην χωρική εκτίμηση των ηχητικών σημάτων που λαμβάνονται από μια συστοιχία μικροφώνων. Η τεχνική αυτή παράγει ένα σήμα εξόδου, το οποίο προσομοιώνει τη συμπεριφορά της συστοιχίας των μικροφώνων σε μια καθορισμένη κατεύθυνση άφιξης.



Εικόνα 3.7 Απεικόνιση του δομικού διαγράμματος μορφοποίησης δέσμης στο πεδίο του χρόνου. (www.qfaitech.com)

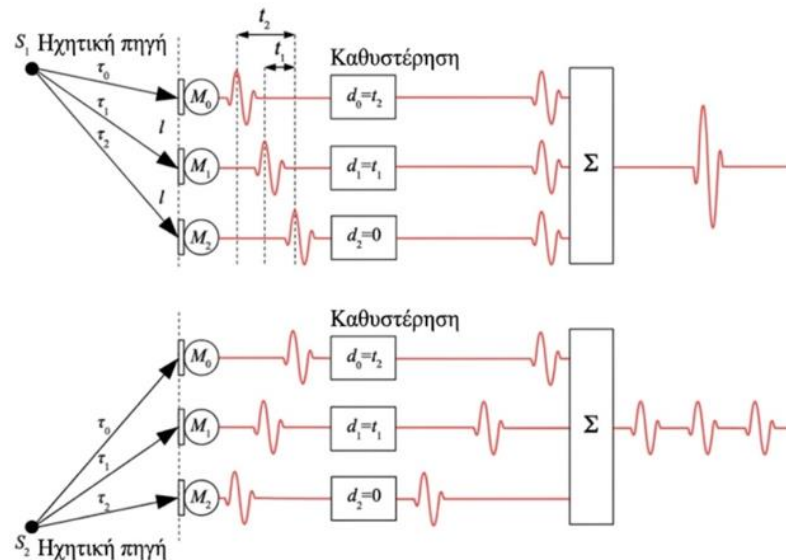
Στην συνέχεια, παρουσιάζονται τα στάδια υλοποίησης των μεθόδων εκτίμησης της κατεύθυνσης άφιξης που βασίζονται στην μορφοποίηση δέσμης.

1. Αρχικά ορίζεται ένα σύνολο L πιθανών κατευθύνσεων άφιξης, για τις οποίες γίνεται αναζήτηση των ηχητικών πηγών.
2. Επιπλέον σχεδιάζεται ένας διαμορφωτής δέσμης για κάθε μια από τις υποψήφιες κατευθύνσεις, καθορίζοντας τη λήψη του προς τη κατεύθυνση αυτή.
3. Στη συνέχεια, πραγματοποιείται μέτρηση της απόκρισης για κάθε διαμορφωτή δέσμης, ελέγχοντας την ενέργεια στην έξοδο του.
4. Έπειτα, κατασκευάζεται το αντίστοιχο φάσμα με κατευθυνόμενη απόκριση.
5. Τέλος, προσδιορίζονται οι κορυφές στο φάσμα κατευθυνόμενης απόκρισης, οι οποίες καθορίζουν τις τελικές εκτιμώμενες θέσεις των κατευθύνσεων άφιξης της πηγής.

Τα αποτελέσματα της μορφοποίησης δέσμης μέσω σάρωσης παράγουν ένα διάνυσμα $1 \times L$, το οποίο αποτυπώνει το φάσμα της κατευθυνόμενης απόκρισης.

Στην περίπτωση που χρησιμοποιείται η έξοδος του διαμορφωτή δέσμης, το φάσμα αυτό εκλαμβάνεται ως φάσμα ενέργειας που περιγράφει την λαμβανόμενη ενέργεια ανά κατεύθυνση άφιξης. Συνεπώς, ένας χάρτης αναζήτησης πλέγματος χρησιμοποιείται για την εκτίμηση πολλαπλών κατευθύνσεων άφιξης με βάση τη μορφοποίηση δέσμης. (Chen, 2003)

Η απλούστερη περίπτωση της συγκεκριμένης τεχνικής είναι η μέθοδος μορφοποίησης καθυστέρησης και αθροίσματος (DAS), όπως παρουσιάζεται από το σχήμα που ακολουθεί



Εικόνα 3.8 Απεικόνιση δομικού διαγράμματος της μορφοποίησης δέσμης καθυστέρησης και αθροίσματος (Krol, 2015).

Στην περίπτωση ενός μοντέλου διάδοσης ελεύθερου πεδίου, η κατεύθυνση άφιξης μιας ηχητικής πηγής σχετίζεται με τη χρονική διαφορά άφιξης του σήματος μεταξύ των μικροφώνων. Στο πλαίσιο αυτό, η μέθοδος μορφοποίησης δέσμης καθυστέρησης και αθροίσματος, θα υλοποιήσει μια τεχνητή χρονική μετατόπιση των λαμβανόμενων ηχητικών σημάτων. Η διαδικασία αυτή αποσκοπεί να αντισταθμίσει τη σχετική καθυστέρηση και στη συνέχεια να προσθέσει τα σήματα για την παραγωγή του σήματος εξόδου. Συνεπώς, τα στοιχεία των καταγεγραμμένων λαμβανόμενων σημάτων που προέρχονται από την καθορισμένη κατεύθυνση ευθυγραμμίζονται χρονικά. Ως εκ τούτου, το επιθυμητό ηχητικό σήμα ενισχύεται στην έξοδο της διάταξης, σε σύγκριση με τα σήματα από ηχητικές πηγές που βρίσκονται σε διαφορετικές κατευθύνσεις, τα οποία εξασθενούν. Το βασικό ρόλο στη διαδικασία αυτή διαδραματίζει ο αριθμός των χρησιμοποιούμενων μικροφώνων.

Ωστόσο, στην περίπτωση όπου τα μικρόφωνα είναι πανκατευθυντικά, η έξοδος του διαμορφωτή δέσμης με μορφοποίηση καθυστέρησης και αθροίσματος παρουσιάζεται από την παρακάτω εξίσωση

$$\hat{s}_\theta[t] = \sum_{n=1}^N x_n [t - \tau_n(\theta)] \quad (3.11)$$

όπου

s_θ : η έξοδος του διαμορφωτή δέσμης, η οποία αποτελεί την εκτίμηση του πραγματικού σήματος s_θ που λαμβάνεται από τη συστοιχία μικροφώνων στην κατεύθυνση θ .

x_n : το ηχητικό σήμα που λαμβάνεται από το n -οστό μικρόφωνο

N : ο αριθμός των μικροφώνων

$\tau_n(\theta)$: χρονική διαφορά άφιξης από την ηχητική πηγή στο n -οστό μικρόφωνο για την κατεύθυνση θ .

Η χρονική διαφορά άφιξης υπολογίζεται με βάση ένα μικρόφωνο αναφοράς. Υπό την προϋπόθεση ότι καθοριστούν η γεωμετρία της διάταξης και το μοντέλο διάδοσης, η μέθοδος αυτή δύναται να χρησιμοποιηθεί σε οποιαδήποτε χωρική διάσταση. Επιπλέον, είναι δυνατή η εφαρμογή της και σε μοντέλα διάδοσης μη ελεύθερου πεδίου, όπως το μοντέλο της σφαιρικής κεφαλής.

Η ενέργεια στην έξοδο του διαμορφωτή δέσμης με κατεύθυνση προς θ (E_θ) υπολογίζεται με βάση την εξίσωση που παρουσιάζεται παρακάτω

$$E_\theta = \sum_{t=1}^T \hat{s}_\theta[t]^2 \quad (3.12)$$

Επιπλέον, η χρονική μετατόπιση δύναται να συμβεί και στο πεδίο της συχνότητας, όπως παρουσιάζεται με την παρακάτω εξίσωση

$$X_{n_{\tau n(\theta)}}[f] = X_n[f]e^{-2\pi f \tau_n(\theta)} \quad (3.13)$$

όπου

X_n : μετασχηματισμός Fourier του σήματος x_n

f : η συχνότητα

$X_{n_{\tau n(\theta)}}$: μετασχηματισμός Fourier του χρονικά μετατοπισμένου σήματος $x_n(t - \tau_n(\theta))$

Παράλληλα, για την καταγραφή των χρονικών μετατοπίσεων που σχετίζονται με την κατεύθυνση θ των λαμβανόμενων ηχητικών σημάτων, εισάγεται ένας πίνακας βαρών W_θ με διαστάσεις $N \times F$, η δομή του οποίου παρουσιάζεται στην συνέχεια

$$W_{\theta} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-2\pi f_1 \tau_2(\theta)} & e^{-2\pi f_2 \tau_2(\theta)} & \dots & e^{-2\pi f_N \tau_2(\theta)} \\ e^{-2\pi f_1 \tau_3(\theta)} & e^{-2\pi f_2 \tau_3(\theta)} & \dots & e^{-2\pi f_N \tau_3(\theta)} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi f_1 \tau_N(\theta)} & e^{-2\pi f_2 \tau_N(\theta)} & \dots & e^{-2\pi f_N \tau_N(\theta)} \end{bmatrix} \quad (3.14)$$

όπου

f: η συχνότητα

N: ο αριθμός των μικροφώνων

F: μέγεθος του παραθύρου ανάλυσης στο πεδίο της συχνότητας

W_{θ} : πίνακας ευρείας ζώνης, του οποίου κάθε στήλη αποτελεί ένα διάνυσμα διεύθυνσης στενής ζώνης

Ο μετασχηματισμός Fourier για το σήμα εξόδου του διαμορφωτή δέσμης (S), παρουσιάζεται από την παρακάτω εξίσωση

$$\widehat{S}_{\theta}[f] = W_{\theta}[f]^H X[f] \quad (3.15)$$

όπου

τελεστής $\{\cdot\}^H$: ο τελεστής του συζυγούς ανάστροφου

X : πίνακας με διαστάσεις $N \times F$, με τα σήματα x_n να αποτελούν τις σειρές του

$W_{\theta}[f]$: διάστημα των στηλών με διαστάσεις $N \times 1$ με περιεχόμενο τα βάρη W_{θ} της μεθόδου μορφοποίησης που αφορούν την συχνότητα f

$X[f]$: διάστημα των στηλών με διαστάσεις $N \times 1$, το οποίο περιέχει συχνοτικές πληροφορίες για τα ηχητικά σήματα που έχουν καταγραφεί

$S_{\theta}[f]$: έξοδος του διαμορφωτή δέσμης με κατεύθυνση προσανατολισμού θ για τη συχνότητα f

S_{θ} : μετασχηματισμός Fourier του σήματος s_{θ} , για κάθε $s_{\theta} = F^{-1}(S_{\theta})$

Στο πεδίο της συχνότητας, ωστόσο, παρέχεται η δυνατότητα επίτευξης σημαντικά βελτιωμένων αποτελεσμάτων για τον διαμορφωτή δέσμης με βάση επιπλέον χαρακτηριστικών που αφορούν τη μήτρα W_{θ} . Αξίζει να σημειωθεί ότι, για τη βελτίωση των αποτελεσμάτων για το διαμορφωτή δέσμης, απαιτείται ένα σύνολο βαρών ($A_{\theta MVD R}$) με ελάχιστη ενέργεια, ενώ παράλληλα οφείλει να διατηρείται η ίδια κατεύθυνση, όπως και στην περίπτωση του W_{θ} . Η τεχνική αυτή παρουσιάζεται από την παρακάτω εξίσωση.

$$A_{\theta MVD R}[f] = \frac{R[f]^{-1} W_{\theta}[f]}{W_{\theta}[f]^H R[f]^{-1} W_{\theta}[f]} \quad (3.16)$$

όπου

$R[f]$: πίνακας συνδιακύμανσης $N \times N$ του $X[f]$

Με βάση τη παραπάνω εξίσωση, ορίζεται ο διαμορφωτής δέσμης Caron ή διαμορφωτής δέσμης ελάχιστης διακύμανσης χωρίς παραμόρφωση (MVDR). Για την εφαρμογή του, είναι απαραίτητος ο υπολογισμός του $R[f]$. Η εκτίμηση του πίνακα αυτού πραγματοποιείται μέσω του πίνακα δείγματος συνδιακύμανσης, ο οποίος δύναται να υπολογιστεί ως ο μέσος όρος των $X[f, t - T : t]$ και $X[f, t - T : t]^H$ σε χρονικά παράθυρα με διάρκεια T , και με το t να συμβολίζει το τρέχον χρονικό παράθυρο. (Valin, 2004)

Ένα βασικό μειονέκτημα του διαμορφωτή δέσμης ελάχιστης διακύμανσης χωρίς παραμόρφωση είναι ότι η σύγκλιση του εκτιμώμενου πίνακα στην πραγματική τιμή, δηλαδή $\hat{R}[f] = R[f]$. Η διαδικασία αυτή, ωστόσο, απαιτεί τη παρέλευση σημαντικού αριθμού χρονικών παραθύρων.

Στην συνέχεια, αναφέρεται ότι η ενέργεια E_θ εξαρτάται κυρίως από την εκτιμώμενη $\hat{S}_\theta$, το οποίο έχει ως συνέπεια ο εντοπισμός της ηχητικής πηγής να βασίζεται κυρίως στον βαθμό ενίσχυσης του ακουστικού σήματος στην κατεύθυνση άφιξης. Κατά βάση, εξετάζεται ο λόγος σήματος προς παρεμβολή (SIR) στην έξοδο του διαμορφωτή δέσμης. Όταν ο διαμορφωτής προσανατολίζεται προς την ηχητική πηγή, ο λόγος σήματος προς παρεμβολή επιδρά στο ύψος της αντίστοιχης κορυφής του φάσματος απόκρισης. Αντίθετα, στην περίπτωση όπου ο λόγος του σήματος προς παρεμβολή είναι ιδιαίτερα χαμηλός, η αντίστοιχη κορυφή δεν μπορεί να εντοπιστεί στο φάσμα. Επιπλέον, καθώς η στάθμη της ακουστικής πίεσης είναι ανάλογη του αριθμού των χρησιμοποιούμενων μικροφώνων, μια ικανοποιητική λύση στο πρόβλημα του χαμηλού λόγου σήματος προς θόρυβο αποτελεί η χρήση μιας διάταξης πολλαπλών μικροφώνων.

Ένα βασικό μειονέκτημα του διαμορφωτή δέσμης καθυστέρησης και αθροίσματος για την εκτίμηση πολλαπλών κατευθύνσεων άφιξης, είναι όταν οι κορυφές στο φάσμα κατευθυνόμενης απόκρισης παρουσιάζουν μεγάλη διασπορά, με αποτέλεσμα η διακριτική ικανότητα για την αναζήτηση στο πλέγμα να είναι χαμηλή. Μια μέθοδος για την άρση του συγκεκριμένου περιορισμού προκύπτει από τον επαναπροσδιορισμό της εξόδου του διαμορφωτή δέσμης με κύριο κριτήριο τα διανύσματα εγκάρσιας συσχέτισης, η οποία μέθοδος περιγράφεται από την εξίσωση που ακολουθεί.

$$E_\theta = \sum_{n=1}^N \sum_{t=1}^T x_n[t - \tau_n(\theta)]^2 + 2 \sum_{n_1=1}^N \sum_{n_2=1}^{n_1-1} CC V_{x_{n_1}, x_{n_2}} [\tau_{n_1}(\theta) - \tau_{n_2}(\theta)] \quad (3.17)$$

Για να επιτευχθεί οξύτερη μορφή στις κορυφές στο φάσμα κατευθυνόμενης απόκρισης δύναται να γίνει χρήση του διανύσματος μετασχηματισμού φάσης. Εντούτοις, με το διάνυσμα αυτό πραγματοποιείται κανονικοποίηση μεγέθους, με αποτέλεσμα κάθε ζώνη συχνότητας να συνεισφέρει ισοσταθμισμένα για τον υπολογισμό της συσχέτισης. Η διαδικασία αυτή, ωστόσο, έχει ως αποτέλεσμα την ευαισθησία του συστήματος στο θόρυβο. Μια μέθοδος για την αντιμετώπιση του προβλήματος συνίσταται στην ετεροβαρή συνεισφορά των συχνοτήτων, με βασική επιδίωξη τη μείωση της συμβολής όσων παρουσιάζουν χαμηλό λόγο σήματος προς θόρυβο στενής ζώνης. Για την πραγματοποίηση αυτής της προσέγγισης, απαιτείται σε κάθε χρονικό παράθυρο ο υπολογισμός της μέσης φασματικής πυκνότητας ισχύος των μικροφώνων, η οποία αναπαρίσταται από το διάνυσμα μιγαδικών τιμών $1 \times F$ της παραμέτρου X_{mean} . Συνεπώς, στο πεδίο της συχνότητας, το σήμα του θορύβου θα εκτιμηθεί βάση του χρονικού μέσου όρου X_{mean} . Τέλος, για τον υπολογισμό της συσχέτισης με χρήση της συνάρτησης στάθμισης $\psi[f]$, χρησιμοποιείται η εξίσωση που παρουσιάζεται στη συνέχεια

$$\Psi[f] = \begin{cases} \Psi_{\text{PHAT}}[f], & X_{\text{mean}}[f] \leq Y[f] \\ \Psi_{\text{PHAT}}[f] \left(\frac{X_{\text{mean}}[f]}{Y[f]} \right)^\gamma, & X_{\text{mean}}[f] > Y[f] \end{cases} \quad (3.18)$$

όπου

$$\Psi_{\text{PHAT}}[f] = \frac{1}{|X_1[f]X_2[f]^*|} : \text{μετασχηματισμός φάσης}$$

Η κλίμακα $0 < \gamma < 2$ καθορίζει το εύρος τιμών για τη στάθμιση των συχνοτήτων. Οι συχνότητες οι οποίες παρουσιάζουν υψηλό λόγο σήματος προς θόρυβο στενής ζώνης λαμβάνουν μεγαλύτερο βάρος, με αποτέλεσμα την αύξηση της ανθεκτικότητας του συστήματος στο θόρυβο.

Ένα πλεονέκτημα που υφίσταται λόγω της χρήσης του διαμορφωτή δέσμης για την εκτίμηση πολλαπλών κατευθύνσεων άφιξης δύναται να οδηγήσει στον διαχωρισμό των ηχητικών πηγών. Ακολουθώντας, τα δεδομένα που συλλέγονται από τις ξεχωριστές πηγές αξιοποιούνται ως μετρική για τη χαρτογράφηση κατά την αναζήτηση στο πλέγμα.

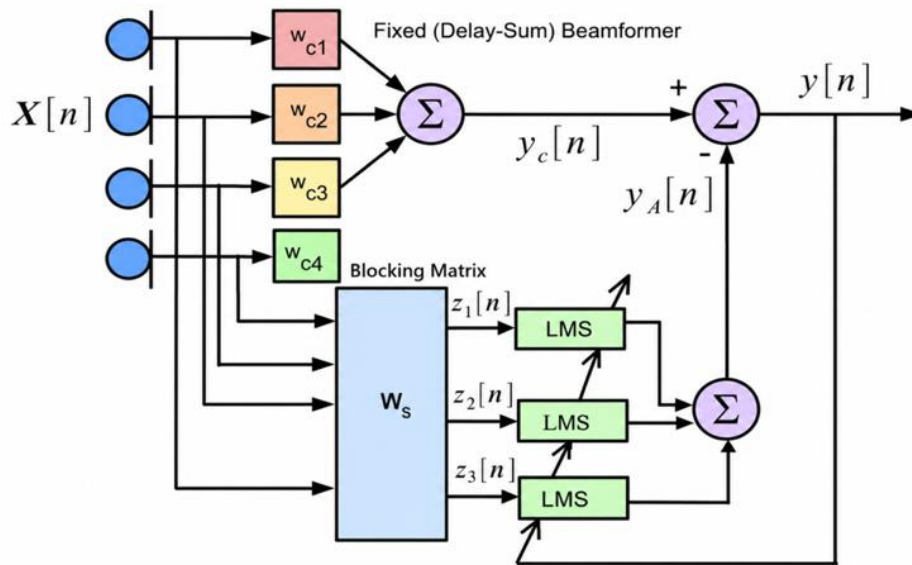
Σε εφαρμογές που απαιτείται να γίνει εντοπισμός της ανθρώπινης φωνής, χρησιμοποιείται μια μετρική φωνητικής ομοιότητας για την εξομάλυνση του κατευθυνόμενου φάσματος απόκρισης. Με τον τρόπο αυτό, οι έντονες κορυφές αναγνωρίζονται ως πηγές ανθρώπινης φωνής και δεν εκλαμβάνονται ως θόρυβος. Ενδεικτικά, για την ανίχνευση της φωνητικής δραστηριότητας (VAD), τα ανθρώπινα φωνήεντα παράγουν ήχους με συγκεκριμένο συνδυασμό κορυφών στο συχνοτικό φάσμα. Αυτός ο συνδυασμός των κορυφών (J), διαφέρει από τους άλλους τύπους φασμάτων, διευκολύνοντας έτσι την ανίχνευση της ανθρώπινης φωνής. Η διαδικασία αυτή υλοποιείται μέσω της καταγραφής ηχογραφήσεων φωνηέντων από ανθρώπινες ομιλίες, επιτρέποντας την κατηγοριοποίηση προτύπων υπογραφών για τις κορυφές στο συχνοτικό φάσμα. Τα εν λόγω πρότυπα λαμβάνουν την τιμή ένα για το σύνολο των κορυφών του φάσματος. (Lim, 2015)

Η ομοιότητα ανάμεσα στην έξοδο του μορφοποιητή δέσμης για μια προτεινόμενη κατεύθυνση άφιξης ($\widehat{S_\theta}[f]$) και στις κορυφές του συχνοτικού φάσματος (J), πραγματοποιείται μέσω της διαφοράς κοιλιάδας και κορυφής (PVD), όπως παρουσιάζεται στην εξίσωση που ακολουθεί

$$\text{PVD} = \frac{\sum_{f=1}^F \widehat{S_\theta}[f]}{\sum_{f=1}^F J[f]} - \frac{\sum_{f=1}^F \widehat{S_\theta}(1-J[f])}{\sum_{f=1}^F J[f]} \quad (3.19)$$

Η διαφορά κοιλιάδας και κορυφής (PVD) υπολογίζει την απόσταση μεταξύ δύο φασμάτων, ως τη διαφορά της μέσης ενέργειας των κορυφών που εμφανίζονται σε αυτά σε σχέση με τη μέση ενέργεια των υπόλοιπων συχνοτικών ζωνών.

Στη συνέχεια, παρουσιάζεται ένας άλλος διαμορφωτής δέσμης, ο γενικευμένος ακύρωτης πλευρικών λοβών (GSC). Ο διαμορφωτής αυτός εστιάζει στη μέθοδο για τον προσδιορισμό των βέλτιστων κατευθύνσεων άφιξης των σημάτων που λαμβάνονται από τα μικρόφωνα, εξασφαλίζοντας μια ομαλή απόκριση φάσματος ώστε να πραγματοποιηθεί η ενίσχυση τους. Επιπλέον, ο γενικευμένος ακύρωτης πλευρικών λοβών είναι σχεδιασμένος να προσαρμόζεται στις μεταβολές του θορύβου.



Εικόνα 3.9 Απεικόνιση δομικού διαγράμματος γενικευμένου ακύρωτη πλευρικών λοβών. (Mamun, 2012)

Όπως παρουσιάζεται στο παραπάνω σχήμα, ο γενικευμένος ακύρωτης πλευρικών λοβών χρησιμοποιεί τον διαμορφωτή δέσμης καθυστέρησης και αθροίσματος στην κύρια διαδρομή του ηχητικού σήματος. Στη δευτερεύουσα διαδρομή πραγματοποιείται εκτίμηση του θορύβου, ο οποίος στη συνέχεια αφαιρείται από την έξοδο του διαμορφωτή. Για το σκοπό αυτό χρησιμοποιείται ένας πίνακας αποκλεισμού, ο οποίος υπολογίζει τον θόρυβο για διαφορετικές κατευθύνσεις από την επιθυμητή κατεύθυνση άφιξης αποκλείοντας κυρίως τα σήματα της κύριας διαδρομής. Τα ηχητικά σήματα που προκύπτουν από την παραπάνω διαδικασία, φιλτράρονται, ειδικά κατά τη διαδικασία της πρόσθεσης, για την παραγωγή της τελικής εκτίμησης του θορύβου, η οποία δύναται να αφαιρεθεί απευθείας από την έξοδο του διαμορφωτή καθυστέρησης και αθροίσματος. Τα φίλτρα βελτιώνονται συνεχώς με την πάροδο του χρόνου, με βάση το αλγόριθμο ελαχίστων μέσων τετραγώνων (LMS), χρησιμοποιώντας την τρέχουσα εκτιμώμενη έξοδο και τις προηγούμενες εκτιμήσεις για το θόρυβο, γεγονός που επιτρέπει την προσαρμογή στις περιβαλλοντικές μεταβολές. (Lim, 2015)

Μια επιπλέον μέθοδος βασίζεται στην επιλογή συχνοτήτων στην έξοδο του διαμορφωτή δέσμης, ώστε να πραγματοποιηθεί αφαίρεση του εξασθενημένου θορύβου. Αρχικά, υπολογίζεται το φάσμα της κατευθυνόμενης απόκρισης που προκύπτει από τον διαμορφωτή καθυστέρησης και αθροίσματος. Μέσω αυτών των υπολογισμών εκτιμάται η ηχητική πηγή με την υψηλότερη ενέργεια κατεύθυνσης άφιξης, καθώς και η πηγή με τη δεύτερη υψηλότερη τιμή. Στη συνέχεια, φιλτράρονται οι συχνότητες που παρουσιάζουν χαμηλότερες τιμές στην έξοδο του διαμορφωτή δέσμης μέγιστης ενέργειας, σε σύγκριση με τις συχνότητες της δεύτερης υψηλότερης τιμής. Ακολουθεί η αφαίρεση μεταξύ της φιλτραρισμένης εξόδου και της εξόδου των διαμορφωτών δέσμης, με τη διαδικασία αυτή να επαναλαμβάνεται μέχρι να απομείνει αποκλειστικά ο θόρυβος υποβάθρου.

Το πλεονέκτημα αυτής της μεθόδου έγκειται στη μεγάλη ευαισθησία της ως προς τον εντοπισμό ηχητικών πηγών χαμηλής ενέργειας. Μοναδικό κριτήριο για την επίτευξη αυτής της ευαισθησίας είναι η ενέργεια από τις ηχητικές πηγές να υπερβαίνει την αντίστοιχη του θορύβου υποβάθρου.

Συμπερασματικά, αξίζει να σημειωθεί, ότι ένα μειονέκτημα της μεθόδου διαμόρφωσης δέσμης για πολλαπλή εκτίμηση κατεύθυνσης άφιξης είναι η ανάγκη ύπαρξης ενός διαμορφωτή δέσμης για κάθε προτεινόμενη κατεύθυνση. Κατά συνέπεια, η υλοποίηση της συγκεκριμένης μεθόδου απαιτεί ιδιαίτερα υψηλή υπολογιστική ισχύ.

3.4.3.2 Παρουσίαση των μεθόδων υποχώρου

Μια χαρακτηριστική μέθοδος υποχώρου είναι η ταξινόμησης πολλαπλών σημάτων (MUSIC). Η μέθοδος αυτή βασίζεται στην αναζήτηση της κατεύθυνσης άφιξης, η οποία τέμνει το υποχώρο που αντιπροσωπεύει τα σημαντικότερα ηχητικά σήματα.

Αρχικά παρουσιάζεται το μοντέλο σήματος μέσω της εξίσωσης που ακολουθεί

$$X = W_s S + V \quad (3.20)$$

όπου

X : μιγαδικός πίνακας με διαστάσεις $N \times F$, ο οποίος περιλαμβάνει στις γραμμές του τα διανύσματα X_n
 X_n : τα μετασχηματισμένα διακριτά σήματα των εισόδων (x_n), που αντιστοιχούν στο λαμβανόμενο ηχητικό σήμα από τον αισθητήρα n

S : μιγαδικός πίνακας με διαστάσεις $D \times F$, ο οποίος περιέχει στις γραμμές του το σύνολο των σημάτων της πηγής στο πεδίο της συχνότητας

V : μιγαδικός πίνακας με διαστάσεις $N \times F$, ο οποίος περιέχει στις γραμμές του το σύνολο των σημάτων θορύβου που καταγράφει κάθε μικρόφωνο, στο πεδίο της συχνότητας

$W_s[f]$: μιγαδικός πίνακας με διαστάσεις $N \times D$, ο οποίος αποδίδει τη χρονική διαφορά της άφιξης για κάθε σήμα της ηχητικής πηγής (d) για δεδομένη συχνότητα (f) η οποία σχετίζεται με την κατεύθυνση άφιξης.

Ο πίνακας που αποδίδει τη χρονική διαφορά της άφιξης για κάθε σήμα της ηχητικής πηγής παρουσιάζεται στην συνέχεια

$$W_s[f] = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-2\pi f_1 \tau_{2:1}} & e^{-2\pi f_2 \tau_{2:2}} & \dots & e^{-2\pi f_N \tau_{2:D}} \\ e^{-2\pi f_1 \tau_{3:1}} & e^{-2\pi f_2 \tau_{3:2}} & \dots & e^{-2\pi f_N \tau_{3:D}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi f_1 \tau_{N:1}} & e^{-2\pi f_2 \tau_{N:2}} & \dots & e^{-2\pi f_N \tau_{N:D}} \end{bmatrix} \quad (3.21)$$

Ο πίνακας $W_s[f]$ αναπαριστά τη χρονική διαφορά άφιξης για τα ηχητικά σήματα που λαμβάνονται από κάθε ηχητική πηγή. Η συγκεκριμένη μέθοδος βασίζεται στην υπόθεση εργασίας ότι το σήμα είναι στενής ζώνης με κεντρική συχνότητα f , παραδοχή η οποία δεν ισχύει για τα ηχητικά σήματα που σχετίζονται με την ανθρώπινη ομιλία. Ωστόσο, στη συνέχεια της μελέτης και χάριν απλοποίησης, υιοθετείται ότι αυτή η παραδοχή εφαρμόζεται και για την ανθρώπινη ομιλία.

Αρχικά, η μέθοδος MUSIC πραγματοποιεί την εκτίμηση των υποχώρων που παράγονται από τις στήλες του πίνακα W_s . Η διαδικασία αυτή βασίζεται στον υπολογισμό του πίνακα συνδιακύμανσης των δειγμάτων $\hat{R}[f]$ με διαστάσεις $N \times N$, για τη συχνότητα των καταγεγραμμένων ηχητικών σημάτων, η οποία παρουσιάζεται από την εξίσωση που ακολουθεί

$$\hat{R}[f] = Q[f]\Lambda[f]Q[f]^{-1} \quad (3.22)$$

όπου

$\Lambda[f]$: διαγώνιος πίνακας μιγαδικών τιμών διαστάσεων $N \times N$ με ιδιοτιμές του πίνακα συνδιακύμανσης για τη συχνότητα f

Οι ιδιοτιμές αυτές είναι τοποθετημένες σε φθίνουσα σειρά, όπως παρουσιάζεται από την παραπάνω εξίσωση

$$\Lambda[f] = \begin{bmatrix} \lambda_1[f] & 0 & \cdots & 0 \\ 0 & \lambda_2[f] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N[f] \end{bmatrix} \quad (3.23)$$

$Q[f]$: πίνακας μιγαδικών τιμών διαστημάτων $N \times N$, όπου στις στήλες του τα διατεταγμένα ιδιοδιανύσματα του πίνακα συνδιακύμανσης για την συχνότητα f , όπως καθορίζεται από τον πίνακα Λ . Ο πίνακας αυτός παρουσιάζεται στην συνέχεια

$$Q[f] = \begin{bmatrix} q_{1:1}[f] & q_{2:1}[f] & \cdots & q_{N:1}[f] \\ q_{1:2}[f] & q_{2:2}[f] & \cdots & q_{N:2}[f] \\ \vdots & \vdots & \ddots & \vdots \\ q_{1:N}[f] & q_{2:N}[f] & \cdots & q_{N:N}[f] \end{bmatrix} \quad (3.24)$$

Ο πίνακας $Q[f]$ έχει τη δυνατότητα να διαιρεθεί σε δύο υποχώρους με κριτήριο τις ιδιοτιμές λ_i , με τη σχετική διαμέριση να παρουσιάζεται με τις παρακάτω εξισώσεις

$$Q_s[f] = Q[f][1:\lambda_l]$$

$$Q_u[f] = Q[f][\lambda_l + 1:N] \quad (3.25)$$

όπου

$Q_s[f]$: πίνακας διαστάσεων $N \times \lambda_l$, ο οποίος περιλαμβάνει στις στήλες του τα ιδιοδιανύσματα που παράγουν τον υποχώρο σήματος στη συχνότητα f

$Q_u[f]$: πίνακας διαστάσεων $N \times [N - \lambda_l]$, ο οποίος περιλαμβάνει στις στήλες του τα ιδιοδιανύσματα που παράγουν τον υποχώρο θορύβου στην συχνότητα f

Η παράμετρος λ_i δύναται να ορίζει την τάξη των $W_s[f]$, ωστόσο, στις περιπτώσεις όπου η εκτίμηση είναι δυσχερής όπως σε περιβάλλον με την παρουσία θορύβου, το λ_i ορίζεται ως ο δείκτης της ιδιοτιμής η οποία είναι μικρότερη συγκριτικά με τις υψηλές ιδιοτιμές, εκείνες δηλαδή τις τιμές που απέχουν σημαντικά από το μηδέν. Η επιλογή αυτή στηρίζεται στην παραδοχή, ότι οι ιδιοτιμές του υποχώρου θορύβου είναι μικρότερες από τις τιμές του υποχώρου σήματος. (Chen, 2003)

Μια εναλλακτική προσέγγιση αφορά τη δημιουργία ενός σταθερού και ομοιόμορφου θορύβου πριν από την υλοποίηση της ιδιοανάλυσης του πίνακα συνδιακύμανσης. Επιδίωξη είναι η ιδιοτιμή του υποχώρου θορύβου να είναι γύρω από τη μονάδα και αντιστοίχως η ιδιοτιμή του υποχώρου σήματος να λαμβάνει τιμή μεγαλύτερη από το ένα.

Για να διασφαλιστεί η διαμόρφωση αυτού του σταθερού θορύβου, οφείλεται αρχικά να πραγματοποιηθεί η εκτίμηση για τις τιμές του πίνακα συνδιακύμανσης διαστάσεων $N \times N$ για τον πίνακα $K[f]$ και το θόρυβο $V[f]$, με βασική προϋπόθεση ότι σε αυτή τη χρονική περίοδο οι ηχητικές πηγές δεν παράγουν ηχητικό σήμα.

Η $K[f]$ υπολογίζεται με την εξίσωση που παρουσιάζεται παρακάτω

$$K[f] = V[f]V[f]^H \quad (3.26)$$

Αντιστοίχως, η ιδιοανάλυση παρουσιάζεται με βάση την παρακάτω εξίσωση

$$K[f]^{-1}\hat{R}[f] = Q[f]\Lambda[f]Q[f]^{-1} \quad (3.27)$$

Συνοψίζοντας, ο πίνακας $K[f]^{-1}$ της ιδιοανάλυσης μπορεί να είναι οποιοσδήποτε ορισμένος πίνακας, ώστε το γινόμενο των πινάκων $K[f]^{-1}\hat{R}[f]$ να είναι ένας τετραγωνικός πίνακας. Η τεχνική αυτή είναι η γενικευμένη ιδιοανάλυση (GEVD), η οποία μπορεί να ενσωματωθεί στη μέθοδο MUSIC, με βασικό πλεονέκτημα ότι είναι ανθεκτική στο μη συσχετισμένο περιβαλλοντικό θόρυβο. (Nakamura, 2009)

Επιπλέον, προκειμένου να γίνει εκτίμηση της κατεύθυνσης άφιξης με την υπομέθοδο της γενικής μεθόδου MUSIC, απαιτείται να υπολογιστούν πολλαπλές υποψήφιες εκτιμήσεις. Ειδικότερα, υπολογίζονται οι τιμές του πίνακα διαστάσεων $N \times 1$ για το διάνυσμα $b[\theta, f]$ για κάθε υποψήφια τιμή της γωνίας θ . Το διάνυσμα αυτό παρουσιάζεται από την εξίσωση που ακολουθεί

$$b[\theta, f] = \begin{bmatrix} 1 \\ e^{-2\pi f \tau_1(\theta)} \\ e^{-2\pi f \tau_2(\theta)} \\ \vdots \\ e^{-2\pi f \tau_N(\theta)} \end{bmatrix} \quad (3.28)$$

όπου

N : ο αριθμός των μικροφώνων

$\tau_n(\theta)$: χρονική διαφορά άφιξης στο μικρόφωνο n για την υποψήφια γωνία θ

Ο υπολογισμός του $\tau_n(\theta)$ πραγματοποιείται πάντοτε με βάση το μοντέλο διάδοσης που θα χρησιμοποιηθεί στη μέθοδο.

Στην περίπτωση που το διάνυσμα $b[\theta, f]$ παρουσιάζει ορθογώνιο σχήμα ως προς το ιδιοδιάνυσμα $Q_n[f]$, τότε αντιστοιχεί σε μια από τις ηχητικές πηγές του υποχώρου σήματος $Q_s[f]$, και επιπλέον είναι

μια εκτίμηση για την κατεύθυνση του σήματος από την ηχητική πηγή. Προκειμένου να πραγματοποιηθεί η διαδικασία του ελέγχου ορθογωνικότητας του διανύσματος $b[\theta, f]$, χρησιμοποιείται η εξίσωση που παρουσιάζεται παρακάτω

$$P_{\text{MUSIC}}[\theta, f] = \frac{1}{b[\theta, f]^H Q_u[f] Q_u[f]^H b[\theta, f]} \quad (3.29)$$

όπου

$P_{\text{MUSIC}}[f]$: ψευδοφάσμα MUSIC για τη συχνότητα f

Οι θέσεις των κορυφών στη συνάρτηση του φάσματος προσδιορίζουν την κατεύθυνση άφιξης των ηχητικών πηγών. Στη μέθοδο αυτή, υφίσταται το ενδεχόμενο ο παρονομαστής της εξίσωσης να μηδενιστεί. Αυτό έχει ως αποτέλεσμα να μην μπορεί να οριστεί η συνάρτηση $P_{\text{MUSIC}}[\theta, f]$, γεγονός που δύναται να συμβεί στη περίπτωση που το διάνυσμα $b[\theta, f]$ καθίσταται ορθογώνιο ως προς το $Q_n[f]$. Εντούτοις, η συγκεκριμένη περίπτωση είναι ιδιαίτερα σπάνια, καθώς ο θόρυβος δεν συσχετίζεται απόλυτα με το ηχητικό σήμα της πηγής. Συνεπώς, οι τιμές του παρονομαστή βρίσκονται κοντά στο μηδέν, χωρίς όμως να ισούνται με αυτό, με αποτέλεσμα οι κορυφές της $P_{\text{MUSIC}}[f]$ να παρουσιάζουν ιδιαίτερα υψηλές τιμές.

Η μέθοδος MUSIC απαιτεί ιδιοδιανύσματα τα οποία ανήκουν στον υποχώρο του θορύβου. Ωστόσο, στην περίπτωση που ο πίνακας $Q_n[f]$ δεν είναι κενός, η μέθοδος δύναται να διαχειριστεί έως $N - 1$ ηχητικές πηγές. Μια λύση για την περίπτωση αυτή αποτελεί η χρήση μεγάλου αριθμού μικρόφωνων, συγκεκριμένα οκτώ ή και περισσότερων.

Επιπλέον, για την εφαρμογή της μεθόδου αυτής σε ηχητικά σήματα ομιλίας, απαιτείται ο υπολογισμός της σε μια στενή ζώνη του φάσματος για κάθε συχνότητα ξεχωριστά. Στη συνέχεια, μπορεί να εξαχθεί ο μέσος όρος της κατεύθυνσης άφιξης του σήματος σε μια ευρεία ζώνη του φάσματος.

Ωστόσο, ένα βασικό μειονέκτημα της μεθόδου MUSIC σε πραγματικό χρόνο είναι ότι χρειάζεται αυξημένους υπολογιστικούς πόρους. Η απαίτηση αυτή συμβαίνει κυρίως όταν εφαρμόζεται η γενικευμένη ιδιοανάλυση σε περιβάλλον με έντονες διακυμάνσεις του θορύβου. Μια λύση για την αντιμετώπιση του προβλήματος αυτού αποτελεί η χρήση της γενικευμένης ανάλυσης ιδιοζουσών τιμών (GSVD), η οποία εξασφαλίζει την απαιτούμενη ορθογωνικότητα για τη μέθοδο MUSIC. Στην συνέχεια, παρουσιάζεται η εξίσωση για την μέθοδο αυτή

$$K[f]^{-1} \hat{R}[f] = Q_l[f] \Lambda[f] Q_r[f]^H \quad (3.30)$$

όπου

$Q_l[f]$, $Q_r[f]$: τα αριστερά και δεξιά ιδιάζοντα διανύσματα για την συχνότητα f , τα οποία είναι ορθογώνια μεταξύ τους

Με τη μέθοδο της γενικευμένης ανάλυσης ιδιοζουσών τιμών της MUSIC απαιτούνται σημαντικά λιγότεροι υπολογιστικοί πόροι και αντίστοιχα μειωμένος χρόνος επεξεργασίας. Η μέθοδος αυτή παρουσιάζει ανθεκτικότητα σε περιβάλλον με θόρυβο, ακολουθώντας την εξής διαδικασία.

1. Με διόρθωση του πίνακα συνδιακύμανσης, δηλαδή διόρθωση των θορυβωδών φάσεων μέσω γραμμικής παλινδρόμησης με τις συστοιχίες της συχνότητας.

2. Αφαίρεση των τμημάτων που περιέχουν θόρυβο στη συχνότητα, μέσω εκτίμησης του λόγου σήματος προς τον θόρυβο, η οποία βασίζεται στον υποχώρο.

Τα δεδομένα που σχετίζονται με τη φάση του σήματος για τη μέθοδο αυτή περιλαμβάνονται στον πίνακα συνδιακύμανσης $\hat{R}[f]$. Η μέθοδος MUSIC δύναται να χρησιμοποιηθεί για την εκτίμηση πολλαπλών κατευθύνσεων άφιξης μέσω μιας συστοιχίας μικροφώνων, με αποτέλεσμα ο πίνακας $X[f]$ να διαθέτει τόσες σειρές όσα και τα χρησιμοποιούμενα μικρόφωνα. Για τη μέθοδο αυτή, η σειρά $X_1[f]$ αντιστοιχεί στο σήμα του μικροφώνου αναφοράς, με συνέπεια να ισχύει η εξίσωση που παρουσιάζεται στην συνέχεια

$$X_2[f] = X_1[f]e^{-j2\pi f\tau} \quad (3.31)$$

Στη συνέχεια, παρουσιάζεται ο πίνακας των δεδομένων της φάσης που προκύπτουν από τη μέθοδο που χρησιμοποιήθηκε

$$\hat{R}[f] = \begin{bmatrix} X_1[f]^2 & X_1[f]^2 e^{i2\pi f\tau} \\ X_1[f]^2 e^{-i2\pi f\tau} & X_2[f]^2 \end{bmatrix} \quad (3.32)$$

όπου

$\zeta(f, \tau) = e^{-i2\pi f\tau}$: τα δεδομένα φάσης, τα οποία εκφράζουν την αντιστοιχία ανάμεσα στη συχνότητα f και το χρόνο τ , και παρουσιάζουν γραμμική σχέση στη περίπτωση απουσίας θορύβου.

Σε περιβάλλον με θόρυβο, ενδέχεται να δημιουργηθούν σφάλματα για την εκτίμηση της κατεύθυνσης άφιξης. Σε αυτή την περίπτωση, για να διορθωθούν τα δεδομένα φάσης και να περιοριστούν τα σφάλματα, μπορεί να εφαρμοστεί η παρακάτω διαδικασία

- α) Αρχικά υπολογισμός του δείγματος συνδιακύμανσης για τον πίνακα $R[f]$.
- β) Στη συνέχεια, δημιουργία των δεδομένων φάσης $\zeta(f, \tau)$ από τον πίνακα $\hat{R}[f]$ για τη συχνότητα f
- γ) Έπειτα, εκτέλεση επαναλαμβανόμενης παλινδρόμησης για μείωση του σφάλματος, η οποία περιλαμβάνει τα εξής στάδια.
 - I. Εφαρμογή επαναλαμβανόμενης παλινδρόμησης στα δεδομένα φάσης για τη δημιουργία γραμμικής φάσης.
 - II. Αναπαράσταση των δεδομένων φάσης με χαμηλότερη τιμή σε σχέση με την εκτιμώμενη γραμμή φάσης.
 - III. Σύγκριση μεταξύ των ενημερωμένων δεδομένων φάσης με την εκτιμώμενη γραμμή φάσης, με σκοπό τον υπολογισμό του σφάλματος.
- δ) Τέλος, επαναδημιουργία του πίνακα συνδιακύμανσης με βάση τα ενημερωμένα δεδομένα φάσης $\zeta(f, \tau)$

Η επιλογή των συχνοτικών ζωνών βασίζεται στη μετρική, όπως παρουσιάζεται από την εξίσωση που ακολουθεί

$$\varepsilon[f] = \log \left(\frac{\lambda_1 \hat{R}[f]}{\lambda_2 \hat{R}[f]} \right) \quad (3.33)$$

όπου

$\lambda_1[f]$: πρώτη ιδιοτιμή του πίνακα $\hat{R}[f]$, η οποία αντιστοιχεί στον υποχώρο σήματος

$\lambda_2[f]$: δεύτερη ιδιοτιμή του πίνακα $\hat{R}[f]$, η οποία αντιστοιχεί στον υποχώρο θορύβου

Η συνάρτηση $\varepsilon(f)$ εκφράζει τον λόγο σήματος προς θόρυβο στενής ζώνης στον υποχώρο κατά τον υπολογισμό του πίνακα $\hat{R}[f]$ για τη συχνότητα f . Σχετικά με το πρόβλημα των ζωνών συχνότητας, μπορεί να οριστεί ένα κριτήριο στη μέθοδο, ώστε να αποφεύγεται η χρήση εκείνων των ζωνών που παρουσιάζουν χαμηλό λόγο σήματος προς θόρυβο. (Vincent, 2015)

3.4.3.3 Παρουσίαση των μεθόδων χρονικής ομαδοποίησης των ηχητικών πηγών

Οι μέθοδοι εκτίμησης πολλαπλών κατευθύνσεων άφιξης στηρίζονται πρωτίστως σε ένα συγκεκριμένο χρονικό διάστημα. Αυτό το διάστημα καθορίζεται από έναν προκαθορισμένο αριθμό δειγμάτων και ορίζεται ως χρονικό παράθυρο.

Παράλληλα, οι τεχνικές χρονικής ομαδοποίησης των ηχητικών πηγών βασίζονται αρχικά στον διαχωρισμό των κατευθύνσεων άφιξης ανά χρονικό παράθυρο και στη συνέχεια, στη δημιουργία συστάδων με διαφορετικές κατευθύνσεις άφιξης. Η κάθε συστάδα μπορεί να αντιπροσωπεύει μόνο μια ηχητική πηγή.

Στο πλαίσιο αυτό, μια συγκεκριμένη μέθοδο εκτίμησης κατεύθυνσης άφιξης αποτελεί η εκτίμηση πλεονάζουσας κατεύθυνσης άφιξης, η οποία βασίζεται στον υπολογισμό της χρονικής διαφοράς άφιξης για τρία ζεύγη μιας τριγωνικής διάταξης μικροφώνων. Η διαδικασία αυτή υλοποιείται με τη χρήση μιας εναλλακτικής μεθόδου γενικευμένου μετασχηματισμού φάσης διασταυρούμενης συσχέτισης (GCC-PHAT). Με χρήση των μετρήσεων της πλεονάζουσας κατεύθυνσης άφιξης και με συνθήκες μη επικαλυπτόμενης ανθρώπινης ομιλίας, ο εκτιμητής δύναται να προχωρήσει σε εκτίμηση κατεύθυνσης άφιξης για μια ηχητική πηγή, ακόμη και όταν στο περιβάλλον συνυπάρχουν πολλοί ομιλητές.

Ο προσδιορισμός μιας κατεύθυνσης άφιξης για το χρονικό παράθυρο $t + 1$ είναι δυνατόν να προκύψει από μια διαφορετική ηχητική πηγή από αυτή στην οποία πραγματοποιείται η τρέχουσα εκτίμηση. Συγχρόνως, υφίσταται η δυνατότητα εφαρμογής της μεθόδου ομαδοποίησης με τεχνικές που βασίζονται στην παρακολούθηση μέσω μικροφώνων. Οι τεχνικές αυτές παράγουν δεδομένα από τις διαδοχικά εκτιμώμενες μεμονωμένες κατευθύνσεις άφιξης για κάθε πηγή ήχου. Σε κάθε χρονικό παράθυρο, τα δεδομένα από όλες τις κατευθύνσεις άφιξης τροφοδοτούν ένα φίλτρο Kalman. Συνεπώς, η μέθοδος μπορεί να περιέχει μια εκτίμηση πολλαπλών κατευθύνσεων άφιξης, βασιζόμενη στην πρόβλεψη που προκύπτει από τα δεδομένα αυτά.

Μια επιπλέον προσέγγιση για την εκτίμηση πολλαπλών κατευθύνσεων άφιξης βασίζεται σε έναν μηχανισμό ιχνηλάτησης, ο οποίος αξιοποιεί ένα μοντέλο μείγματος Gauss. Συγκεκριμένα, χρησιμοποιείται ένα φίλτρο υπόθεσης πυκνότητας πιθανότητας μείγματος Gauss (GM-PHD), ώστε να γίνει εκτίμηση σε πραγματικό χρόνο της συνάρτησης πυκνότητας πιθανότητας για τις πιθανές θέσεις των ηχητικών πηγών. Για τον λόγο αυτό, σε κάθε χρονικό παράθυρο πραγματοποιείται ανανέωση της συνάρτησης στάθμης της ακουστικής πίεσης με νέες εκτιμήσεις κατεύθυνσης άφιξης. Κατά τη διαδικασία αυτή, συνυπολογίζονται τόσο οι νέες εκτιμήσεις της κατεύθυνσης άφιξης όσο και οι αντίστοιχες νέες εκτιμήσεις αναφοράς.

Ένα πλεονέκτημα αυτής της μεθόδου είναι ότι η συγκεκριμένη τεχνική δύναται να παρέχει κατά προσέγγιση την απόσταση κάθε ηχητικής πηγής, σε συνδυασμό με την αντίστοιχη εκτίμηση της κατεύθυνσης άφιξης. Συνεπώς, η παρούσα μέθοδος επιτρέπει τον συσχετισμό της απόστασης με τη χρονική απόκλιση.

Μια συμπληρωματική επιλογή για την εκτίμηση της κατεύθυνσης άφιξης αποτελεί η χρήση μιας προσαρμοστικής παραλλαγής της μεθόδου ομαδοποίησης K-Means++. Ο συγκεκριμένος αλγόριθμος βασίζεται σε κέντρα δεδομένων που προκύπτουν από ένα καθορισμένο σύνολο πιθανοτήτων, οι οποίες εξάγονται από τα ίδια τα διαθέσιμα δεδομένα.

Αντιστοίχως, οι προσεγγίσεις που βασίζονται σε συστάδες για την εκτίμηση της κατεύθυνσης άφιξης έχουν ως βάση τις συστάδες εκτίμησης κατεύθυνσης άφιξης σε βάθος χρόνου. Μέσω της προσέγγισης αυτής πραγματοποιείται εκτίμηση για μια μόνο ηχητική πηγή ανά χρονικό παράθυρο. Το κύριο πλεονέκτημα αυτής της επιλογής είναι ότι απαιτεί αποκλειστικά τα μικρόφωνα που χρησιμοποιούνται για την εκτίμηση της κατεύθυνσης άφιξης, μειώνοντας δραστικά τον απαιτούμενο αριθμό τους συγκριτικά με άλλες μεθόδους. Ένα επιπλέον πλεονέκτημα αυτής της μεθόδου αφορά τη δυνατότητα ενσωμάτωσης αλγορίθμων φιλτραρίσματος που αξιοποιούν συστάδες, όπως τα φίλτρα Kalman ή τα φίλτρα PHD, τα οποία περιορίζουν τις θορυβώδεις κατευθύνσεις άφιξης στο τελικό αποτέλεσμα.

Στον αντίποδα, ένα μειονέκτημα της προσέγγισης αυτής αποτελεί ότι η εκτίμηση της κατεύθυνσης άφιξης, χρειάζεται χρονικά επαναλαμβανόμενους υπολογισμούς και πολλαπλά χρονικά παράθυρα. Η ιδιαιτερότητα αυτή έχει ως συνέπεια ότι μπορεί να προκαλέσει πρόβλημα χαμηλής απόκρισης στο σύστημα.

Τέλος, οι τεχνικές που απαιτούν τον υπολογισμό ενός πίνακα συνδιακύμανσης, όπως ο αλγόριθμος MUSIC ή ο MVDR, εξασφαλίζουν μια αξιόπιστη εκτίμηση για πολλαπλές κατευθύνσεις άφιξης. Το οποίο επιτυγχάνεται έπειτα από επαρκή χρονικά παράθυρα κατά τα οποία ο πίνακας συνδιακύμανσης υπολογίζεται εναλλακτικά. (Pourmehr, 2015).

3.4.4 Εκτίμηση της απόστασης της ηχητικής πηγής

Ένα επιπλέον χαρακτηριστικό για τον εντοπισμό της ηχητικής πηγής αποτελεί η απόσταση της από τα μικρόφωνα. Μια συνήθεις τεχνική εκτίμησης της απόστασης βασίζεται στην παραδοχή ότι η πηγή βρίσκεται πάνω στην υπερβολή που ορίζεται από τη χρονική διαφορά άφιξης των λαμβανόμενων ηχητικών σημάτων σε ένα ζεύγος αισθητήρων. Στην περίπτωση, ωστόσο, που η διάταξη περιλαμβάνει πολλαπλά μικρόφωνα, η θέση της ηχητικής πηγής προσδιορίζεται από το σημείο τομής των υπερβολικών καμπύλων που δημιουργούνται για κάθε συνδυασμό αισθητήρων.

Επομένως, το δυναμικό εύρος της χρονικής διαφοράς άφιξης, το οποίο δημιουργείται, λόγω της μεταβολής της απόστασης, είναι μη γραμμικό για κοντινές αποστάσεις και είναι ιδιαίτερα μικρό για μακρινές. Στην περίπτωση, ωστόσο, που μια ηχητική πηγή βρίσκεται σε μεγάλη απόσταση, μια περαιτέρω αλλαγή της θέσης της ως προς τα μικρόφωνα ελάχιστα επιδρά στη χρονική διαφορά άφιξης. Αυτό έχει ως αποτέλεσμα να μην δημιουργείται κάποια αξιοσημείωτη αλλαγή στην τομή των υπερβολικών καμπύλων.

Παράλληλα, αυτό το χαρακτηριστικό είναι η κύρια αιτία κατά την εμφάνιση σημαντικών αποκλίσεων για την εκτίμηση της απόστασης της ηχητικής πηγής. Ένα επιπλέον μειονέκτημα της συγκεκριμένης μεθόδου είναι ότι διάφορα ηχητικά φαινόμενα όπως οι ανακλάσεις και οι αυξομειώσεις στην ένταση της ακουστικής πίεσης του σήματος, ενδέχεται να μεταβάλουν την ακρίβεια της εκτίμησης. (Huang, 1999)

Η αξιοποίηση τεχνικών μηχανικής μάθησης αποτελεί μια εναλλακτική προσέγγιση για τον προσδιορισμό της απόστασης της ηχητικής πηγής. Τα στοιχεία μπορούν να συλλεγούν τόσο από το λαμβανόμενο ακουστικό σήμα όσο και από πηγές ήχου που βρίσκονται σε διαφορετικές αποστάσεις.

Το σύνολο των δεδομένων αυτών τροφοδοτείται ως περιβάλλον εκπαίδευσης στον αλγόριθμο, ώστε να μπορεί να γίνει η εκτίμηση απόστασης της ηχητικής πηγής.

Ένας περιορισμός της μεθόδου αυτής σχετίζεται με την παραδοχή της ανεξάρτητης και ταυτόσημης κατανομής (IID). Η διαφορά της στάθμης ακουστικής πίεσης του ηχητικού σήματος μεταξύ των ηχητικών πηγών δύναται να είναι μεγαλύτερη από την ίδια την επίδραση της πηγής ήχου, όταν η καταγραφή πραγματοποιείται σε διαφορετικές αποστάσεις. Κατά συνέπεια, η παραδοχή της ανεξάρτητης και ταυτόσημης κατανομής εμφανίζει μειωμένη ευαισθησία στις μεταβολές της απόστασης στα εύρη λειτουργίας των ρομποτικών συστημάτων. Για την αντιμετώπιση του προβλήματος αυτού, προτείνεται η εισαγωγή ενός συντελεστή διόρθωσης, ο οποίος σχετίζεται με το αζιμούθιο της ηχητικής πηγής, ώστε να εισαχθεί η απαιτούμενη μεταβλητότητα ως προς το λαμβανόμενο ηχητικό σήμα.

Μια εναλλακτική προσέγγιση είναι η εφαρμογή της μεθόδου γονιμοποίησης, η οποία μπορεί να εφαρμοστεί αποκλειστικά με κινούμενα ρομπότ. Στην περίπτωση αυτή, η εκτίμηση της απόστασης της ηχητικής πηγής υπολογίζεται τριγωνομετρικά, υπό την προϋπόθεση ότι η κατεύθυνση της ηχητικής πηγής μπορεί να εκτιμηθεί σε διαφορετικές προκαθορισμένες θέσεις στο χώρο. Επιπλέον, για την εκτίμηση της απόστασης δύναται να γίνει χρήση συστοιχιών μικροφώνων, οι οποίες χωρίζονται σε επιμέρους υποσυστοιχίες. Κατά τη διαδικασία αυτή, κάθε υποσυστοιχία εκτιμά την κατεύθυνση άφιξης και στην συνέχεια, η απόσταση υπολογίζεται μέσω τριγωνομετρικών μεθόδων. (Rodemann, 2010)

Μια εναλλακτική προσέγγιση για την εκτίμηση της απόστασης της πηγής ήχου, στηρίζεται στην εφαρμογή τεχνικών βελτιστοποίησης εντός του καρτεσιανού συστήματος συντεταγμένων. Πιο συγκεκριμένα, ο εντοπισμός της θέσης της ηχητικής πηγής καθίσταται εφικτός με βάση τη χρονική διαφορά άφιξης και τη μέθοδο γενικευμένου μετασχηματισμού φάσης διασταυρούμενης συσχέτισης. Η επίλυση του εν λόγω προβλήματος επιτυγχάνεται μέσω της χρήσης των πολλαπλασιαστών Lagrange. Ακολουθώντας, εφαρμόζεται η τεχνική της γραμμικής διόρθωσης ελαχίστων τετραγώνων. Η μέθοδος αυτή επεξεργάζεται δεδομένα, τα οποία προέρχονται από έναν ελαχίστως παραμετρικό αυτοοργανούμενο χάρτη (PL-SOM) και είναι δυνατή η αξιοποίηση της διαφοράς στάθμης ακουστικής πίεσης μεταξύ των μικροφώνων (RIID). Μέσω της διαδικασίας αυτής, μια από τις παραγόμενες εξόδους λαμβάνει μια τιμή που καθορίζει το μέγεθος της απόστασης.

Μια συμπληρωματική τεχνική για την εκτίμηση της απόστασης της ηχητικής πηγής βασίζεται στο φαινόμενο των ανακλάσεων του ήχου. Στην περίπτωση μιας κινούμενης ηχητικής πηγής, η απόσταση της από τα μικρόφωνα αλλάζει διαρκώς εξαιτίας της κίνησης, η ενέργεια που προέρχεται από τις ανακλάσεις του χώρου δύναται να θεωρηθεί σταθερή. Ο λόγος που προκύπτει από την σύγκριση της ενέργειας απευθείας διάδοσης και της αντίστοιχης των ανακλάσεων ορίζεται ως αναλογία άμεσης ενέργειας προς αντήχηση (DRR). Η σχέση αυτή μεταξύ της ενέργειας που λαμβάνεται απευθείας από την κινούμενη ηχητική πηγή και την ενέργεια των ανακλάσεων είναι ανάλογη με την απόσταση της ηχητικής πηγής. Παρά ταύτα, ένα βασικό μειονέκτημα της συγκεκριμένης προσέγγισης είναι η προαπαιτούμενη γνώση των χαρακτηριστικών του χώρου και διεξαγωγή προκαταρκτικών μετρήσεων πριν από την εφαρμογή της.

Αναφορικά με την αξιοποίηση της μεθόδου αναλογίας άμεσης ενέργειας προς την αντήχηση για κινούμενη πηγή ήχου, είναι δυνατή η εφαρμογή μιας τεχνικής εξίσωσης ακύρωσης που βασίζεται στον προσδιορισμό της κατεύθυνσης άφιξης. Με την προσέγγιση αυτή επιτυγχάνεται ο εντοπισμός και η ανάλυση των δειγμάτων του ηχητικού σήματος, γεγονός που διευκολύνει τον ακριβή υπολογισμό της ενέργειας απευθείας διάδοσης από την ηχητική πηγή.

Στην πράξη, εντούτοις, η προσέγγιση αυτή έχει το μειονέκτημα, ότι η ενέργεια της αντήχησης μεταβάλλεται σε σχέση με τον χρόνο. Προκειμένου να ξεπεραστεί η δυσκολία αυτή, προτείνεται η εφαρμογή ενός μοντέλου μίξης Γκαουσιανών (GMM). Η συγκεκριμένη προσέγγιση επιτρέπει την παρουσίαση της αλληλεξάρτησης μεταξύ της απόστασης της ηχητικής πηγής και του λόγου της άμεσης ενέργειας που προέρχεται από την πηγή προς την ανακλώμενη ενέργεια.

Παράλληλα, μπορούν να επιστρατευτούν διάφορες εναλλακτικές προσεγγίσεις για τον προσδιορισμό της απόστασης της ηχητικής πηγής. Μια περίπτωση είναι η εφαρμογή αλγορίθμων μηχανικής

μάθησης, όπου ο υπολογισμός της απόστασης πραγματοποιείται με τεχνικές εκτίμησης θέσης στο καρτεσιανό σύστημα συντεταγμένων. Με την αξιοποίηση της αρχιτεκτονικής των τεχνητών νευρωνικών δικτύων, καθίσταται εφικτή η εκτίμηση της καρτεσιανής θέσης της ηχητικής πηγής ακόμα και με χρήση ενός μόνο μικροφώνου. (Lu, 2010)

3.5 Παρουσίαση μεθόδων παρακολούθησης της ηχητικής πηγής

Οι τεχνικές εντοπισμού ηχητικής πηγής, είτε πρόκειται για μια είτε για περισσότερες κατευθύνσεις προσέγγισης, βασίζονται συνήθως σε ένα μοναδικό χρονικό παράθυρο για την εκτίμηση της κατεύθυνσης άφιξης. Προκειμένου να πραγματοποιηθεί εκτίμηση για πολλαπλά χρονικά παράθυρα, μπορούν να ληφθούν δεδομένα μέσα από τη διαρκή καταγραφή της ηχητικής πηγής στην πορεία του χρόνου. Η διαδικασία αυτή επιτρέπει τον προσδιορισμό της μιας τελικής θέσης, μέσω του υπολογισμού πολλαπλών εκτιμήσεων ανά παραθύρου με τη χρήση ενός προκαθορισμένου μοντέλου κίνησης.

Παράλληλα, παρέχεται η δυνατότητα πρόβλεψης της μελλοντικής τροχιάς της ηχητικής πηγής για τα επόμενα χρονικά παράθυρα, η οποία εξαρτάται κατά βάση από το δυναμικό μοντέλο κίνησης που ενσωματώνει η εκάστοτε τεχνική παρακολούθησης. Κατά συνέπεια, το αποτέλεσμα της παρακολούθησης είναι μια φιλτραρισμένη εκτίμηση της θέσης για το εκάστοτε παράθυρο.

Στην περίπτωση, ωστόσο, που εφαρμόζεται ένα ευσταθές μοντέλο κίνησης, οι παρεμβολές και ο περιβαλλοντικός θόρυβος μπορούν να ενσωματωθούν στο σύστημα και να εξουδετερωθούν. Το γεγονός αυτό βελτιώνει την ακρίβεια εντοπισμού, τόσο για μια όσο και για πολλαπλές ηχητικές πηγές ταυτόχρονα.

Οι τεχνικές που χρησιμοποιούνται για την παρακολούθηση ηχητικών πηγών σε ρομποτικά συστήματα, παρουσιάζονται παρακάτω, και οι οποίες θα αναλυθούν στη συνέχεια

- α. Τα φίλτρα Kalman (Kalman Filters)
- β. Τα φίλτρα σωματιδίων (Particle Filters)
- γ. Και εναλλακτικές μέθοδοι

3.5.1 Τα φίλτρα Kalman

Τα φίλτρα Kalman αποτελούν μια κατηγορία φίλτρων Bayes και αποσκοπούν στον προσδιορισμό της κατάστασης ενός συστήματος μέσω των θορυβωδών μετρήσεων στη διάρκεια ενός χρονικού διαστήματος. Στο πλαίσιο αυτό, ο θόρυβος θεωρείται ότι ακολουθεί μια κανονική κατανομή. Τα φίλτρα Kalman χρησιμοποιούν τον θόρυβο, στους υπολογισμούς τους, με αποτέλεσμα να γίνεται χρήση των φίλτρων αυτών για τον προσδιορισμό της πορείας κινούμενων στόχων.

Γενικά, τα φίλτρα Kalman, για να εκτιμήσουν την κατάσταση ενός συστήματος, έχουν δύο φάσεις

1. Τη φάση πρόβλεψης, στην οποία πραγματοποιείται μια αρχική εκτίμηση, βασιζόμενη σε αποτελέσματα από προηγούμενες καταστάσεις και τις τρέχουσες εισόδους του συστήματος.
2. Τη φάση διόρθωσης, στην οποία η αρχική εκτιμώμενη κατάσταση ενημερώνεται και βελτιστοποιείται με την εισαγωγή των πρόσφατων δεδομένων των μετρήσεων.

Το τελικό αποτέλεσμα που εξάγεται από το φίλτρο Kalman διαμορφώνεται από το στάδιο της διόρθωσης.

Επιπλέον, η φύση των εισερχόμενων σημάτων του συστήματος είναι καθοριστικός παράγοντας για το στάδιο της πρόβλεψης. Η μεταβλητή εισόδου (c) έχει τη δυνατότητα να είναι μέρος του μοντέλου του. Συγκεκριμένα, μπορεί να ενσωματωθεί στο σύστημα χώρου-κατάστασης, όπως παρουσιάζεται από την εξίσωση που ακολουθεί

$$p_t = Ap_{t-1} + Bc_{t-1} + u_{t-1} \quad (3.34)$$

Η αντίστοιχη μέτρηση για το συγκεκριμένο μοντέλο παρουσιάζεται στην εξίσωση που ακολουθεί

$$z_t = Hp_t + u_t \quad (3.35)$$

όπου

t : χρονικός δείκτης

p : κατάσταση συστήματος τη χρονική στιγμή t

c_t : διάνυσμα εισόδου που λαμβάνεται την χρονική στιγμή t

z_t : μέτρηση τη χρονική στιγμή t

A : πίνακας μετάβασης

H : πίνακας μέτρησης

V_t : διεργασία τη χρονική στιγμή t

U_t : θόρυβος που μετρήθηκε την χρονική στιγμή t

Σε περιπτώσεις όπου η μέθοδος καλείται να διαχειριστεί μια πηγή ήχου, το διάνυσμα εισόδου c ενδέχεται να μην είναι εκ των προτέρων διαθέσιμο. Αυτό οφείλεται στο γεγονός ότι προαπαιτείτε η πληροφορία που αφορά την κίνηση της εν λόγω πηγής. Με αυτό ως βάση, και με δεδομένο ότι το σύστημα επεξεργάζεται μόνο τα δεδομένα της παρούσας κατάστασης, προκύπτει η εξίσωση που παρουσιάζεται στη συνέχεια

$$p_t = Ap_{t-1} + u_{t-1} \quad (3.36)$$

Ένα μειονέκτημα της τεχνικής αυτής συνίσταται στη στασιμότητα που δημιουργείται κατά την πρόβλεψη του συστήματος, δεδομένου ότι δεν ανιχνεύονται νέα δεδομένα, παρά μόνο ο θόρυβος της διεργασίας. Μια προτεινόμενη λύση για την αντιμετώπιση του προβλήματος αυτού είναι να πραγματοποιούνται παρατηρήσεις και μετά τη μετάβαση για τον υπολογισμό της διαφοράς της θέσης της ηχητικής πηγής ανάμεσα στα χρονικά πλαίσια. (Liang, 2008)

Ακολουθώντας, η κατάσταση του συστήματος δύναται να οριστεί μέσω του διανύσματος $p_t = [\theta_t, r_t, \dot{\theta}_t, \dot{r}_t]^T$, το οποίο περιέχει τις καρτεσιανές συντεταγμένες και την ταχύτητα της κινούμενης ηχητικής πηγής.

Επιπλέον, αξίζει να σημειωθεί ότι ο θόρυβος διαδικασίας και ο θόρυβος μέτρησης χαρακτηρίζονται από μηδενική μέση τιμή, είναι ανεξάρτητοι μεταξύ τους και ακολουθούν κανονική κατανομή με συνδιακυμάνσεις R_v και R_u , όπως παρουσιάζεται στη συνέχεια

$$v \sim N(0, R_v) \quad u \sim N(0, R_u) \quad (3.37)$$

Συνεπώς για την επίλυση του τυπικού φίλτρου Bayes χρησιμοποιούνται οι μέσες τιμές καθώς και οι συνδιακυμάνσεις. Τόσο για το στάδιο της πρόβλεψης όσο και για το στάδιο της διόρθωσης χρησιμοποιούνται τεχνικές, για τις οποίες θα πραγματοποιηθεί μια συνοπτική αναφορά στη συνέχεια

1. Η προηγούμενη εκτίμηση της κατάστασης του συστήματος $\hat{p}_t^-$, υπολογίζεται με βάση το σύνολο των πληροφοριών πριν από το χρονικό διάστημα t .
2. Η επόμενη εκτίμηση της κατάστασης του συστήματος $\hat{p}_t$, προσδιορίζεται για τη χρονική στιγμή t , ενσωματώνοντας και την τρέχουσα μέτρηση z_t .

Η διαδικασία υλοποίησης του φίλτρου Kalman περιγράφεται συνοπτικά στη συνέχεια.

- α) Στη φάση της πρόβλεψης, η προηγούμενη εκτίμηση $\hat{p}_{t-1}$ χρησιμοποιείται για τον υπολογισμό της εκ των προτέρων εκτίμησης $\hat{p}_t^-$.
- β) Στη φάση της διόρθωσης, η προηγούμενη εκτίμηση $\hat{p}_t^-$ χρησιμοποιείται για τον προσδιορισμό της επόμενης εκτίμησης $\hat{p}_t$, η οποία αποτελεί και το αποτέλεσμα του φίλτρου Kalman για τη χρονική στιγμή t .

Η εκ των προτέρων εκτίμηση της κατάστασης $\hat{p}_t^-$ παρουσιάζεται από την εξίσωση που ακολουθεί

$$\hat{p}_t^- = A\hat{p}_{t-1} \quad (3.38)$$

Όσο αφορά τη φάση της πρόβλεψης, η αναμενόμενη συνδιακύμανση του εκτιμώμενου σφάλματος τη χρονική στιγμή $t(R_{et})$ παρουσιάζεται από την παρακάτω εξίσωση

$$R_{et}^- = AR_{e_{t-1}}A^T + R_v \quad (3.39)$$

όπου

$R_{e_{t-1}}$: συνδιακύμανση του εκ των υστέρων σφάλματος τη χρονική στιγμή $t-1$

Παράλληλα, όσο αφορά τη φάση της διόρθωσης, η εκ των υστέρων εκτίμηση, η έξοδος του φίλτρου Kalman, μπορεί να υπολογιστεί με την εξίσωση που παρουσιάζεται στη συνέχεια

$$\hat{p}_t = \hat{p}_t^- + K(z_t - H\hat{p}_t^-) \quad (3.40)$$

όπου

$K(z_t - H\hat{p}_t^-)$: η διαφορά μεταξύ της πραγματικής μέτρησης z_t και της εκτιμώμενης μέτρησης $H\hat{p}_t^-$, η οποία πολλαπλασιάζεται με το κέρδος K

Το κέρδος του φίλτρου Kalman K , οι τιμές το οποίου κυμαίνονται στο εύρος $[0, H^{-1}]$, υπολογίζεται μέσω της εξίσωσης που παρουσιάζεται στη συνέχεια

$$K = \frac{R_{e_t}^- H^T}{H R_{e_t}^- H^T + R_u} \quad (3.41)$$

όπου

$R_{e_t}^-$: συνδιακύμανση του προβλεπόμενου σφάλματος

e_t^- : σφάλμα τη χρονική στιγμή t

Η συνδιακύμανση του προβλεπόμενου σφάλματος και το σφάλμα την χρονική στιγμή t υπολογίζονται μέσω των εξισώσεων που παρουσιάζονται στη συνέχεια

$$R_{e_t}^- = e_t^- e_t^{-T} \quad e_t^- = p_t - \hat{p}_t^- \quad (3.42)$$

Το κέρδος του φίλτρου Kalman είναι ανάλογο τόσο με τη συνδιακύμανση του εκ των προτέρων σφάλματος $R_{e_t}^-$, όσο και με τη συνδιακύμανση του θορύβου μέτρησης R_u . Με τον τρόπο αυτό, επιτυγχάνεται χαμηλό σφάλμα εκτίμησης $R_{e_t}^-$, γεγονός που καθιστά την εκτίμηση $H\hat{p}_t^-$ ιδιαίτερα αξιόπιστη. Αντιθέτως, στην περίπτωση όπου ο θόρυβος μέτρησης R_u είναι σε χαμηλό επίπεδο, η μέτρηση z_t χαρακτηρίζεται από μεγάλη αξιοπιστία. Συμπερασματικά, η προσεκτική παρατήρηση του αποτελέσματος αναδεικνύει ότι η μέθοδος Kalman αποτελεί μια ιδανική τεχνική σε περιβάλλοντα με έντονη παρουσία θορύβου.

Επιπλέον, στη φάση της διόρθωσης, η παράμετρος R_{e_t} μπορεί να υπολογιστεί από την εξίσωση που παρουσιάζεται στη συνέχεια

$$R_{e_t} = (I - KH)R_{e_t}^- \quad (3.43)$$

η οποία αξιοποιείται με σκοπό τον υπολογισμό της συνδιακύμανσης του σφάλματος εκτίμησης προτεραιότητας για το χρονικό διάστημα $t+1$.

Η χρήση των φίλτρων Kalman εφαρμόζεται κυρίως σε γραμμικά συστήματα προσδιορισμού θέσης ηχητικών πηγών. Αντίθετα, στην περίπτωση μη γραμμικών συστημάτων, επιστρατεύονται τα εκτεταμένα φίλτρα Kalman (EKF), τα οποία στηρίζονται στη παραδοχή ότι τα συστήματα είναι διαφορισμά και με τη χρήση των πινάκων Jacobi τα μη γραμμικά συστήματα δύναται να προσεγγιστούν ως γραμμικά.

Παράλληλα, τα φίλτρα Kalman συνεισφέρουν στον υπολογισμό της κατεύθυνσης εντοπισμού μιας ηχητικής πηγής. Στην περίπτωση, ωστόσο, που απαιτείται η χρήση τους σε συστήματα εντοπισμού πολλαπλών ηχητικών πηγών, πραγματοποιείται κατηγοριοποίηση των αρχικών χωρικών εκτιμήσεων. Η ομαδοποίηση αυτή γίνεται κυρίως με βάση την ομοιότητα των εκτιμήσεων για τις προβλέψεις και τη χρήση ξεχωριστών φίλτρων Kalman για κάθε επιμέρους ομάδα. (Ogiso, 2015)

3.5.2 Τα φίλτρα σωματιδίων

Η εφαρμογή των φίλτρων σωματιδίων εστιάζει στον προσδιορισμό της θέσης μιας ηχητικής πηγής, βασιζόμενη στο ίδιο μοντέλο κατάστασης χώρου με τα φίλτρα Kalman. Αντιθέτως, όμως με τα φίλτρα Kalman, τα φίλτρα σωματιδίων δύναται να χρησιμοποιηθούν σε μη γραμμικά συστήματα, καθώς και σε συστήματα με μη Γκαουσιανή κατανομή. Τα φίλτρα αυτά αξιοποιούν μεθόδους δειγματοληψίας, οι οποίες επιτρέπουν την εκτίμηση της συνάρτησης πυκνότητας πιθανότητας (PDF) αναφορικά με τη κατάσταση του συστήματος.

Επιπλέον, η μέθοδος αυτή υλοποιείται με χρήση του αλγορίθμου διαδοχικής σημαντικής δειγματοληψίας (SIS). Πρόκειται ουσιαστικά για μια παραλλαγή της γενικής μεθόδου Monte Carlo, η οποία δημιουργεί ένα σύνολο με σταθμισμένα δείγματα με βάση μια υπάρχουσα πυκνότητας πιθανότητας. Στη συνέχεια, τα δείγματα αυτά χρησιμοποιούνται για να εκτιμηθεί η συνάρτηση κατανομής πιθανότητας της κατάσταση του συνολικού συστήματος. Ωστόσο, επειδή η συγκεκριμένη μέθοδος εφαρμόζει τεχνικές τυχαίας δειγματοληψίας, ενδέχεται να περιορίσει την ακρίβεια των παραγόμενων αποτελεσμάτων, προκαλώντας φαινόμενα αστάθειας σε ορισμένες περιπτώσεις.

Η αναμενόμενη τιμή που προκύπτει από την εκ των υστέρων συνάρτηση κατανομής πιθανότητας είναι το αποτέλεσμα από το φίλτρο σωματιδίων. Παράλληλα, η υπολογιζόμενη διακύμανση χρησιμεύει ως δείκτης μέτρησης της αξιοπιστίας των αποτελεσμάτων αυτών.

Ο προσδιορισμός της εκ των υστέρων συνάρτηση πυκνότητας πιθανότητας επιτυγχάνεται με εφαρμογή της εξίσωσης που παρουσιάζεται στη συνέχεια

$$P(p_t | z_{1:t}) \approx \sum_{n=1}^{N_p} w_t^n \delta(p_t - p_t^n) \quad (3.44)$$

όπου

$P(p_t | z_{1:t})$: εκ των υστέρων συνάρτηση πυκνότητας πιθανότητας του συστήματος τη χρονική στιγμή t .

$\{p_{0:t}^n, w_t^n\}_{n=1}^{N_p}$: σύνολο τυχαίων σταθμισμένων δειγμάτων

w_t^n : συντελεστής στάθμισης μετά την κανονικοποίηση, για να ισχύει $\sum_{n=1}^{N_p} w_t^n = 1$

N_p : αριθμός σωματιδίων

$\delta(i)$: συνάρτηση Dirac με κεντρική τιμή το i

Η συνάρτηση Dirac υλοποιεί την κανονικοποίηση για τα βάρη w_t^n με τις τιμές από τα δείγματα $[p_t - p_t^n]$ της συνάρτησης πυκνότητας πιθανότητας. (Doucet, 2008)

Στη συνέχεια, λαμβάνει χώρα η διαδικασία προσδιορισμού των βαρών της μεθόδου, η οποία είναι απαραίτητη τόσο για την εξαγωγή της πυκνότητας πιθανότητας όσο και τη δημιουργία των ίδιων των σωματιδίων. Η τεχνική που επιλέγεται για την εκτίμηση της πυκνότητας αυτής είναι το βασικό κριτήριο για τον σχεδιασμό του φίλτρου των σωματιδίων. Εντούτοις, ένα βασικό μειονέκτημα της μεθόδου είναι ότι όσο επαναλαμβάνεται η διαδικασία, η πλειονότητα των σωματιδίων μπορεί να αποκτήσει αμελητέο βάρος. Με αυτό τον τρόπο δημιουργείται το φαινόμενο του εκφυλισμού, το οποίο συνεπάγεται τη σταδιακή διεύρυνση της διακύμανσης των βαρών συναρτήσει του χρόνου.

Μια τεχνική για τον περιορισμό του μειονεκτήματος αυτού, επιτάσσει η χρησιμοποιούμενη πυκνότητα να προέρχεται από την πρότερη κατανομή της συνάρτησης πιθανότητας. Με τον τρόπο

αυτό, τα βάρη δύνανται να υπολογιστούν με εφαρμογή του θεωρήματος του Bayes, όπως παρουσιάζεται από την εξίσωση που ακολουθεί

$$w_t^n = w_{t-1}^n P(z_t | p_t^n) \quad (3.45)$$

Μια επιπλέον λύση για το μειονέκτημα της μεθόδου αυτής, και πιο συγκεκριμένα για το πρόβλημα του εκφυλισμού, είναι η εφαρμογή της τεχνικής της επαναδειγματοληψίας, η οποία τεχνική αποτελεί τη διαδικασία που ανανεώνονται οι τιμές των βαρών. Αυτό πραγματοποιείται με τη δημιουργία ενός νέου συνόλου σωματιδίων από τις υπολογισμένες σταθμισμένες μονάδες, δηλαδή από τα σωματίδια που έχουν απομείνει.

Η παραγωγή των σωματιδίων αυτών βασίζεται στη δειγματοληψία, κατά την οποία τα νέα σωματίδια αποτελούν αναπαραγωγή των προηγούμενων σωματιδίων. Ο προσδιορισμός του πλήθους των αντιγραφών ενός παλαιότερου σωματιδίου (N_n) πραγματοποιείται με χρήση ακέραιων τιμών. Η διαδικασία αυτή πραγματοποιείται μέσω ενός καθορισμένου σύνολο αριθμών N_p , το οποίο ορίζεται ως σύνολο U . Για κάθε αριθμό ισχύει ο τύπος $U_i = U_1 + \frac{i-1}{N_p}$, όπου το U_1 αποτελεί μια τυχαία επιλεγμένη τιμή στο διάστημα $(0, N_p)$. Συνοψίζοντας, η συστηματική επαναδειγματοληψία καθορίζει την τιμή N με βάση το σύνολο των στοιχείων που ανήκουν στο U , υπό την προϋπόθεση ότι η τιμή βρίσκεται εντός του διαστήματος $[\sum_{k=1}^{n-1} w_t^k, \sum_{k=1}^n w_t^k]$.

Στη συνέχεια παρουσιάζονται συνοπτικά τα στάδια εκτέλεσης ενός φίλτρου σωματιδίων για τη χρονική στιγμή $t > 1$.

- α) Αρχικά, προσδιορίζεται η τρέχουσα πυκνότητα δεδομένων των σωματιδίων για τη χρονική στιγμή $t-1$.
- β) Στη συνέχεια, πραγματοποιείται τυχαία επιλογή των σωματιδίων N_p με βάση την τρέχουσα πυκνότητα.
- γ) Έπειτα, υλοποιείται ο υπολογισμός των βαρών των σωματιδίων w_t^n .
- δ) Ακολούθως, εκτελείται η κανονικοποίηση των βαρών w_t^n , ώστε $\sum_{n=1}^{N_p} w_t^n = 1$.
- ε) Μετά, λαμβάνει χώρα η αξιολόγηση της εκφύλισης των βαρών.
- στ) Ύστερα, ακολουθεί η επαναδειγματοληψία, εφόσον ο βαθμός εκφύλισης έχει μειωθεί κάτω από ένα προκαθορισμένο όριο.
- ζ) Εν συνεχεία, εκτιμάται η συνάρτηση πυκνότητας πιθανότητας για την κατάσταση του συστήματος $P(p_t | z_{1:t})$.
- η) Τέλος, ολοκληρώνεται ο υπολογισμός της εκτίμησης και της διακύμανσης της $P(p_t | z_{1:t})$, καθώς και η παρουσίαση των αποτελεσμάτων για την ιδανική προτεινόμενη τιμή.

Τα φίλτρα σωματιδίων, παρόλο που δεν αποτελούν πάντοτε την πλέον αξιόπιστη μέθοδο, χρησιμοποιούνται συχνά για τον εντοπισμό των ηχητικών πηγών. Ειδικότερα, εφαρμόζονται για εντοπισμό πολλαπλών ηχητικών πηγών ή σε περιπτώσεις όπου οι τροχιές τους διασταυρώνονται.

Τα φίλτρα σωματιδίων είναι στην πραγματικότητα μια γενίκευση των φίλτρων Bayes. Συγκεκριμένα, τα στάδια για τον υπολογισμό μπορούν να εκτελεστούν με φίλτρα Bayes για κάθε σωματίδιο ξεχωριστά, επιτρέποντας έτσι την ενημέρωση των βαρών του. Το βασικό μειονέκτημα των φίλτρων σωματιδίων είναι ότι εφαρμόζουν ταυτόχρονα πολλά φίλτρα Bayes για κάθε σωματίδιο. Καθώς τα σωματίδια αναπαράγονται κατά το στάδιο της επαναδειγματοληψίας, η συγκεκριμένη διαδικασία απαιτεί αυξημένους υπολογιστικούς πόρους. (Asoh, 2013)

3.5.3 Διάφορες άλλες προσεγγίσεις

Επιπλέον, υπάρχουν αρκετές εναλλακτικές προσεγγίσεις για τον εντοπισμό της ηχητικής πηγής. Στη συνέχεια, παρουσιάζεται μια συνοπτική ανασκόπηση σε ορισμένες από αυτές.

Λογαριθμική πιθανοφάνεια.

Με τη μέθοδο αυτή, δημιουργείται ένα πλέγμα που αναπαριστά τον χώρο κίνησης του ρομπότ. Κύριος στόχος είναι κάθε κελί του πλέγματος να παρέχει θεωρητικά μια λογαριθμική πιθανοφάνεια ύπαρξης μιας ηχητικής πηγής. Βασική προϋπόθεση για την εφαρμογή της μεθόδου αυτής αποτελεί ότι το ρομπότ πρέπει να ακούει και να κινείται ταυτόχρονα, δεδομένου ότι το πλέγμα των ακουστικών στοιχείων επικαιροποιείται και συγκλίνει με τη χρήση των ηχητικών δεδομένων, καθώς και με τον προσδιορισμό της θέσης του ίδιου του πλέγματος.

Ένα μειονέκτημα της συγκεκριμένης προσέγγισης αποτελεί η περιορισμένη αποτελεσματικότητα της σε εφαρμογές που βασίζονται σε συστήματα με πολλαπλές ηχητικές πηγές. (Martinson, 2007)

Προσδοκία – Μεγιστοποίηση.

Με την μέθοδο αυτή πραγματοποιείται αρχικά το στάδιο της προσδοκίας, όπου πραγματοποιείται η εκτίμηση της αναμενόμενης τιμής της συνάρτησης πιθανοφάνειας με βάση τα διαθέσιμα δεδομένα. Ακολούθως, κατά το στάδιο της μεγιστοποίησης προσδιορίζονται οι παράμετροι εκείνοι που μεγιστοποιούν τη συγκεκριμένη αναμενόμενη τιμή. Μια βασική εφαρμογή της μεθόδου είναι ο εντοπισμός ανθρώπων. (Kim, 2009)

Επαναλαμβανόμενα νευρωνικά δίκτυα (RNN).

Η μέθοδος αυτή αξιοποιεί μια κατηγορία τεχνητών νευρωνικών δικτύων, η οποία χαρακτηρίζεται από δυναμικές εσωτερικές καταστάσεις και παράλληλα αξιοποιεί δεδομένα που έχουν αποθηκευτεί κατά τη διαδικασία αναγνώρισης. Στο πλαίσιο αυτό, οι πληροφορίες για την εκτίμηση της κατεύθυνσης άφιξης του ακουστικού σήματος εισάγονται σε ένα επαναλαμβανόμενο νευρωνικό δίκτυο, με σκοπό τον προσδιορισμό της πιθανής θέσης της πηγής. Η εσωτερική κατάσταση του δικτύου αυτού τροφοδοτείται συνεχώς τόσο με τα τρέχουσα όσο και με τα προηγούμενα χωρικά δεδομένα της ηχητικής πηγής. (Murray, 2009)

Τυχαία Συναίνεση Δειγμάτων (RANSAC).

Η μέθοδος αυτή βασίζεται στη στοχαστική και επαναλαμβανόμενη δειγματοληψία των παρατηρούμενων δεδομένων. Για την εφαρμογή αυτής της μεθόδου, είναι απαραίτητη η ύπαρξη τόσο κανονικών τιμών όσο και ακραίων τιμών στο σύνολο των δεδομένων. Στη συνέχεια, μέσω της επαναληπτικής διαδικασίας και με την παραδοχή ότι οι ακραίες τιμές αποκλείονται συστηματικά, ενώ οι κανονικές συμπεριλαμβάνονται στο μοντέλο, καθίσταται δυνατός ο εντοπισμός της ηχητικής πηγής. Το βασικό πλεονέκτημα της μεθόδου αυτής είναι η ικανότητα της για τον ταυτόχρονο εντοπισμό πολλαπλών ηχητικών πηγών. (Sasaki, 2006)

3.6 Σύνοψη 3^{ου} κεφαλαίου

Στο συγκεκριμένο τμήμα της εργασίας εξετάστηκαν οι τεχνικές προσδιορισμού θέσης και υπολογισμού της απόστασης της ηχητικής πηγής από τον δέκτη, στο πλαίσιο ρομποτικών εφαρμογών. Παράλληλα, αναλύθηκαν οι διάφορες μέθοδοι ανάλογα με την κατηγοριοποίησή τους. Οι κατηγορίες αυτές αφορούν μια μεμονωμένη ή πολλαπλές ηχητικές πηγές, οι οποίες μπορεί να είναι είτε σταθερές είτε κινούμενες.

Οι προαναφερθείσες μέθοδοι εφαρμόζονται βάσει των κριτηρίων που αναλύθηκαν. Ειδικότερα, ορισμένες προσεγγίσεις αξιοποιούνται σε περιβάλλοντα χωρίς θόρυβο, ενώ άλλες εφαρμογές είναι κατάλληλες για συνθήκες με πολλαπλές ηχητικές πηγές ή με ανθρώπινη παρουσία. Κοινό στοιχείο όλων των μεθόδων αποτελεί η εξάρτηση τους από συγκεκριμένες παραμέτρους, οι οποίες σχετίζονται με την ίδια τη μέθοδο, το κόστος υλοποίησης, το περιβάλλον εφαρμογής και τα επιμέρους τεχνικά χαρακτηριστικά. Ένα από τα βασικότερα κριτήρια είναι η ταχύτητα, τόσο κατά τη λήψη όσο και κατά την επεξεργασία των ηχητικών δεδομένων. Η συγκεκριμένη παράμετρος είναι ιδιαιτέρως κρίσιμη για ρομποτικά συστήματα, καθώς καλούνται να ανταποκριθούν σε ταχύτητα μεταβαλλόμενες συνθήκες, αλλά και σε συχνά τιθέμενα ζητήματα ασφάλειας.

Επιπλέον, αξίζει να σημειωθεί ότι ο εντοπισμός της ηχητικής πηγής αποτελεί κρίσιμη παράμετρο, παρόλο που δεν είναι η πλέον διαδεδομένη στις ρομποτικές εφαρμογές. Αν και στα ρομποτικά συστήματα χρησιμοποιούνται κυρίως οπτικοί αισθητήρες για την πλοήγηση, σε πολλές περιπτώσεις κρίνεται αναγκαία η χρήση ακουστικών αισθητήρων. Αυτό οφείλεται στο ότι η απότομη εναλλαγή της φωτεινότητας ή η απουσία φωτός καθιστούν αναποτελεσματική τη χρήση οπτικών μεθόδων, καθιστώντας απαραίτητη την υιοθέτηση ακουστικών μεθόδων.

Στη συνέχεια, για την πληρέστερη κατανόηση της σημασίας που έχει η ταχεία λήψη και ανάλυση των ακουστικών δεδομένων, παρατίθενται δύο παραδείγματα. Στην περίπτωση της αυτόνομης πλοήγησης οχήματος, ο επεξεργαστής οφείλει να λαμβάνει και να επεξεργάζεται ταχύτατα τα ηχητικά σήματα, ακόμη και σε περιβάλλοντα με υψηλά επίπεδα θορύβου. Η γρήγορη επεξεργασία επιτρέπει τον ακριβή αναγνώριση της κατεύθυνσης και της απόστασης της πηγής, αποτρέποντας τυχόν παρερμηνείες, όπως για παράδειγμα την εσφαλμένη λήψη οδηγιών πλοήγησης από τους επιβαίνοντες των πίσω θέσεων. Μια άλλη εφαρμογή αφορά τα ρομποτικά συστήματα διάσωσης που εντοπίζουν εγκλωβισμένους, όπου η συνολική ταχύτητα λήψης και επεξεργασίας των σημάτων είναι ζωτικής σημασίας.

Επιπλέον, καθοριστικό παράγοντα στον εντοπισμό αποτελεί η ίδια η ηχητική πηγή, και συγκεκριμένα τα χαρακτηριστικά της, όπως το είδος και το μέγεθος της. Παράλληλα, ο αριθμός των ηχητικών πηγών, και η μεταξύ τους απόσταση επηρεάζουν άμεσα την ακρίβεια της εκτίμησης της θέσης τους.

Από την άλλη πλευρά, εξίσου σημαντικό ρόλο διαδραματίζει το σύστημα λήψης των ηχητικών σημάτων. Οι βασικοί παράμετροι που επηρεάζουν την απόδοση του, περιλαμβάνουν τη φύση του συστήματος, την ποιότητα και την κατευθυντικότητα του ακουστικού αισθητήρα, καθώς και την γεωμετρική διάταξη των μικροφώνων. Η διάταξη των μικροφώνων επηρεάζει την ακρίβεια προσδιορισμού της κατεύθυνσης άφιξης του ηχητικού σήματος. Για παράδειγμα, οι γραμμικές διατάξεις των μικροφώνων περιορίζονται στον εντοπισμό της θέσης της ηχητικής πηγής σε οριζόντιο επίπεδο, αντίθετα όμως οι κυκλικές ή οι τρισδιάστατες διατάξεις δύναται να πραγματοποιήσουν σημαντικά ακριβέστερη εκτίμηση της θέσης, τόσο σε αζιμουθιακό όσο και σε κατακόρυφο επίπεδο.

Τέλος, αξίζει να σημειωθεί ότι η ενσωμάτωση μεθοδολογιών μηχανικής μάθησης και τεχνητής νοημοσύνης έχει αναβάθμιση αισθητά τα ρομποτικά συστήματα εντοπισμού. Οι σύγχρονες προσεγγίσεις αξιοποιώντας τεχνητά νευρωνικά δίκτυα και μοντέλα βαθιάς μάθησης, προσφέρουν βέλτιστα αποτελέσματα όχι μόνο για την εκτίμηση της θέσης της ηχητικής πηγής, αλλά και στην ταυτοποίηση ή τον αποτελεσματικό διαχωρισμό των ηχητικών σημάτων.

Κεφάλαιο 4

4. Ανάπτυξη ρομποτικής εφαρμογής

4.1 Εισαγωγικό μέρος - Σκοπός

Στο συγκεκριμένο τμήμα της εργασίας παρουσιάζεται η μελέτη και η κατασκευή ενός ρομποτικού οχήματος, το οποίο κινείται σύμφωνα με τη διαφορά ενέργειας του ηχητικού σήματος που λαμβάνεται από τα δύο μικρόφωνα που έχουν τοποθετηθεί πάνω σε αυτό. Το ρομποτικό όχημα έχει σχήμα που προσομοιώνει προσεγγιστικά το ανθρώπινο κεφάλι, ενώ η απόσταση μεταξύ των μικροφώνων είναι 23 εκατοστά σε ευθεία γραμμή.

Ο βασικός σκοπός της κατασκευής είναι ο εντοπισμός της ηχητικής πηγής και, ακολούθως, η κίνηση του ρομποτικού οχήματος προς την κατεύθυνση της. Η αρχική ιδέα, η οποία τελικά δεν υλοποιήθηκε, προέβλεπε τον εντοπισμό της ηχητικής πηγής τόσο στο αζιμουθιακό επίπεδο όσο και στον κατακόρυφο άξονα. Επιπλέον, έχει πραγματοποιηθεί μελέτη σχετικά με την επίδραση των ηχητικών φαινομένων όπως οι ανακλάσεις και τα περιβαλλοντικά φαινόμενα όπως η θερμοκρασία και η υγρασία, τα οποία θα επηρέαζαν αρχικά τον εντοπισμό της ηχητικής πηγής και κατ' επέκταση τη συμπεριφορά του ρομποτικού οχήματος. Ωστόσο, η ανωτέρω προσέγγιση δεν εφαρμόστηκε, με αποτέλεσμα να υλοποιηθεί μόνο η βασική εκδοχή του ρομποτικού οχήματος για εντοπισμό της ηχητικής πηγής.

Το ρομποτικό όχημα βρίσκεται αρχικά σε κατάσταση ηρεμίας, περιμένοντας την ανίχνευση κάποιου ηχητικού σήματος. Ωστόσο, με τον εντοπισμό του σήματος, κατευθύνεται προς τη θέση της ηχητικής πηγής εκπομπής. Μόλις επιτευχθεί η προσέγγιση της ηχητικής πηγής στο αζιμουθιακό επίπεδο, το όχημα σταματά, παράγει έναν ηχητικό τόνο και παράλληλα ανάβει η αντίστοιχη φωτεινή ένδειξη. Στη συνέχεια, επανέρχεται στην αρχική κατάσταση ηρεμίας, περιμένοντας επιπλέον την ανίχνευση κάποιου ηχητικού σήματος, επαναλαμβάνοντας την διαδικασία.

4.2 Ανάλυση υλικού του ρομποτικού οχήματος

Στη συνέχεια, παρουσιάζεται μια συνοπτική περιγραφή του ρομποτικού οχήματος και ακολούθως πραγματοποιείται αναλυτική παρουσίαση των επιμέρους περιφερειακών διατάξεων και των λοιπών στοιχείων που το απαρτίζουν.

Το ρομποτικό όχημα αποτελείται από μια επίπεδη πλαστική βάση και ένα αντίστοιχο κέλυφος, τα οποία διαμορφώνουν τελικά ένα ημισφαιρικό σχήμα. Ο βασικός σκελετός του ρομποτικού οχήματος προέρχεται από μια έτοιμη κατασκευή της εταιρείας De Agostini. Αντίθετα, το υπόλοιπο τμήμα αποτελεί προϊόν προσωπικής μελέτης και σχεδιασμού, χωρίς να σχετίζεται με την αρχική κατασκευή που πραγματοποίησε η εταιρεία De Agostini.



Εικόνα 4.1 Απεικόνιση του ρομποτικού οχήματος.

Τροφοδοσία του ρομποτικού οχήματος

Για τη λειτουργία του, το ρομποτικό όχημα απαιτεί δύο ανεξάρτητες πηγές τροφοδοσίας. Η πρώτη αφορά μια μπαταρία των 9V, η οποία χρησιμοποιείται για την τροφοδοσία της πλατφόρμας Arduino Mega. Επιπλέον, απαιτείται μια τροφοδοσία της τάξης των 6V DC, η οποία παρέχεται από μια συστοιχία μπαταριών για την τροφοδοσία των υπόλοιπων ηλεκτρονικών κυκλωμάτων του ρομποτικού οχήματος.



Εικόνα 4.2 Απεικόνιση των πηγών τροφοδοσίας του ρομποτικού οχήματος. Στο κέντρο της εικόνας απεικονίζεται η συστοιχία μπαταριών που παρέχει τάση 6V DC, ενώ στο επάνω αριστερό τμήμα διακρίνεται η μπαταρία που παρέχει τάση 9V DC.

Η πηγή τροφοδοσίας των 9V DC συνδέεται απευθείας με την πλατφόρμα Arduino, εξυπηρετώντας αποκλειστικά την λειτουργία της.

Αντιστοίχως, η τροφοδοσία των 6V DC οδηγείται σε έναν διανομέα ισχύος με ενσωματωμένο διακόπτη ON-OFF, και στη συνέχεια κατανέμεται ανεξάρτητα σε κάθε κινητήρα των τροχών. Παράλληλα, η ίδια γραμμή τροφοδοσίας συνδέεται με ένα κύκλωμα υποβιβασμού τάσης, το οποίο χρησιμοποιείται αποκλειστικά για την επιτήρηση και ένδειξη της τάσης των μπαταριών. Η ένδειξη αυτή αξιοποιείται κατά την διάρκεια των δοκιμών για τη συνεχή ενημέρωση σχετικά με την κατάσταση των μπαταριών, ώστε να αποφευχθεί τυχόν πτώση της κάτω από ένα κρίσιμο όριο. Μια τέτοια πτώση θα προκαλούσε διακοπή λειτουργίας στα υπόλοιπα ηλεκτρονικά κυκλώματα και κατ' επέκταση, στο ίδιο το ρομποτικό όχημα.

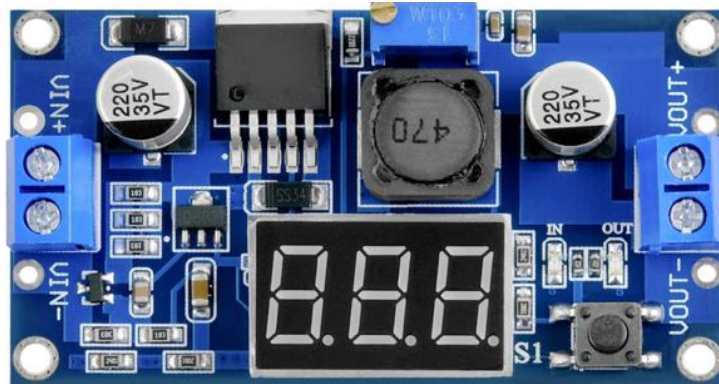
Εναλλακτικά, το εν λόγω ηλεκτρονικό κύκλωμα θα μπορούσε να παραλειφθεί, ώστε η μέτρηση της τάση των μπαταριών να πραγματοποιείται από το Arduino και η αντίστοιχη ένδειξη να προβάλλεται στη βασική οθόνη πληροφοριών.



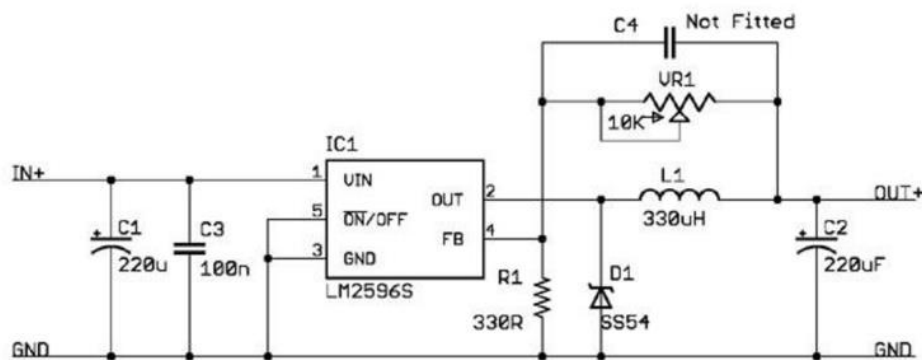
Εικόνα 4.3 Απεικόνιση του υποβιβαστή τάσης με ενσωματωμένη ένδειξη τάσης μπαταριών

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Το συγκεκριμένο ηλεκτρονικό στοιχείο αποτελεί τη μονάδα υποβιβασμού τάσης AZ-Delivery LM2596S LM2596 step-down DC-DC Buck-Converter Modul with 3 Digits Voltmeter Display.



Εικόνα 4.4 Απεικόνιση του υποβιβαστή τάσης από την επίσημη ιστοσελίδα της AZ-Delivery. (<https://www.az-delivery.de>)

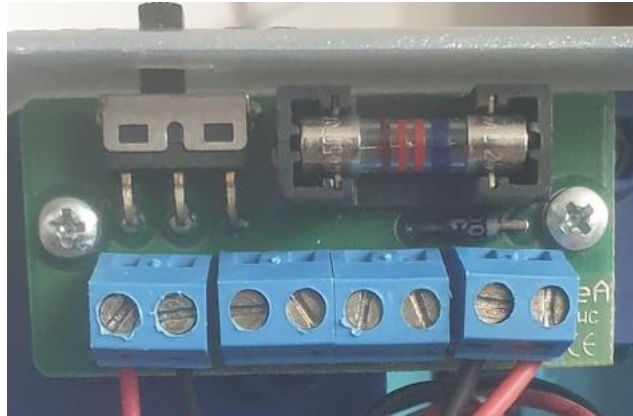


Εικόνα 4.5 Απεικόνιση του σχηματικού διαγράμματος του υποβιβαστή τάσης. (www.handsontec.com)

Η ρύθμιση της τάσης εξόδου επιτυγχάνεται μέσω του LM2596S, το οποίο κύκλωμα συγκρίνει την τάση του ακροδέκτη 4 με μια εσωτερική τάση αναφοράς 1,23 V. Η τάση εξόδου προσδιορίζεται βάσει της εξίσωσης που παρουσιάζεται στην συνέχεια

$$V_{out} = V_{ref} \left(1 + \frac{R_2}{R_1} \right) \quad (4.1)$$

Τεχνικά εγχειρίδια : <https://www.ti.com/lit/ds/symlink/lm2596.pdf#page=37.16>



Εικόνα 4.6 Απεικόνιση του κυκλώματος διανομής ισχύος, το οποίο περιλαμβάνει τον διακόπτη για την τροφοδοσία του κυκλώματος.

Το κύκλωμα διανομής ισχύος περιλαμβάνει έναν ακροδέκτη σύνδεσης για την είσοδο τροφοδοσίας από τις μπαταρίες, καθώς και τρεις ακροδέκτες εξόδου. Επιπλέον, ενσωματώνει έναν διακόπτη δύο θέσεων ON-OFF, μια δίοδο προστασίας έναντι αναστροφής τάσης και μια ασφάλεια. Η δίοδος και η ασφάλεια αποτελούν τη διάταξη προστασίας του κυκλώματος έναντι πιθανής ανάστροφης πολικότητας κατά τη σύνδεση των καλωδίων τροφοδοσίας των μπαταριών. Παράλληλα, η δίοδος παρέχει προστασία από αύξηση του ρεύματος που προκαλείται λόγω ανάστροφου επαγωγικού ρεύματος από τους κινητήρες ή κατά τη διακοπή λειτουργίας του συστήματος, φαινόμενα τα οποία ενδέχεται να προκαλέσουν δυσλειτουργίες στα υπόλοιπα ηλεκτρονικά κυκλώματα.

Σύστημα κίνησης του ρομποτικού οχήματος

Για την κίνηση του ρομποτικού οχήματος χρησιμοποιούνται τρεις αυτόνομοι τροχοί. Ο οπίσθιος τροχός είναι ελεύθερος και δεν δέχεται κινητήρια δύναμη, επιτρέποντας στο ρομποτικό όχημα να στρίβει προς όλες τις κατευθύνσεις. Επιπλέον, το ρομποτικό όχημα διαθέτει δύο εμπρόσθιους τροχούς, οι οποίοι κινούνται μέσω κινητήρων συνεχούς ρεύματος με συλλέκτη (DC Motors). Οι τροχοί αυτοί εξασφαλίζουν τη γραμμική κίνηση του ρομποτικού οχήματος, προσφέροντας τη δυνατότητα κίνησης προς τα εμπρός ή προς τα πίσω.



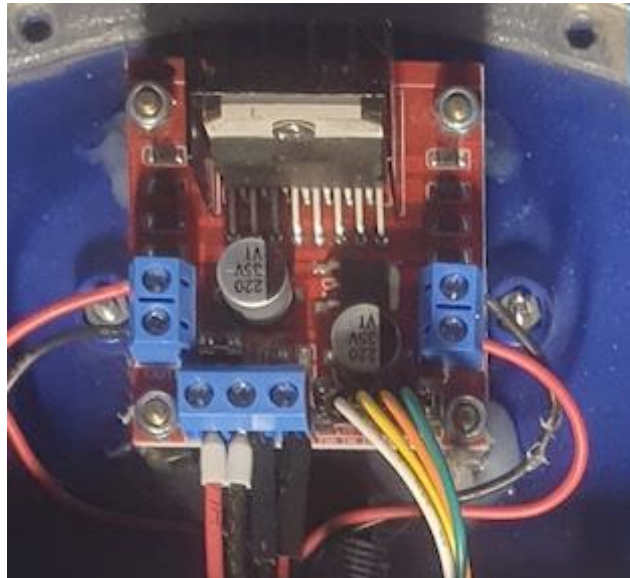
Εικόνα 4.7 Απεικόνιση του οπίσθιου ελεύθερου τροχού του ρομποτικού οχήματος



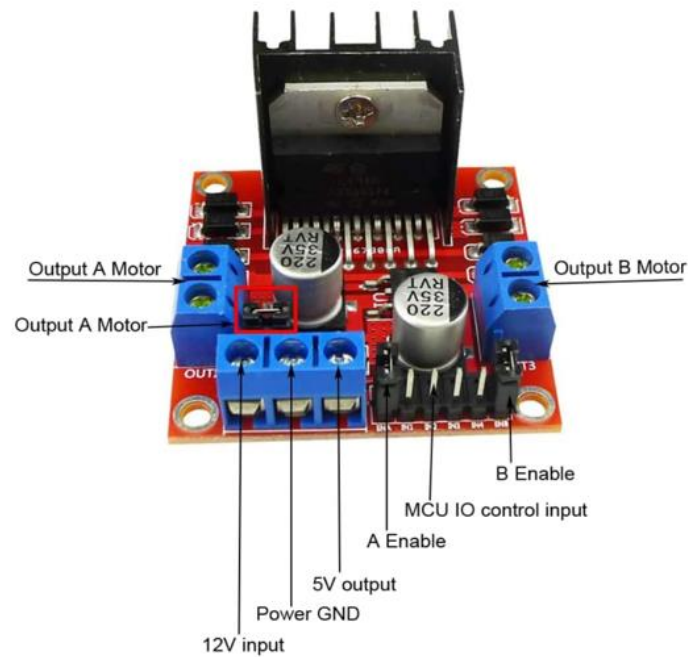
Εικόνα 4.8 Απεικόνιση των δύο εμπρόσθιων τροχών, των κινητήρων και του συστήματος οδοντωτών τροχών (γρاناζιών).

Έλεγχος των δύο κινητήρων

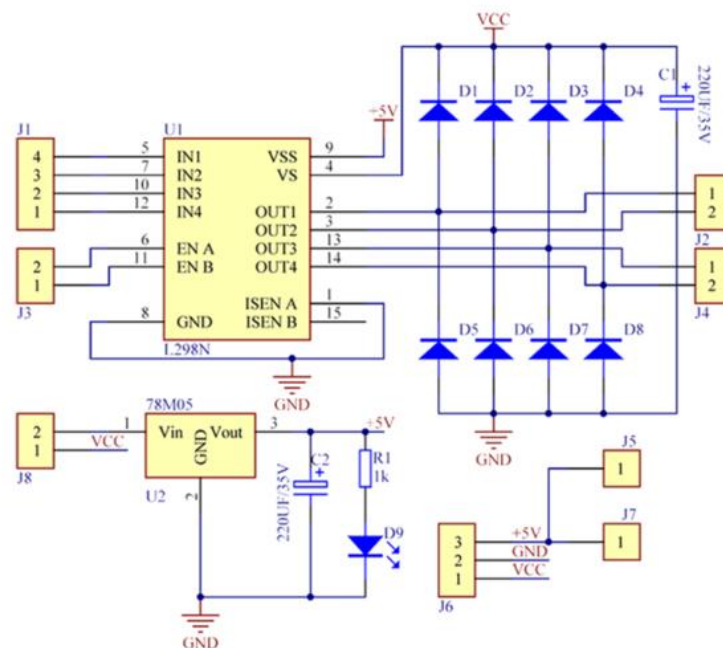
Για την οδήγηση των δύο κινητήρων χρησιμοποιήθηκε η μονάδα ελέγχου L289N Motor Drive Controller Board Modul Dual H Bridge DC Stepper Modul της εταιρείας Aihasd. Η συγκεκριμένη μονάδα επιλέχτηκε με σκοπό τον έλεγχο τόσο της κατεύθυνσης περιστροφής όσο και της ταχύτητας των δύο κινητήρων.



Εικόνα 4.9 Απεικόνιση της ηλεκτρονικής πλακέτας οδήγησης των δύο κινητήρων.



Εικόνα 4.10 Απεικόνιση της πλακέτας οδήγησης των κινητήρων από τα τεχνικά εγχειρίδια του κατασκευαστή. (www.amazon.de).



Εικόνα 4.11 Απεικόνιση του σχηματικού διαγράμματος της ηλεκτρονικής πλακέτας οδήγησης των κινητήρων από τα τεχνικά εγχειρίδια του κατασκευαστή. (www.handsontec.com)

Τεχνικά εγχειρίδια : <https://www.handsontec.com/dataspecs/L298N%20Motor%20Driver.pdf>

Αισθητήρας υπερήχων

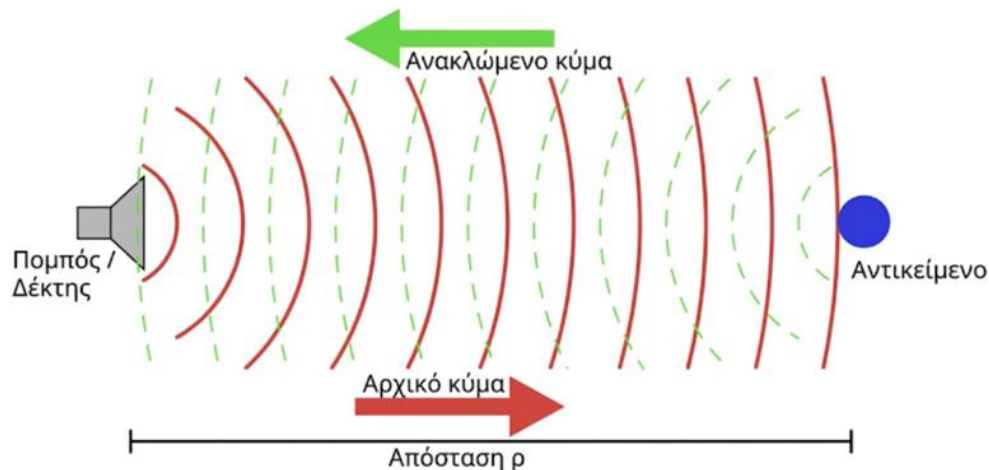
Ο υπερήχος ορίζεται ως ένα μηχανικό ακουστικό κύμα, του οποίου οι συχνότητες είναι υψηλότερες από το ανώτατο ακουστικό όριο της ανθρώπινης ακοής. Οι φυσικές του ιδιότητες δεν διαφέρουν από εκείνες του ακουστού ήχου, με μόνη εξαίρεση το γεγονός ότι δεν είναι ακουστός από τον άνθρωπο. Συγκεκριμένα, η συχνότητα του είναι πάνω από τα 20 KHz, όριο το οποίο αποτελεί δυνητικά το ανώτερο όριο της ανθρώπινης ακουστικής ικανότητας. Οι συσκευές υπερήχων λειτουργούν σε ένα ευρύ φάσμα συχνοτήτων, το οποίο περιλαμβάνει τις συχνότητες από 20 kHz έως και τις συχνότητες των GHz.

Οι υπερήχοι αξιοποιούνται σε πληθώρα εφαρμογών της ανθρώπινης δραστηριότητας, με βασικότερες από αυτές την ανίχνευση αντικειμένων και τη μέτρηση αποστάσεων.



Εικόνα 4.12 Απεικόνιση του φάσματος των ηχητικών κυμάτων, με ιδιαίτερη έμφαση στην περιοχή των συχνοτήτων των υπερήχων. (www.ardumotive.com)

Οι αισθητήρες υπερήχων βασίζονται στην ίδια αρχή λειτουργίας με τα συστήματα ραντάρ και με τα σόναρ. Συγκεκριμένα, εκτιμούν την απόσταση ενός στόχου, λαμβάνοντας υπόψη την ανάκλαση του εκπεμπόμενου ηχητικού σήματος επάνω σε αυτόν, όπως παρουσιάζεται από την παρακάτω εικόνα.



Εικόνα 4.13 Απεικόνιση της αρχής λειτουργίας του αισθητήρα υπερήχων. (www.ardumotive.com)

Οι αισθητήρες υπερήχων έχουν τη δυνατότητα παραγωγής κυμάτων υψηλής συχνότητας και κάνοντας χρήση του ανακλώμενου σήματος, μπορούν να προσδιορίσουν την απόσταση ή ακόμη και την ταχύτητα του στόχου. Για τον καθορισμό τόσο της απόστασης όσο και της ταχύτητας, χρησιμοποιούν το χρονικό διάστημα που μεσολαβεί προκειμένου το σήμα να διανύσει την απόσταση από τον αισθητήρα υπερήχων προς το στόχο και να επιστρέψει. (www.ardumotive.com)

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Στο ρομποτικό όχημα ενσωματώνεται ο αισθητήρας υπερήχων HC-SR04, με σκοπό την αποφυγή πρόσκρουσης με διάφορα εμπόδια ή την ηχητική πηγή. Το ρομποτικό όχημα έχει προγραμματιστεί ώστε να ακινητοποιείται σε απόσταση δέκα εκατοστών από το εκάστοτε εμπόδιο, ενεργοποιώντας παράλληλα μια κόκκινη λυχνία (LED). Στην περίπτωση, όμως, που το εμπόδιο αποτελεί την πηγή του ήχου, το όχημα σταματά στην ίδια απόσταση από αυτήν, ενεργοποιεί παράλληλα την κόκκινη λυχνία (LED) και ταυτόχρονα εκπέμπει έναν ακουστό τόνο μέσω του βομβητή (Buzzer).



Εικόνα 4.14 Απεικόνιση του αισθητήρα υπερήχων HC-SR04. (www.amazon.de)

Ο αισθητήρας υπερήχων HC-SR04 δύναται να μετρήσει αποστάσεις από δύο εκατοστά έως και τετρακόσια εκατοστά. Ο αισθητήρας λειτουργεί εκπέμποντας έναν παλμό υπερήχων στα 40 kHz και υπολογίζοντας το χρόνο που για την επιστροφή της ηχώ (echo). Για τον υπολογισμό της απόστασης, χρησιμοποιείται η εξίσωση που παρουσιάζεται στη συνέχεια

$$D = \frac{t * c}{2} \quad (4.2)$$

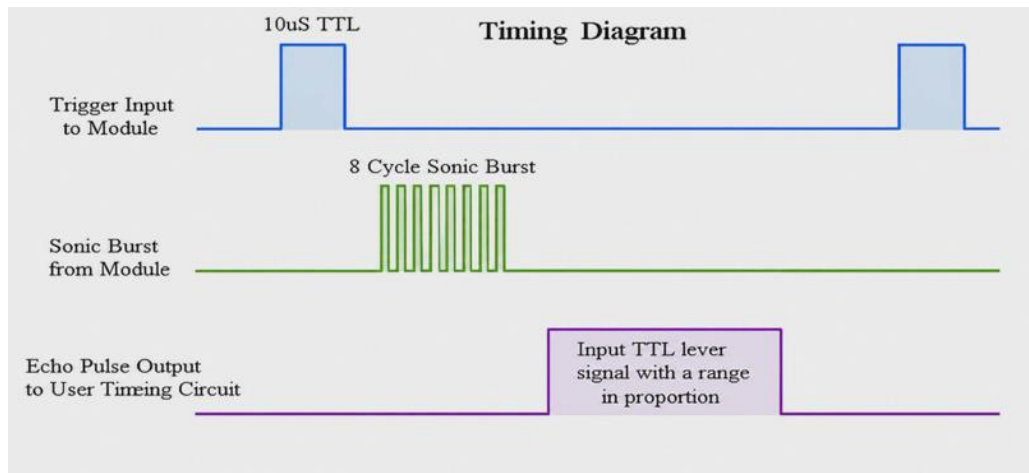
όπου

D: η απόσταση (cm)

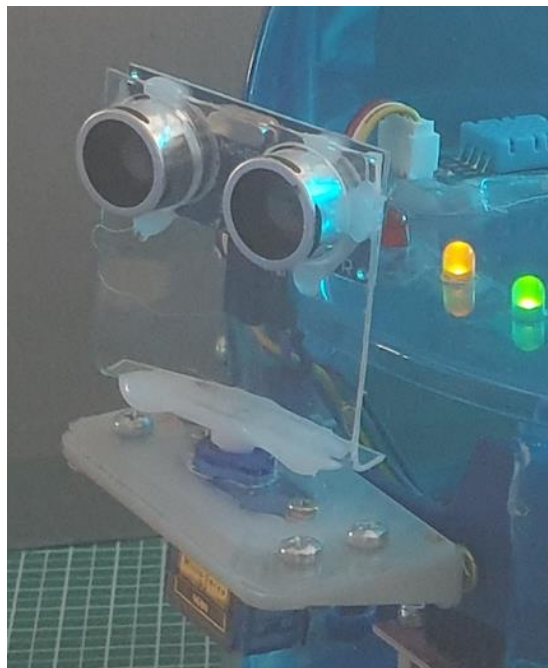
t: ο χρόνος (sec)

c: η ταχύτητα του ήχου

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.



Εικόνα 4.15 Απεικόνιση του διαγράμματος χρονισμού του αισθητήρα υπερήχων HC-SR04. (cdn.sparkfun.com)



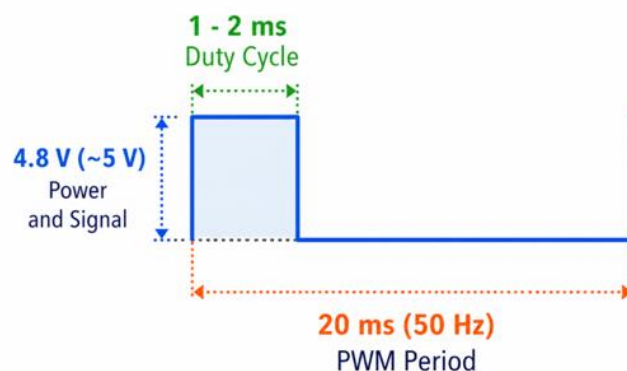
Εικόνα 4.16 Απεικόνιση του αισθητήρα υπερήχων σε συνδυασμό με τον σερβοκινητήρα.

Τεχνικά εγχειρίδια : <https://cdn.sparkfun.com/datasheets/Sensors/Proximity/HCSR04.pdf>

Ο αισθητήρας υπέρηχων είναι τοποθετημένος πάνω σε έναν σερβοκινητήρα, επιτρέποντας τον εντοπισμό εμποδίων υπό γωνία. Για την υλοποίηση αυτή, χρησιμοποιείται ο σερβοκινητήρας SG90 9g Micro Servo.



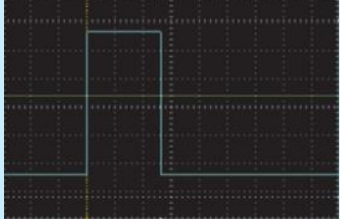
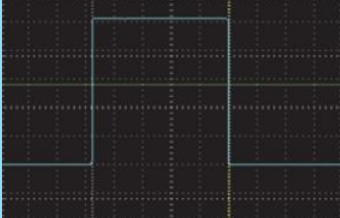
Εικόνα 4.17 Απεικόνιση σερβοκινητήρα SG90 9g Micro Servo. (www.amazon.de)



Εικόνα 4.18 Απεικόνιση του σήματος ελέγχου PWM του σερβοκινητήρα. (handsontec.com)

Η λειτουργία του σερβοκινητήρα βασίζεται στην οδήγηση του από ψηφιακούς παλμούς διαμόρφωσης εύρους παλμού (PWM). Η συχνότητα του συγκεκριμένου σήματος καθορίζεται στα 50 Hz, με συνέπεια η περίοδος του παλμού να είναι ίση με 20 ms. Η γωνία στροφής του άξονα του σερβοκινητήρα είναι ανάλογη με το χρονικό διάστημα κατά το οποίο το σήμα βρίσκεται σε υψηλή κατάσταση (HIGH).

α/α	Γωνιά (μοίρες)	Θέση	Διάρκεια υψηλής κατάστασης (ms)	Οπτική παρουσίαση της διάρκειας
1	-90°	Τέρμα αριστερά	0,5	

2	0°	Κέντρο	1,5	
3	90°	Τέρμα δεξιά	2,5	

Πίνακας 4.1 Απεικόνιση των καταστάσεων λειτουργίας και της αντίστοιχης θέσης του σερβοκινητήρα.

Τεχνικά εγχειρίδια : https://handsontec.com/dataspecs/motor_fan/SG90-Servo.pdf

Οθόνη LCD

Για την οπτική απεικόνιση των μηνυμάτων χρησιμοποιήθηκε μια οθόνη υγρών κρυστάλλων (LCD Display), η οποία διαθέτει ενσωματωμένη μονάδα διεπαφής I2C. Το βασικό χαρακτηριστικό της οθόνης είναι ότι αποτελείται από δύο γραμμές με δεκαέξι χαρακτήρες ανά γραμμή.



Εικόνα 4.19 Απεικόνιση της οθόνης LCD με προβολή μηνύματος κατά τη λειτουργία του ρομποτικού οχήματος.



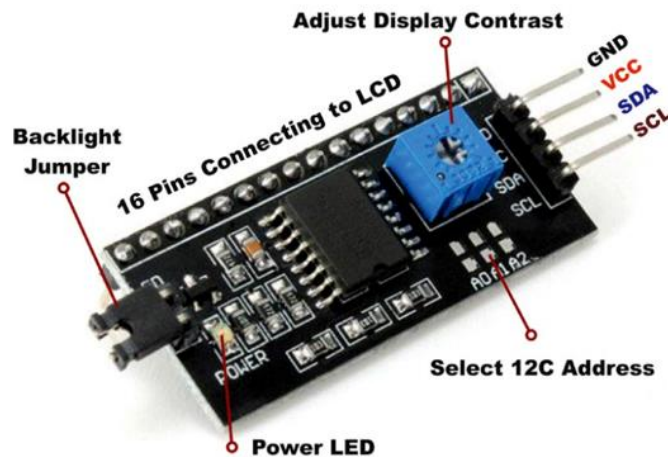
Εικόνα 4.20 Απεικόνιση της οθόνης LCD δύο γραμμών με δεκαέξι χαρακτήρες ανά γραμμή. (www.amazon.de)

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.



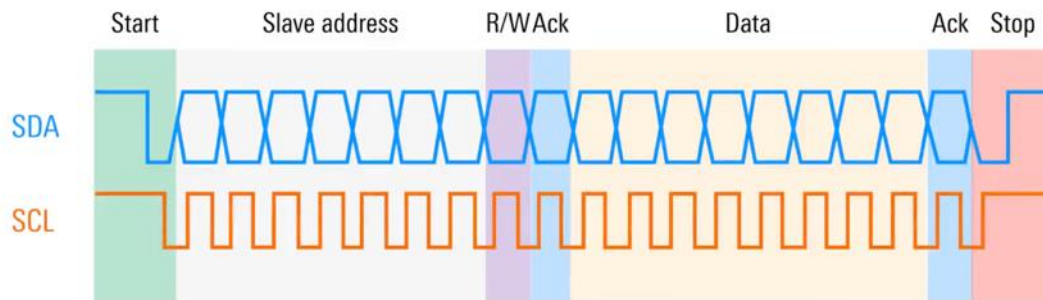
Εικόνα 4.21 Απεικόνιση της οθόνης LCD με ενσωματωμένη μονάδα διεπαφής I2C. (www.amazon.de)

Η οθόνη LCD χρησιμοποιείται για την προβολή του αρχικού μηνύματος, καθώς και για την απεικόνιση της θερμοκρασίας και της υγρασίας.



Εικόνα 4.22 Απεικόνιση της μονάδας διεπαφής I2C. (components101.com)

Κάθε δίαυλος του πρωτοκόλλου I2C αποτελείται από δύο γραμμές σημάτων: τη γραμμή SCL (Serial Clock), η οποία μεταφέρει το σήμα χρονισμού, και τη γραμμή SDA (Serial Data), η οποία μεταφέρει τα δεδομένα. Το σήμα χρονισμού παράγεται αποκλειστικά από την κύρια μονάδα ελέγχου του διαύλου. Ωστόσο, δύναται να διατηρηθεί το σήμα αυτό σε χαμηλή κατάσταση από μια δευτερεύουσα μονάδα για συγκεκριμένο χρονικό διάστημα. Με τον τρόπο αυτό, επιβραδύνεται η κύρια μονάδα ελέγχου από την αποστολή επιπλέον δεδομένων ή εξασφαλίζεται ο απαιτούμενος χρόνος για την προετοιμασία τους, προτού αυτή τα αναγνώσει. Η τεχνική αυτή ονομάζεται επιμήκυνση ρολογιού (clock stretching). (components101.com)



Εικόνα 4.23 Απεικόνιση του διαγράμματος χρονισμού του πρωτοκόλλου I2C.
(<https://www.youtube.com/watch?v=CAvawEcxoPU>)

Τεχνικά εγχειρίδια : <https://cdn.sparkfun.com/assets/9/5/f/7/b/HD44780.pdf>

Αισθητήρας θερμοκρασίας και υγρασίας

Στο πλαίσιο της παρούσας εργασίας χρησιμοποιήθηκε ο αισθητήρας DHT 11, με σκοπό την καταγραφή των περιβαλλοντικών παραμέτρων του χώρου πριν από την εκκίνηση του ρομποτικού οχήματος. Ο αισθητήρας DHT 11 είναι ένα ηλεκτρονικό στοιχείο, το οποίο αξιοποιείται για τη μέτρηση της θερμοκρασίας και της υγρασίας. Ο συγκεκριμένος αισθητήρας χρησιμοποιείται σε πληθώρα εφαρμογών, όπως σε οικιακούς αυτοματισμούς και σε βιομηχανικές κατασκευές.



Εικόνα 4.24 Απεικόνιση του αισθητήρα DHT 11. (www.electronicwings.com)

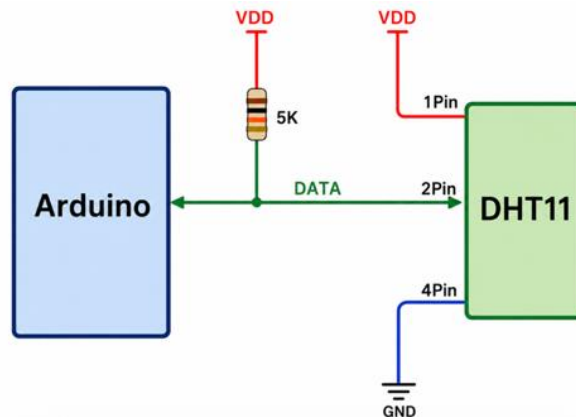
Ο αισθητήρας DHT11 είναι μια ψηφιακή μονάδα, η οποία ενσωματώνει έναν αναλογικό σε ψηφιακό μετατροπέα και δύο ανεξάρτητα αισθητήρια, ένα για την υγρασία και ένα για τη θερμοκρασία.

1. Ο αισθητήρας υγρασίας: Ο αισθητήρας λειτουργεί με βάση μια μεταβαλλόμενη αντίσταση. Αποτελείται από δύο ηλεκτρόδια, μεταξύ των οποίων παρεμβάλλεται ένα υπόστρωμα ευαίσθητο στην υγρασία, το οποίο είναι κατασκευασμένο από αγώγιμο πολυμερές υλικό. Η μεταβολή της υγρασίας του περιβάλλοντος προκαλεί αλλαγή στην αγωγιμότητα του υποστρώματος, με αποτέλεσμα να μεταβάλλεται η ηλεκτρική αντίσταση των ηλεκτροδίων. Με αυτόν τον τρόπο επιτρέποντας τη μέτρηση της σχετικής υγρασίας (% RH).

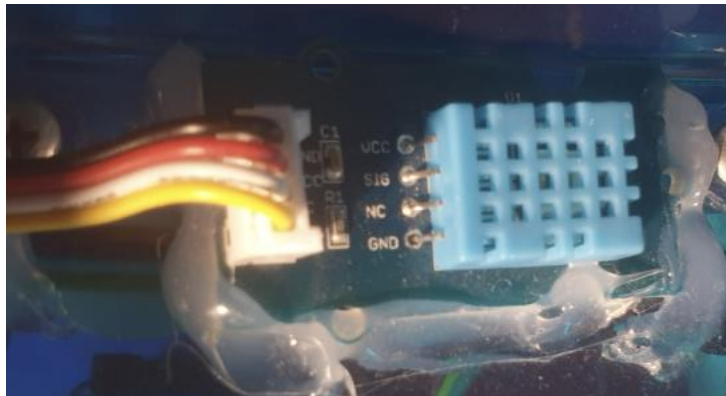
Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

2. Ο αισθητήρας θερμοκρασίας: Ο αισθητήρας ενσωματώνει ένα θερμίστορ αρνητικού συντελεστή θερμοκρασίας (NTC Thermistor), του οποίου η ηλεκτρική αντίσταση μειώνεται αντιστρόφως ανάλογα με την αύξηση της θερμοκρασίας.

Τα αναλογικά σήματα από τις εξόδους των δύο αισθητηρίων οδηγούνται σε έναν μικροελεγκτή. Ο οποίος επεξεργάζεται τα σήματα αυτά εσωτερικά, εφαρμόζει τους απαραίτητους συντελεστές βαθμονόμησης και ακολούθως μετατρέπει τις μετρήσεις σε ψηφιακό σήμα, το οποίο οδηγείται στην έξοδο της μονάδας.



Εικόνα 2.25 Απεικόνιση του σχηματικού διαγράμματος του αισθητήρα DHT11. (www.mouser.com)



Εικόνα 4.26 Απεικόνιση του αισθητήρα θερμοκρασίας και υγρασίας DHT 11.

Τεχνικά εγχειρίδια :

https://www.mouser.com/datasheet/2/758/DHT11-Technical-Data-Sheet-Translated-Version-1143054.pdf?srltid=AfmBOoqsl2SiolXAHBpbk7uWUuXrB_TOFNlswjuKMWtsSKignGaRXM6a

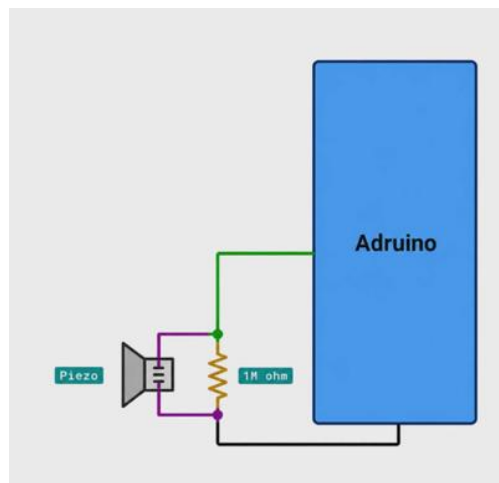
Παθητική ηχητική μονάδα πιεζοηλεκτρικού βομβητή (Buzzer)

Μέσω του βομβητή (Buzzer) παρέχεται η δυνατότητα αναπαραγωγής μονοτονικών ήχων. Ο βομβητής δύναται να παράγει έναν ακουστικό τόνο καθορισμένης συχνότητας και συγκεκριμένης διάρκειας ή μια μονοτονική μελωδία.

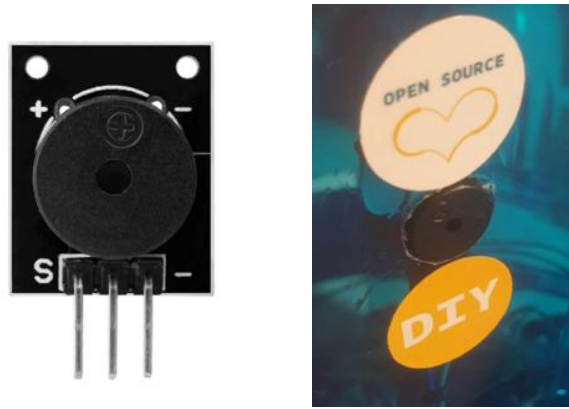
Ο πιεζοηλεκτρικός βομβητής αποτελεί μια ηλεκτροακουστική διάταξη, η οποία έχει τη δυνατότητα να μετατρέπει τα ηλεκτρικά σήματα σε ηχητικά κύματα. Η λειτουργία του βασίζεται στο πιεζοηλεκτρικό φαινόμενο, κατά το οποίο η εφαρμογή εναλλασσόμενης ηλεκτρικής τάσης στα άκρα ενός πιεζοηλεκτρικού υλικού δημιουργεί τη μηχανική παραμόρφωση του. Η κίνηση αυτή μεταφέρεται σε ένα μεταλλικό διάφραγμα, προκαλώντας τη δόνηση του, η οποία με τη σειρά της δημιουργεί διαταραχές στην πίεση του αέρα, παράγοντας έτσι ακουστικό σήμα. Ο πιεζοηλεκτρικός βομβητής διακρίνεται σε δύο κατηγορίες:

1. Ενεργός βομβητής (Active): Ενσωματώνει εσωτερικό κύκλωμα ταλάντωσης. Παράγει ήχο σταθερής συχνότητας με την απλή εφαρμογή συνεχούς τάσης (DC).
2. Παθητικός βομβητής (Passive): Δεν διαθέτει εσωτερικό ταλαντωτή και για να λειτουργήσει απαιτεί εξωτερικό σήμα τετραγωνικού κύματος (PWM). Το πλεονέκτημά του είναι ότι επιτρέπει τον πλήρη έλεγχο της συχνότητας και της διάρκειας, καθιστώντας δυνατή την αναπαραγωγή διαφορετικών τόνων και μελωδιών.

Στο ρομποτικό όχημα, ο βομβητής χρησιμοποιείται αποκλειστικά κατά την προσέγγιση της πηγής ήχου. Τη χρονική στιγμή που το ρομποτικό όχημα πλησιάσει το στόχο, την ηχητική πηγή, ενεργοποιείται η κόκκινη λυχνία (LED) και ο βομβητής παράγει έναν μονοτονικό ήχο συχνότητας 1000 Hz και διάρκειας δύο δευτερολέπτων. Η επιλογή της συχνότητας των 1000 Hz βασίζεται στο γεγονός ότι εμπίπτει στο φάσμα της μέγιστης ακουστικής ευαισθησίας του ανθρώπινου αυτιού, διασφαλίζοντας ότι η ειδοποίηση γίνεται άμεσα αντιληπτή. Για την υλοποίηση αυτή, χρησιμοποιείται ο παθητικός πιεζοηλεκτρικός βομβητής KY-006.



Εικόνα 4.27 Απεικόνιση του σχηματικού διαγράμματος σύνδεσης του βομβητή με το Arduino. (docs.arduino.cc)



Εικόνα 4.28 Απεικόνιση του παθητικού βομβητή KY-006, και η τοποθέτηση του στο ρομποτικό όχημα. (www.az-delivery.de).

Τεχνικά εγχειρίδια :

https://eclass.uth.gr/modules/document/file.php/E-CE_U_269/Sensors/Sensors_%20Datasheets/KY-006-Joy-IT.pdf#page=1.00

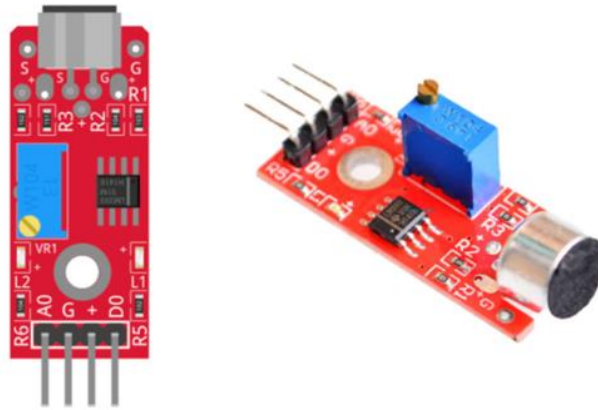
Μικρόφωνο

Στο ρομποτικό όχημα έχουν χρησιμοποιηθεί δύο μονάδες μικροφώνων, οι οποίες είναι τοποθετημένες στην αριστερή και στη δεξιά πλευρά του. Το μικρόφωνο βρίσκεται στο ίδιο ύψος, απέχοντας ίση απόσταση από το έδαφος, ενώ παράλληλα δεν εξέρχουν από το όχημα. Οι συγκεκριμένες μονάδες ανίχνευσης ήχου παρουσιάζουν υψηλή ευαισθησία της τάξης των -42 ± 3 dB και πρόκειται για το ηλεκτρονικό κύκλωμα KY-037.

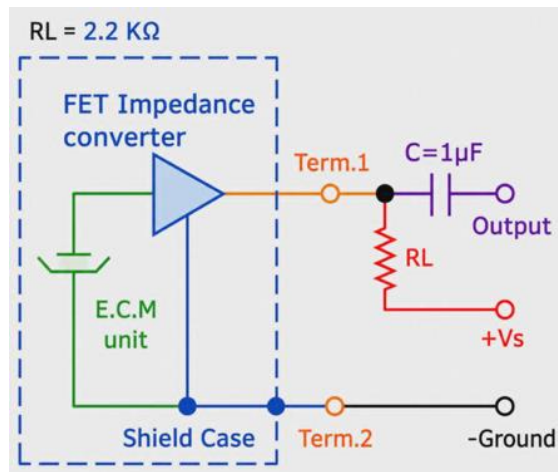
Το βασικό εξάρτημα είναι ένα πυκνωτικό μικρόφωνο τύπου electret, το CMA-6542PF, το οποίο έχει τη δυνατότητα καταγραφής των ηχητικών σημάτων του περιβάλλοντος. Επιπλέον, η μονάδα ενσωματώνει έναν διαφορικό συγκριτή, το LM393, ο οποίος χρησιμοποιείται για τον έλεγχο της ψηφιακής εξόδου.

Η κύρια λειτουργία της μονάδας είναι η ανίχνευση ηχητικών σημάτων που υπερβαίνουν ένα καθορισμένο κατώφλι, το οποίο ρυθμίζεται μέσω του ποτενσιόμετρου που διαθέτει. Η μονάδα διαθέτει τόσο αναλογική όσο και ψηφιακή έξοδο. Η ψηφιακή έξοδος παράγει ένα σήμα υψηλής στάθμης όταν ανιχνεύεται ηχητικό σήμα από ένα επίπεδο και πάνω. Αντίθετα, η αναλογική έξοδος παράγει τιμές σχετιζόμενες με το επίπεδο του ηχητικού σήματος που εντοπίζει το πυκνωτικό μικρόφωνο. Οι τιμές αυτές είναι ανάλογες της τάσης αλλά και της θέσης του ποτενσιόμετρου, γεγονός που καθιστά δυσχερή την ακριβή ανακατασκευή του αρχικού ηχητικού σήματος με βάση τα λαμβανόμενα δεδομένα.

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.



Εικόνα 4.29 Απεικόνιση της μονάδας του μικροφώνου KY-037. (www.az-delivery.de)



Εικόνα 4.30 Απεικόνιση σχηματικού διαγράμματος του μικροφώνου. (www.alldatasheet.com)



Εικόνα 4.31 Απεικόνιση της καμπύλης απόκρισης συχνότητας του μικροφώνου. (www.alldatasheet.com)



Εικόνα 4.32 Απεικόνιση της κυματομορφής των τιμών της αναλογικής εξόδου του μικροφώνου.



Εικόνα 4.33 Απεικόνιση της μονάδας μικροφώνου στη δεξιά πλευρά του ρομποτικού οχήματος

Τα σήματα από την έξοδο των δύο μικροφώνων οδηγούνται στην πλατφόρμα του Arduino. Στη συνέχεια, πραγματοποιείται σύγκριση των ηχητικών σημάτων μέσω λογισμικού, με βάση την ενεργειακή διαφορά μεταξύ τους. Με τον τρόπο αυτό, γίνεται η ακριβής εκτίμηση για τον εντοπισμό της ηχητικής πηγής.

Τεχνικά εγχειρίδια :

https://www.sameskydevices.com/product/resource/cma-6542pf.pdf?srltid=AfmBOqgbVlrOKUabNJwVRe3JMz7wN5TbBGOXH3Uufm2_wNEzYaUQ3Y0Q

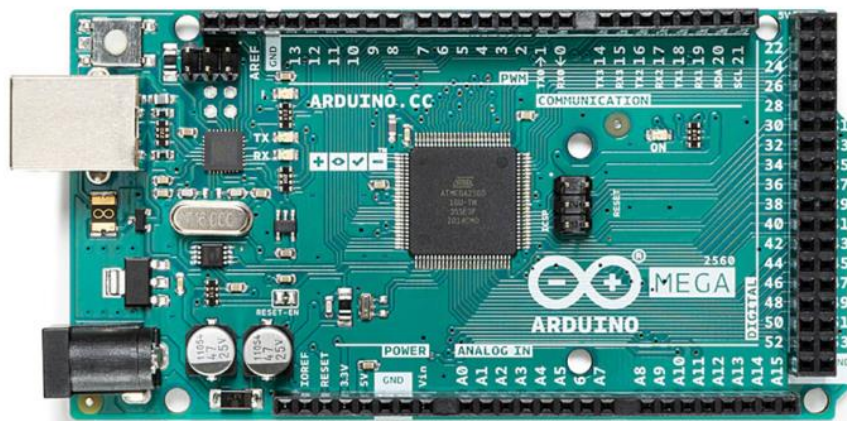
Arduino

Το ρομποτικό όχημα βασίζεται στην πλατφόρμα Arduino Mega 2560. Η συγκεκριμένη πλατφόρμα ενσωματώνει την κεντρική μονάδα επεξεργασίας, καθώς και τη μνήμη στην οποία αποθηκεύεται το λογισμικό ελέγχου των λειτουργιών του ρομποτικού οχήματος.

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

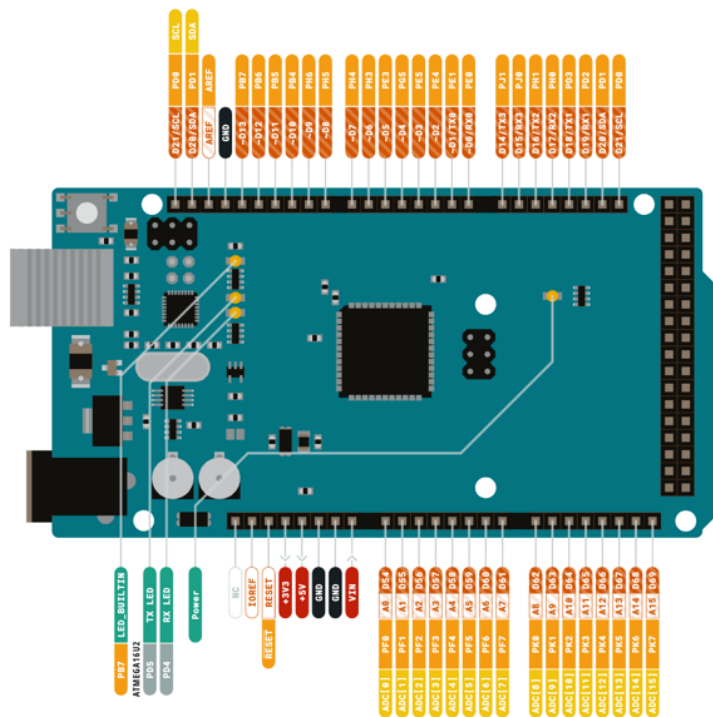
Το Arduino Mega αποτελεί μια ηλεκτρονική πλατφόρμα ανοιχτού κώδικα, η οποία έχει την δυνατότητα εύκολης διασύνδεσης με περιφερειακά εξαρτήματα και παρέχει ευκολία για τον προγραμματισμό της. Η πλατφόρμα διαθέτει 54 ψηφιακές εισόδους και εξόδους. Από αυτές, οι 15 δύναται να διαμορφωθούν ως έξοδοι διαμόρφωσης εύρους παλμού (PWM), προκειμένου να χρησιμοποιηθούν για τον έλεγχο της ταχύτητας των κινητήρων καθώς και της γωνιακής θέσης των σερβοκινητήρων του ρομποτικού οχήματος. Επιπλέον, έχει 16 αναλογικές εισόδους για την λήψη δεδομένων από διάφορους αισθητήρες, όπως οι μονάδες ανίχνευσης ήχου. Διαθέτει 4 θύρες UART, οι οποίες διευκολύνουν την συνδεσιμότητα της πλατφόρμας με εξωτερικές μονάδες τηλεμετρίας, όπως μονάδες Bluetooth.

Η πλακέτα βασίζεται στον 8-bit μικροελεγκτή ATmega 2560 και είναι εξοπλισμένη με ένα κρύσταλλο χρονισμού (ταλαντωτή) με συχνότητα 16 MHz. Η συχνότητα του κρυστάλλου εξασφαλίζει την απαραίτητη ταχύτητα επεξεργασίας για την εκτέλεση των λειτουργιών του ρομποτικού οχήματος σε πραγματικό χρόνο, με παράλληλη διατήρηση της ενεργειακής κατανάλωσης του ρομποτικού οχήματος σε χαμηλά επίπεδα.

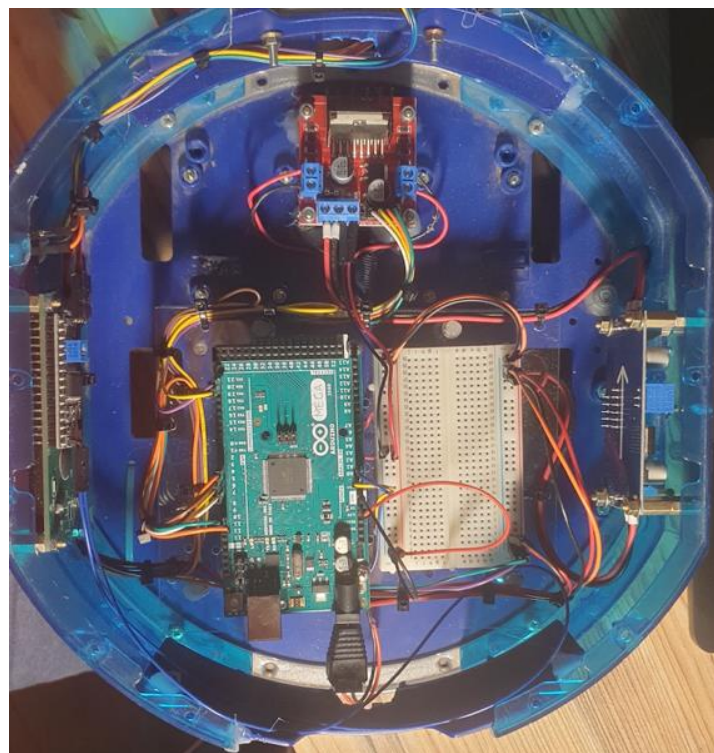


Εικόνα 4.34 Απεικόνιση της πλατφόρμας του Arduino Mega. (www.arduino.cc)

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.



Εικόνα 4.35 Απεικόνιση του σχηματικού διαγράμματος των ακροδεκτών του Arduino Mega 2560. (www.arduino.cc)



Εικόνα 4.36 Απεικόνιση της κάτοψης της βάσης του ρομπότ με την τοποθέτηση της πλατφόρμας του Arduino Mega.

Τεχνικά εγχειρίδια : <https://docs.arduino.cc/resources/datasheets/A000067-datasheet.pdf>

Επιπλέον περιφερειακά στοιχεία

1. Το ρομποτικό όχημα διαθέτει ένα πλήκτρο (Button) για τη χειροκίνητη επανεκκίνηση του, όταν αυτό κριθεί απαραίτητο από τον χρήστη.
2. Μια πράσινη λυχνία (LED), η οποία παρέχει οπτική ένδειξη για τη λειτουργία του ρομποτικού οχήματος.
3. Μια κίτρινη λυχνία (LED), η οποία ενεργοποιείται για να υποδηλώσει ότι το ρομποτικό όχημα βρίσκεται σε κατάσταση κίνησης.
4. Μια κόκκινη λυχνία (LED), η οποία ενεργοποιείται για να υποδηλώσει ότι το ρομποτικό όχημα εντόπισε κάποιο εμπόδιο στην πορεία του ή ότι προσέγγισε επιτυχώς τον τελικό του στόχο, ο οποίος είναι η ηχητική πηγή.

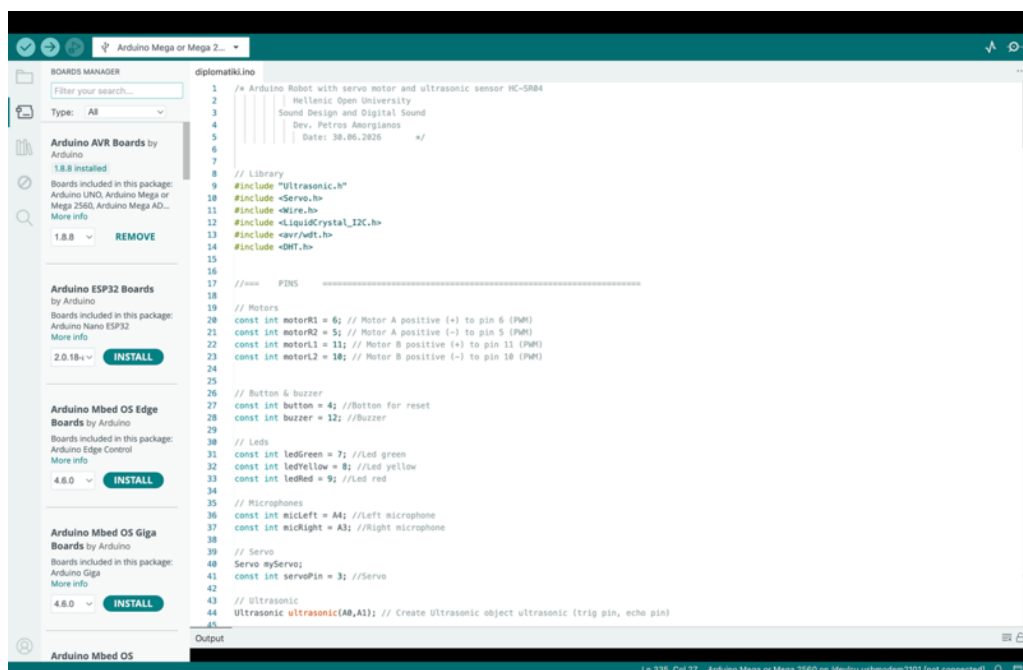


Εικόνα 4.37 Απεικόνιση του ρομποτικού οχήματος κατά τη λειτουργία κίνησης προς την ηχητική πηγή.

4.3 Ανάλυση αλγόριθμου ρομποτικού οχήματος

Στη συγκεκριμένη ενότητα του κεφαλαίου πραγματοποιείται η ανάλυση του λογισμικού για τη λειτουργία του ρομποτικού οχήματος. Στο παράρτημα Α, παρουσιάζεται ο πλήρης κώδικας που χρησιμοποιήθηκε για τη βασική λειτουργία του οχήματος, ενώ στο παράρτημα Β παρουσιάζεται το τροποποιημένος κώδικας για τον εντοπισμό της ηχητικής πηγής κατά την κίνηση του ρομποτικού οχήματος.

Το πρόγραμμα του ρομποτικού οχήματος εκτελείται στο Arduino Mega και αναπτύχθηκε στο περιβάλλον Arduino IDE, το οποίο αποτελεί το επίσημο λογισμικό προγραμματισμού της πλατφόρμας Arduino.



Εικόνα 4.40. Απεικόνιση του περιβάλλοντος ανάπτυξης Arduino IDE, το οποίο χρησιμοποιήθηκε για τον προγραμματισμό του ρομποτικού οχήματος.

Στην αρχή του κώδικα πραγματοποιείται η κλήση των απαραίτητων βιβλιοθηκών, οι οποίες απαιτούνται για τη λειτουργία των περιφερειακών μονάδων του ρομποτικού οχήματος.

```
8 // Library
9 #include "Ultrasonic.h"
10 #include <Servo.h>
11 #include <Wire.h>
12 #include <LiquidCrystal_I2C.h>
13 #include <avr/wdt.h>
14 #include <DHT.h>
15
```

Εικόνα 4.41 Απεικόνιση των απαραίτητων βιβλιοθηκών για τη λειτουργία του ρομποτικού οχήματος.

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Στη συνέχεια του κώδικα καθορίζεται η αντιστοίχιση των ακροδεκτών για τη διασύνδεση των περιφερειακών μονάδων με το Arduino. Με τον τρόπο αυτό, οι φυσικές συνδέσεις του υλικού, ορίζονται σε επίπεδο του λογισμικού.

```
17 //==== PINS =====
18
19 // Motors
20 const int motorR1 = 6; // Motor A positive (+) to pin 6 (PWM)
21 const int motorR2 = 5; // Motor A positive (-) to pin 5 (PWM)
22 const int motorL1 = 11; // Motor B positive (+) to pin 11 (PWM)
23 const int motorL2 = 10; // Motor B positive (-) to pin 10 (PWM)
24
25
26 // Button & buzzer
27 const int button = 4; //Botton for reset
28 const int buzzer = 12; //Buzzer
29
30 // Leds
31 const int ledGreen = 7; //Led green
32 const int ledYellow = 8; //Led yellow
33 const int ledRed = 9; //Led red
34
35 // Microphones
36 const int micLeft = A4; //Left microphone
37 const int micRight = A3; //Right microphone
38
39 // Servo
40 Servo myServo;
41 const int servoPin = 3; //Servo
42
43 // Ultrasonic
44 Ultrasonic ultrasonic(A0,A1); // Create Ultrasonic object ultrasonic (trig pin, echo pin)
45
46 // DHT11
47 #define DHTPIN 13 //Sensor temperatur and humanity
48 #define DHTTYPE DHT11
49 DHT dht(DHTPIN, DHTTYPE);
50
51 // LCD
52 LiquidCrystal_I2C lcd(0x3F, 16, 2); // set the LCD address to 0x3F for a 16 chars and 2 line display
```

Εικόνα 4.42. Απεικόνιση των συνδέσεων των περιφερειακών μονάδων στη πλατφόρμα του Arduino.

Στη συνέχεια, ορίζονται οι απαραίτητες μεταβλητές για τον έλεγχο της λειτουργίας των περιφερειακών μονάδων του ρομποτικού οχήματος.

```
55 //=== Variables =====
56
57 //Variables
58 int distance; // Variable to store distance from an object
59 int threshold = 20;
60
61 int soundLevel = 0;
62
63 // Microphones calibration
64 int baseLeft = 0;
65 int baseRight = 0;
66
67 // Filter
68 float alpha = 0.7; //Filter to reduce environmental noise
69 int smoothLeft = 0;
70 int smoothRight = 0;
71
72 // FSM
73 enum State {
74     WAIT_SOUND, MOVE_TO_SOUND, AVOID_OBSTACLE, TARGET_REACHED
75 };
76
77 State currentState = WAIT_SOUND;
78 int lockedDirection = 0; // -1 left, 0 forward, 1 right
79
80 ~~~
```

Εικόνα 4.43. Απεικόνιση των απαραίτητων μεταβλητών για την λειτουργία των περιφερειακών μονάδων του ρομποτικού οχήματος.

Ακολούθως, ορίζονται οι συναρτήσεις που αφορούν τον έλεγχο και τη διαχείριση της κίνησης του ρομποτικού οχήματος.

```
81 //=== Functions =====
82
83 int readMicLevel(int pin);
84 void stop();
85 void forward();
86 void backward();
87 void left();
88 void right();
89 void resetArduino();
90
```

Εικόνα 4.44 Απεικόνιση των συναρτήσεων ελέγχου κίνησης του ρομποτικού οχήματος.

Στη συνέχεια, στο τμήμα του κώδικα για τη συνάρτηση αρχικοποίησης, καθορίζεται εάν οι ακροδέκτες των περιφερειακών μονάδων θα λειτουργούν ως είσοδοι ή έξοδοι για το Arduino. Παράλληλα, ορίζεται ο ρυθμός μετάδοσης δεδομένων για την επικοινωνία μέσω της σειριακής θύρας. Επιπλέον, πραγματοποιείται η αρχικοποίηση των μικροφώνων, όπου λαμβάνεται ένας αριθμός αρχικών μετρήσεων από το καθένα, προκειμένου να υπολογιστεί η μέση τιμή τους, η οποία θα χρησιμοποιηθεί για τη βαθμονόμηση τους. Επισημαίνεται ότι τα μικρόφωνα έχουν ρυθμιστεί για αναλογική λειτουργία. Τέλος, πραγματοποιείται η αρχικοποίηση του αισθητήρα θερμοκρασίας και υγρασίας.

```
92 //=== Setup =====
93
94 void setup(){
95
96   pinMode(motorR1, OUTPUT);
97   pinMode(motorR2, OUTPUT);
98   pinMode(motorL1, OUTPUT);
99   pinMode(motorL2, OUTPUT);
100
101   pinMode(ledGreen, OUTPUT);
102   pinMode(ledYellow, OUTPUT);
103   pinMode(ledRed, OUTPUT);
104
105   digitalWrite(ledGreen, HIGH);
106
107   pinMode(buzzer, OUTPUT);
108   pinMode(button, INPUT_PULLUP);
109
110   Serial.begin(9600);
111
112   myServo.attach(servoPin); // servo fixed
113   myServo.write(70);
114
115 //=== Microphones =====
116
117   delay(2000);
118   for (int i=0; i<50; i++) {
119     baseLeft += analogRead(micLeft);
120     baseRight += analogRead(micRight);
121     delay(10);
122   }
123
124   baseLeft /= 50;
125   baseRight /= 50;
126
127 //=== DHT sensor =====
128
129   dht.begin();
130   delay (1500);
131 }
```

Εικόνα 4.45 Απεικόνιση του καθορισμού της λειτουργίας των περιφερειακών μονάδων που συνδέονται στην πλατφόρμα του Arduino.

Ακολούθως, παρουσιάζεται ο αλγόριθμος για την προβολή των μηνυμάτων στην οθόνη LCD του ρομποτικού οχήματος, καθώς και των ενδείξεων της υγρασίας, θερμοκρασίας του χώρου. Ειδικότερα, το αρχικό μήνυμα που εμφανίζεται περιλαμβάνει τα εξής στοιχεία

«Hellenic Open University
Acoustic design and Digital Sound
Petros Amorgianos»

Στη συνέχεια, προβάλλονται οι μετρήσεις της θερμοκρασίας και της υγρασίας. Επιπλέον, καθ' όλη τη διάρκεια λειτουργίας του ρομποτικού οχήματος, παραμένει ορατό το μήνυμα

«Petros Amorgianos»

```
132 //=== LCD =====
133
134 lcd.init();           // initialize the lcd
135 lcd.backlight();      // Turn on the LCD screen backlight
136
137 lcd.setCursor(0, 0);
138 lcd.print(" Hellenic Open ");
139 lcd.setCursor(0, 1);
140 lcd.print(" University ");
141 delay(3000);
142 lcd.clear();
143
144 lcd.setCursor(1, 0);
145 lcd.print("Acoustic design");
146 delay(3000);
147 lcd.clear();
148 lcd.setCursor(1, 0);
149 lcd.print(" and ");
150 delay(3000);
151 lcd.clear();
152 lcd.setCursor(2, 0);
153 lcd.print("Digital Sound");
154 delay(3000);
155 lcd.clear();
156
157 lcd.setCursor(0, 0);
158 lcd.print(" Petros ");
159 lcd.setCursor(0, 1);
160 lcd.print(" Amorgianos ");
161 lcd.clear();
162
163 float h = dht.readHumidity();
164 float t = dht.readTemperature();
165
166 lcd.setCursor(0,0);
167 lcd.print("Temp: ");
168 lcd.print(t);
169 lcd.print(" C");
170
171 lcd.print("Hum: ");
172 lcd.print(h);
173 lcd.print(" %");
174
175 delay(3000);
176 lcd.clear();
177
178 lcd.setCursor(0, 0);
179 lcd.print(" Petros ");
180 lcd.setCursor(0, 1);
181 lcd.print(" Amorgianos ");
182
183 }
184
185
```

Εικόνα 4.46 Απεικόνιση του αλγορίθμου προβολής των μηνυμάτων στην οθόνη LCD.

Ακολούθως, παρουσιάζεται το τμήμα του αλγορίθμου που εκτελείται επαναληπτικά και καθορίζει τη συνεχή λειτουργία του ρομποτικού οχήματος. Συγκεκριμένα, έχει ενσωματωθεί η λειτουργία επαναφοράς, ώστε σε περίπτωση εμφάνισης οποιοδήποτε σφάλματος, να παρέχεται η δυνατότητα χειροκίνητης επανεκκίνησης του συστήματος μέσω ενός διακόπτη (reset). Ακολούθως, ενεργοποιείται η περιφερειακή μονάδα των υπερήχων, η οποία επιτρέπει στο ρομποτικό όχημα να αποφεύγει τις προσκρούσεις στα εμπόδια. Στο επόμενο στάδιο, καθορίζεται η διαδικασία ανάγνωσης του ηχητικού σήματος από τα δύο μικρόφωνα και πραγματοποιείται το φιλτράρισμα των μετρήσεων με σκοπό τη μείωση του περιβαλλοντικού θορύβου. Τέλος, η διαδικασία ολοκληρώνεται με τον υπολογισμό της συνολικής στάθμης ακουστικής πίεσης, καθώς και της διαφοράς έντασης που προκύπτει ανάμεσα στον ήχο που λαμβάνεται από τα δύο μικρόφωνα.

```
186 //== LOOP =====
187
188 void loop() {
189
190 //== RESET BUTTON =====
191
192 if (digitalRead(button) == LOW) {
193   resetArduino();
194 }
195
196
197 //== Distance =====
198
199 distance = ultrasonic.Ranging(CM);
200
201 //== SOUND =====
202
203 int rawLeft = readMicLevel(micLeft);
204 int rawRight = readMicLevel(micRight);
205
206 //== FILTER =====
207
208 smoothLeft = alpha * smoothLeft + (1 - alpha) * rawLeft;
209 smoothRight = alpha * smoothRight + (1 - alpha) * rawRight;
210
211 int leftSound = smoothLeft;
212 int rightSound = smoothRight;
213
214 soundLevel = (leftSound + rightSound) / 2;
215 int diff = leftSound - rightSound;
216 // int deadZone = 15;
217
```

Εικόνα 4.47 Απεικόνιση των αρχικών λειτουργιών του κύριου προγράμματος του ρομποτικού οχήματος.

Στη συνέχεια, παρουσιάζεται ο κώδικας που αφορά την κύρια λειτουργία του ρομποτικού οχήματος. Στο στάδιο αυτό, το όχημα βρίσκεται σε κατάσταση αναμονής για την ανίχνευση κάποιου ηχητικού σήματος. Μόλις εντοπιστεί το σήμα, προσδιορίζεται η κατεύθυνση προέλευσης του με βάση το μικρόφωνο που το κατέγραψε, και το ρομποτικό όχημα κινείται προς την αντίστοιχη κατεύθυνση. Εάν κατά τη διαδρομή ανιχνευτεί κάποιο εμπόδιο, ενεργοποιείται η κόκκινη λυχνία, το όχημα αλλάζει πορεία για να το αποφύγει και συνεχίζει προς τον τελικό στόχο. Μόλις προσεγγίσει την ηχητική πηγή, ενεργοποιείται εκ νέου η κόκκινη λυχνία και ο βομβητής, ο οποίος παράγει ένα ακουστικό σήμα. Τέλος, μεσολαβεί παύση πέντε δευτερολέπτων προτού το ρομποτικό όχημα επανέλθει σε κατάσταση αναμονής για το επόμενο ηχητικό σήμα.

```
218 // ===== CURRENT STATES =====
219
220 switch (currentState) {
221
222 // ===== 1. WAIT =====
223
224 case WAIT_SOUND:
225
226     digitalWrite(ledYellow, LOW);
227     digitalWrite(ledRed, LOW);
228     stop();
229
230     if (soundLevel > 20) {
231
232         if (diff > threshold) lockedDirection = -1;
233         else if (diff < -threshold) lockedDirection = 1;
234         else lockedDirection = 0;
235
236         currentState = MOVE_TO_SOUND;
237     }
238
239     break;
240
241 // ===== 2. MOVE =====
242
243 case MOVE_TO_SOUND:
244
245     digitalWrite(ledYellow, HIGH);
246
247     if (distance < 10 ) {
248         currentState = TARGET_REACHED;
249         break;
250     }
251
252     if (distance < 20 ) {
253         currentState = AVOID_OBSTACLE;
254         break;
255     }
256
257     if (distance < 10 ) {
258         currentState = AWAIT;
259         currentState = TARGET_OBSTACLE;
260         break;
261     }
262
263     if (distance < 20 ) {
264         currentState = TARGET_OBSTACLE;
265         break;
266     }
267
268     currentState = MOVE_TO_SOUND;
269
270 }
271
272 if (lockedDirection == -1) left();
273 else if (lockedDirection == 1) right();
274 else forward();
275
276 break;
277
278 // ===== 3. AVOID =====
279
280 case AVOID_OBSTACLE:
281
282     digitalWrite(ledYellow, LOW);
283     digitalWrite(ledRed, HIGH);
284
285     digitalWrite(ledYellow, LOW);
286     digitalWrite(ledRed, LOW);
287     digitalWrite(ledRed, HIGH);
288
289     stop();
290
291     backward();
292     delay(300);
293
294     right();
295     delay(400);
296
297
298     digitalWrite(ledRed, LOW);
299     digitalWrite(ledYellow, HIGH);
300
301     currentState = MOVE_TO_SOUND;
302
303 break;
304
305 // ===== 4. TARGET =====
306
307 case TARGET_REACHED:
308
309     stop();
310
311     digitalWrite(ledYellow, LOW);
312     digitalWrite(ledRed, HIGH);
313
314     tone(buzzer, 1000);
315     delay(3000);
316     delay(3000);
317     noTone(buzzer);
318
319     delay(5000);
320
321
322 // ===== 3. AVOID =====
323
324 void loop() {
325
326     digitalWrite(ledRed, LOW);
327
328     digitalWrite(ledRed, LOW);
329
330     currentState = WAIT_SOUND;
331
332     break;
333
334 }
335
336 delay(50);
337
338 }
339 }
```

Εικόνα 4.48 Απεικόνιση του κύριου τμήματος του προγράμματος που καθορίζει τη λειτουργία του ρομποτικού οχήματος βάσει του λαμβανόμενου ηχητικού σήματος.

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Στη συνέχεια του προγράμματος, υλοποιείται ο υπολογισμός της στάθμης ακουστικής πίεσης ανά μικρόφωνο. Ακολουθώντας, διαμορφώνεται η συνάρτηση επαναφοράς του ρομποτικού οχήματος.

```
310 //=== MICROPHONE =====
311
312 int readMicLevel(int pin) {
313     long sum = 0;
314     for(int i=0; i<15; i++) {
315         int val = analogRead(pin);
316         if(pin == micLeft)
317             sum += abs(val - baseLeft);
318         else
319             sum += abs(val - baseRight);
320     }
321     return sum / 15;
322 }
323
324 //=== RESET =====
325
326 void resetArduino() {
327     wdt_enable(WDTO_15MS);
328     while(1) {}
329 }
330
```

Εικόνα 4.49 Απεικόνιση του τμήματος του προγράμματος που περιλαμβάνει τις συναρτήσεις υπολογισμού της στάθμης ήχου και επαναφοράς του ρομποτικού οχήματος.

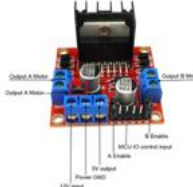
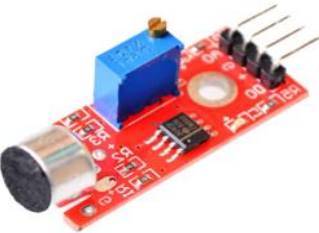
Στο τελευταίο μέρος του προγράμματος, παρουσιάζεται το τμήμα του κώδικα που αφορά την πλοήγηση και την κατεύθυνση του ρομποτικού οχήματος, καθορίζοντας τη λειτουργία των δύο βασικών κινητήριων τροχών του.

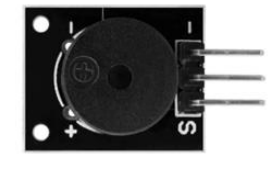
```
331 //=== MOVEMENT =====
332
333 //Movement functions
334 void forward(){
335     digitalWrite(motorR1, HIGH);
336     digitalWrite(motorR2, LOW);
337     digitalWrite(motorL1, HIGH);
338     digitalWrite(motorL2, LOW);
339 }
340
341 void backward(){
342     digitalWrite(motorR1, LOW);
343     digitalWrite(motorR2, HIGH);
344     digitalWrite(motorL1, LOW);
345     digitalWrite(motorL2, HIGH);
346 }
347
348 void right(){
349     digitalWrite(motorR1, LOW);
350     digitalWrite(motorR2, HIGH);
351     digitalWrite(motorL1, HIGH);
352     digitalWrite(motorL2, LOW);
353 }
354
355 void left(){
356     digitalWrite(motorR1, HIGH);
357     digitalWrite(motorR2, LOW);
358     digitalWrite(motorL1, LOW);
359     digitalWrite(motorL2, HIGH);
360 }
361
362 void stop(){
363     digitalWrite(motorR1, LOW);
364     digitalWrite(motorR2, LOW);
365     digitalWrite(motorL1, LOW);
366     digitalWrite(motorL2, LOW);
367 }
368
```

Εικόνα 4.50 Απεικόνιση του τμήματος του προγράμματος που αφορά τη λειτουργία κίνησης του ρομποτικού οχήματος.

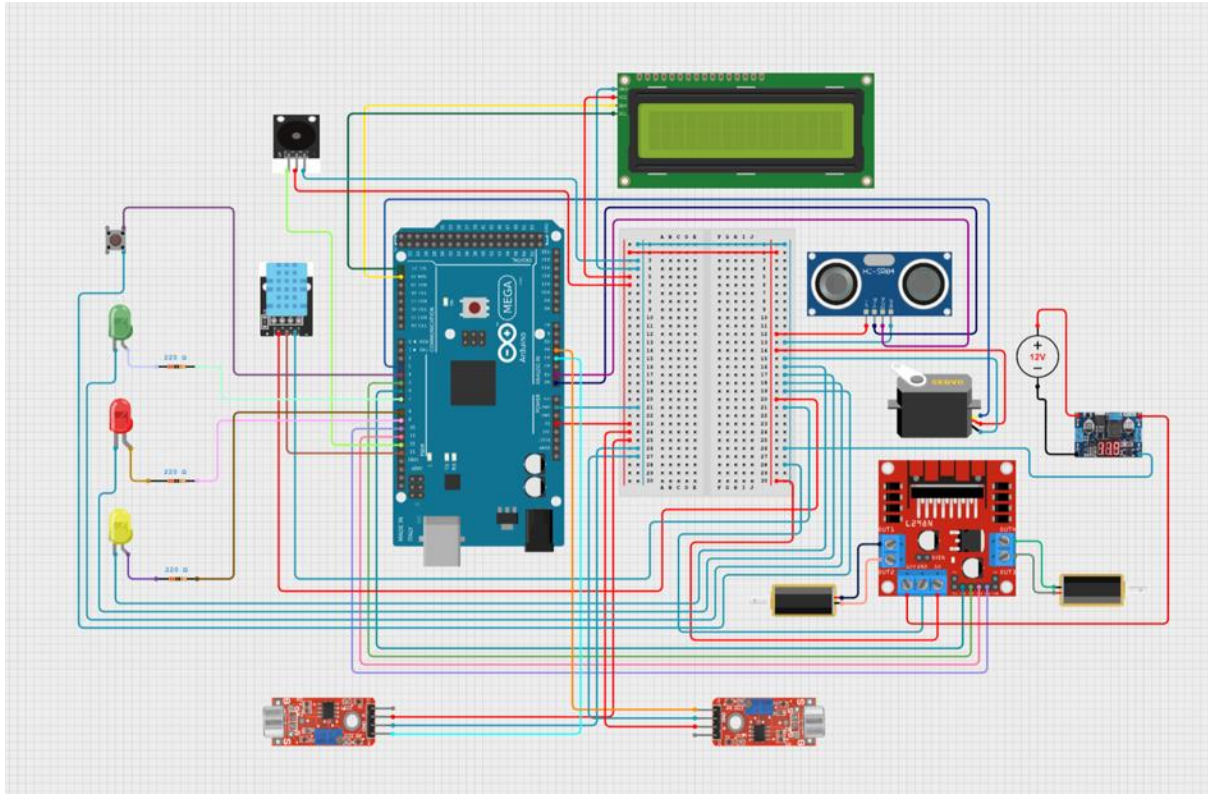
4.4 Σχεδίαση και ανάπτυξη του ρομποτικού οχήματος

Στη συνέχεια, παρατίθεται μια σύντομη παρουσίαση των περιφερειακών μονάδων που απαρτίζουν το ρομποτικό όχημα.

a/a	Διάφορα τμήματα του ρομποτικού οχήματος		Λειτουργία	Σύνδεση
1	Arduino Mega		Κεντρική πλατφόρμα	
2	LCD Display with I2C		Εμφάνιση μηνυμάτων	Pin 20 Pin 21
3	Ultrasonic HC-SR04		Έλεγχος για εμπόδια	Pin A0 Pin A1
4	Servo Motor		Υποβοήθηση για κίνηση του αισθητήρα υπέρυθρων	Pin 3
5	L298N Motor driver controller Board Modul		Οδήγηση των κινητήρων	Pin 5 Pin 6 Pin 10 Pin 11
6	Motor Right		Κίνηση δεξιού τροχού	
7	Motor Left		Κίνηση αριστερού τροχού	
8	Microphone Sensor Modul Right		Εντοπισμός ηχητικών σημάτων από δεξιά	Pin A4
9	Microphone Sensor Modul Left		Εντοπισμός ηχητικών σημάτων από αριστερά	Pin A3

10	Temperature Senso DHT 11		Έλεγχος θερμοκρασίας και υγρασίας του χώρου	Pin 13
11	Passives Piezo-Buzzer Modul		Εκπομπή ενός ηχητικού τόνου κατά τον εντοπισμό της ηχητικής πηγής	Pin 12
12	Button		Πραγματοποιεί χειροκίνητη επανεκκίνηση του ρομποτικού οχήματος	Pin 4
13	Led red		Εντοπισμός εμποδίου ή ηχητικής πηγής	Pin 9
14	Led green		Λειτουργία ρομποτικού οχήματος	Pin 7
15	Led yellow		Κίνηση ρομποτικού οχήματος	Pin 8

Πίνακας 4.2 Απεικόνιση των δομικών στοιχείων του ρομποτικού οχήματος, τις ακριβείς λειτουργίες τους και τη σύνδεση τους με την κεντρική πλατφόρμα.



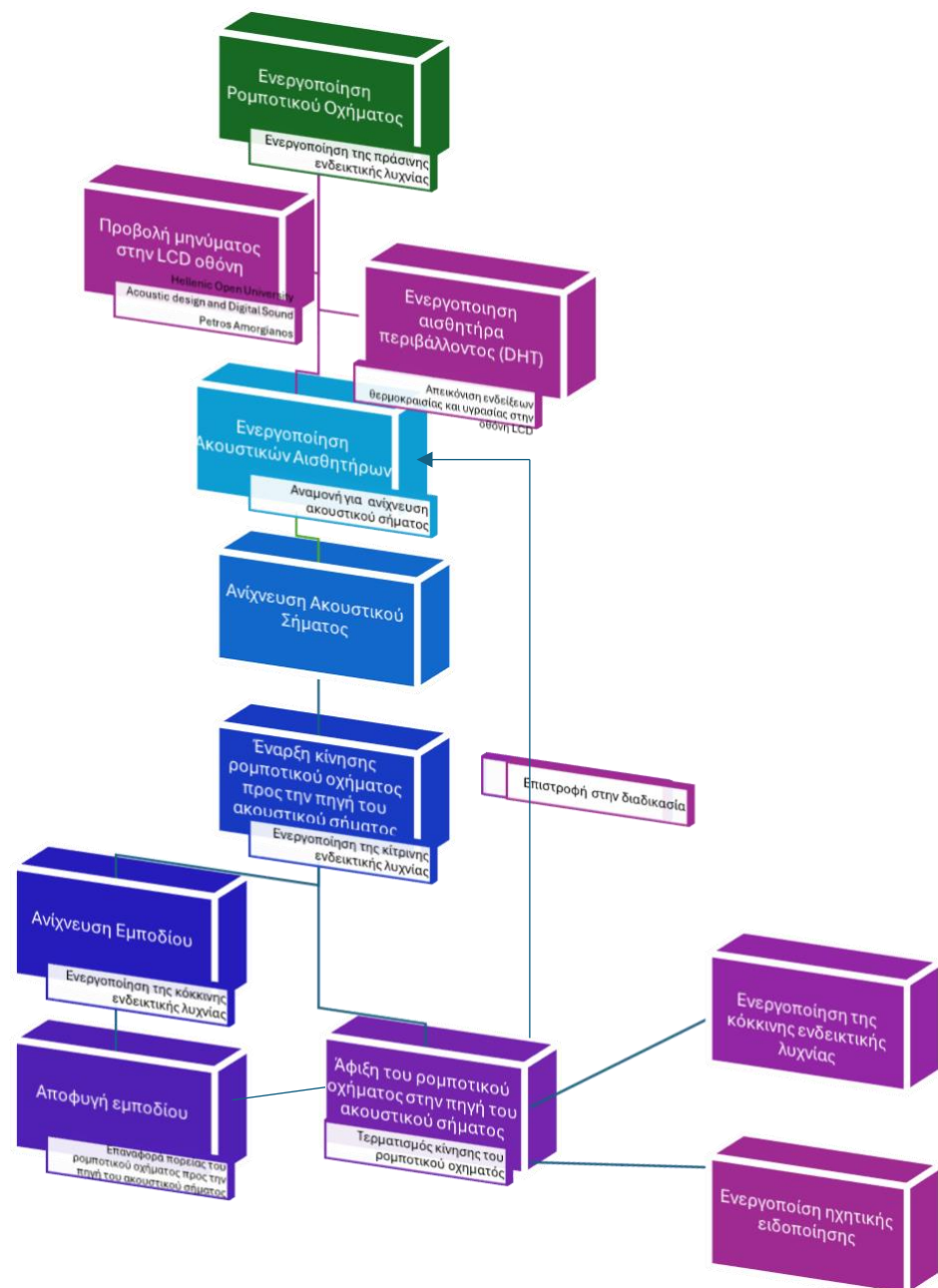
Εικόνα 4.51 Απεικόνιση του σχηματικού διαγράμματος του ηλεκτρονικού κυκλώματος του ρομποτικού οχήματος.

Περιγραφή λειτουργίας του ρομποτικού οχήματος

1. Ενεργοποίηση του ρομποτικού οχήματος.
2. Ενεργοποίηση της πράσινης λυχνίας (LED), η οποία παραμένει αναμμένη καθ' όλη τη διάρκεια λειτουργίας του ρομποτικού οχήματος.
3. Ενεργοποίηση της οθόνης LCD και προβολή του ακόλουθου εισαγωγικού μηνύματος

“ Hellenic Open University
Acoustic design and Digital Sound
Petros Amorgianos ”

4. Προβολή των μετρήσεων θερμοκρασίας και υγρασίας του χώρου στην οθόνη.
5. Μετάβαση σε κατάσταση αναμονής για την ανίχνευση ηχητικού σήματος.
6. Ανίχνευση ηχητικού σήματος και κίνηση του ρομποτικού οχήματος προς την κατεύθυνση της ηχητικής πηγής, με παράλληλη ενεργοποίηση της κίτρινης λυχνίας (LED).
7. Εντοπισμός πιθανού εμποδίου, ακινητοποίηση του ρομποτικού οχήματος και ενεργοποίηση της κόκκινης λυχνίας (LED).
8. Παράκαμψη του εμποδίου και ταυτόχρονη απενεργοποίηση της κόκκινης λυχνίας (LED).
9. Συνέχιση της πορείας προς την κατεύθυνση της ηχητικής πηγής.
10. Εντοπισμός της ηχητικής πηγής, ακινητοποίηση του οχήματος, ενεργοποίηση της κόκκινης λυχνίας (LED) και παραγωγή ακουστικού τόνου από τον βομβητή.
11. Ολοκλήρωση της αποστολής και εκ νέου μετάβαση σε κατάσταση αναμονής για ανίχνευση ηχητικού σήματος.



Εικόνα 4.52 Απεικόνιση δομικού διαγράμματος υλοποίησης των λειτουργιών του ρομποτικού οχήματος

Σημείωση: Διευκρινίζεται ότι η παρούσα εργασία επικεντρώνεται στη βασική έκδοση του ρομποτικού οχήματος. Η ανίχνευση του ακουστικού σήματος πραγματοποιείται αποκλειστικά κατά την εκκίνηση, καθώς και πέντε δευτερόλεπτα μετά την άφιξη στην ηχητική πηγή. Κατά το ενδιάμεσο διάστημα λειτουργίας, δεν πραγματοποιείται ανίχνευση ακουστικών σημάτων. Η εναλλακτική υλοποίηση για τον συνεχή εντοπισμό ακουστικών σημάτων, η οποία δεν απαιτεί τροποποίηση των ηλεκτρονικών κυκλωμάτων, παρουσιάζεται στο παράρτημα Β.

4.5 Παρουσίαση της ηχητικής πηγής

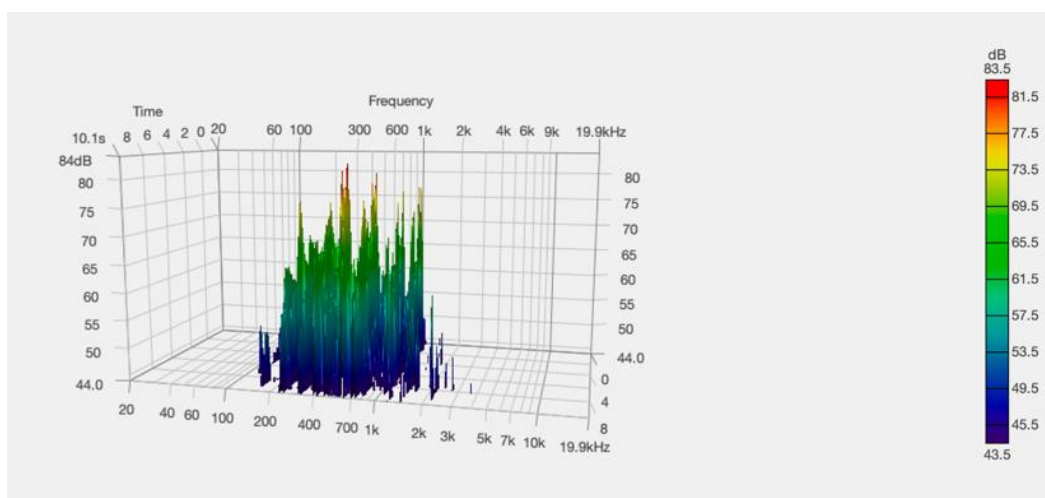
Στην παρούσα ενότητα περιγράφεται η ηχητική πηγή που χρησιμοποιήθηκε κατά την πειραματική διαδικασία, καθώς και τα χαρακτηριστικά των εκπεμπόμενων ήχων. Πρόκειται για ένα φορητό, μονοφωνικό ηχείο, το οποίο αναπαράγει ηχητικό σήμα λευκού θορύβου, καθώς και άλλα ακουστικά σήματα χαμηλών συχνοτήτων. Η συσκευή διαθέτει μέγιστη ισχύ εξόδου 3 Watt, ενώ το ηχείο καλύπτει το πλήρες ακουστικό φάσμα, από 20 Hz έως 20 kHz. Το εύρος αυτό καθιστά την ηχητική πηγή ιδανική για τη βέλτιστη αναπαραγωγή λευκού και ροζ θορύβου.



Εικόνα 4.53 Απεικόνιση της ηχητικής πηγής που χρησιμοποιήθηκε στην πειραματική διαδικασία.



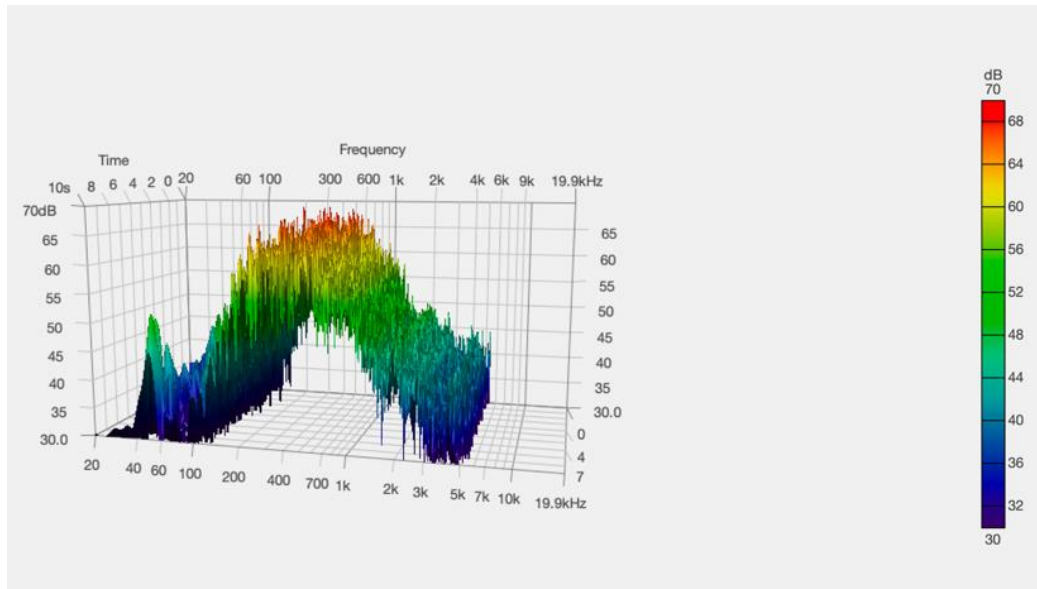
Εικόνα 4.54 Απεικόνιση του πρώτου ηχητικού σήματος στο λογισμικό Audacity.



Εικόνα 4.55 Απεικόνιση του φασματογραφήματος του πρώτου ηχητικού σήματος στο λογισμικό REW.



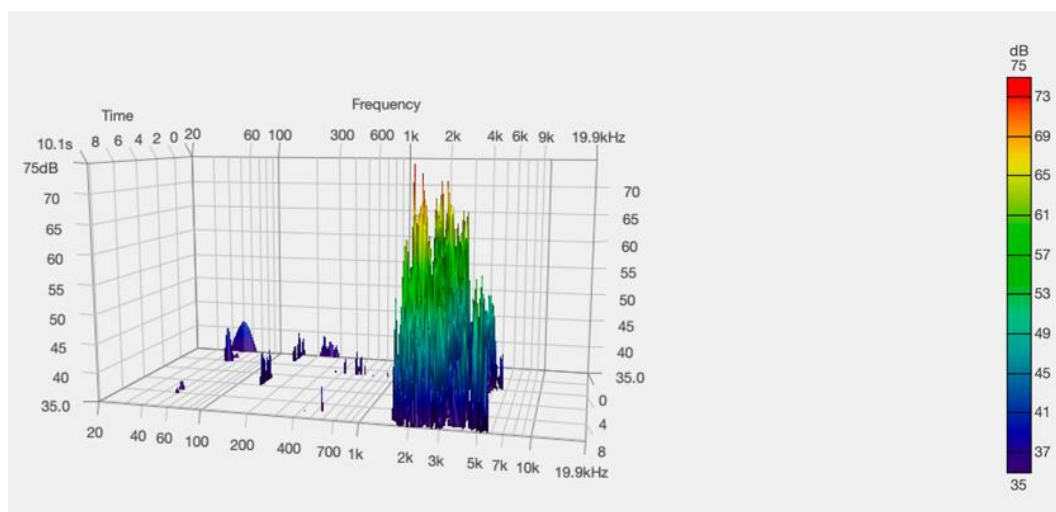
Εικόνα 4.56 Απεικόνιση του δεύτερου ηχητικού σήματος στο λογισμικό Audacity.



Εικόνα 4.57 Απεικόνιση του φασματογραφήματος του δεύτερου ηχητικού σήματος στο λογισμικό REW.



Εικόνα 4.58 Απεικόνιση του τρίτου ηχητικού σήματος στο λογισμικό Audacity.



Εικόνα 4.59 Απεικόνιση του φασματογραφήματος του τρίτου ηχητικού σήματος στο λογισμικό REW.

4.6 Παρουσίαση του χώρου διεξαγωγής του πειράματος

Ο χώρος στον οποίο διεξήχθη το πείραμα αποτελεί τμήμα σαλονιού κατοικίας, χωρίς ιδιαίτερες ακουστικές παρεμβάσεις. Το δάπεδο είναι ξύλινο, οι τοίχοι αποτελούνται από βαμμένο σοβά και η οροφή από βαμμένη γυψοσανίδα. Οι διαστάσεις του εν λόγω χώρου είναι 3,30 μέτρα μήκος επί 2,40 μέτρα πλάτος επί 2,40 μέτρα ύψος.

Ως εμπόδιο χρησιμοποιήθηκε μια στοίβα βιβλίων, με αποτέλεσμα το κύριο υλικό αυτού να είναι το χαρτί. Οι αποστάσεις μεταξύ του ρομποτικού οχήματος, του εμποδίου και της ηχητικής πηγής, καθώς και οι αντίστοιχες θέσεις τους στο χώρο, επιλέχθηκαν και μεταβάλλονταν τυχαία.

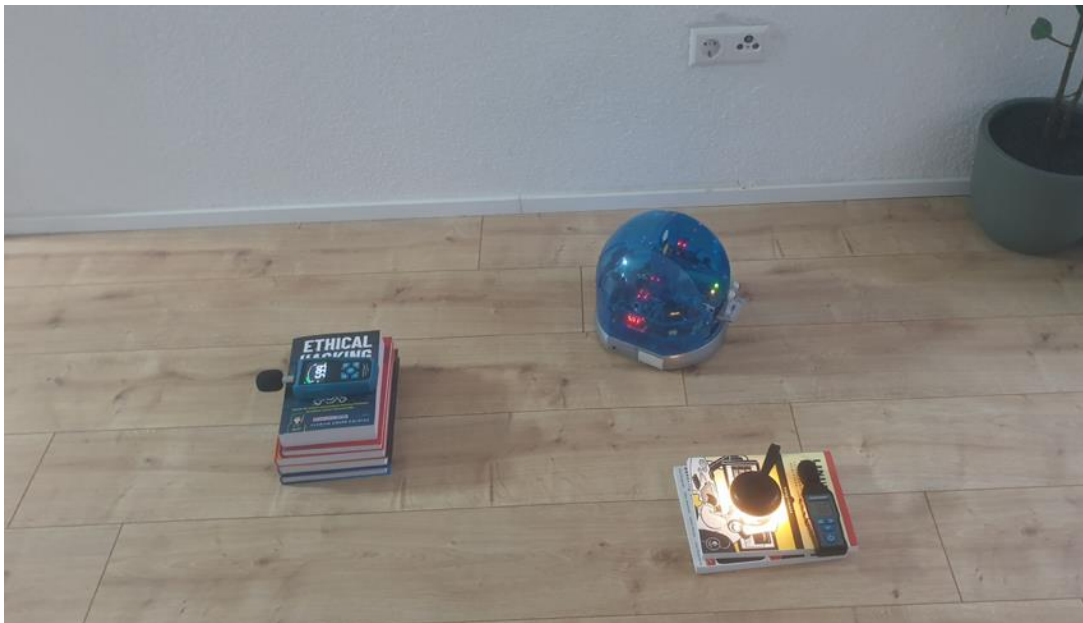
Στη συνέχεια παρατίθεται φωτογραφικό υλικό από τον χώρο κατά την διάρκεια της πειραματικής διαδικασίας.



Εικόνα 4.60 Απεικόνιση του ρομποτικού οχήματος κατά την εκκίνηση της λειτουργίας του.



Εικόνα 4.61 Απεικόνιση του ρομποτικού οχήματος κατά τον εντοπισμό εμποδίου στην διαδρομή του.



Εικόνα 4.62 Απεικόνιση του ρομποτικού οχήματος κατά τη φάση κίνησης του.

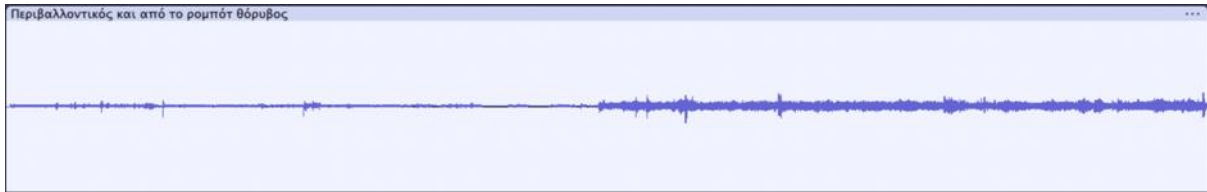


Εικόνα 4.63 Απεικόνιση του ρομποτικού οχήματος κατά τη φάση κίνησης του.

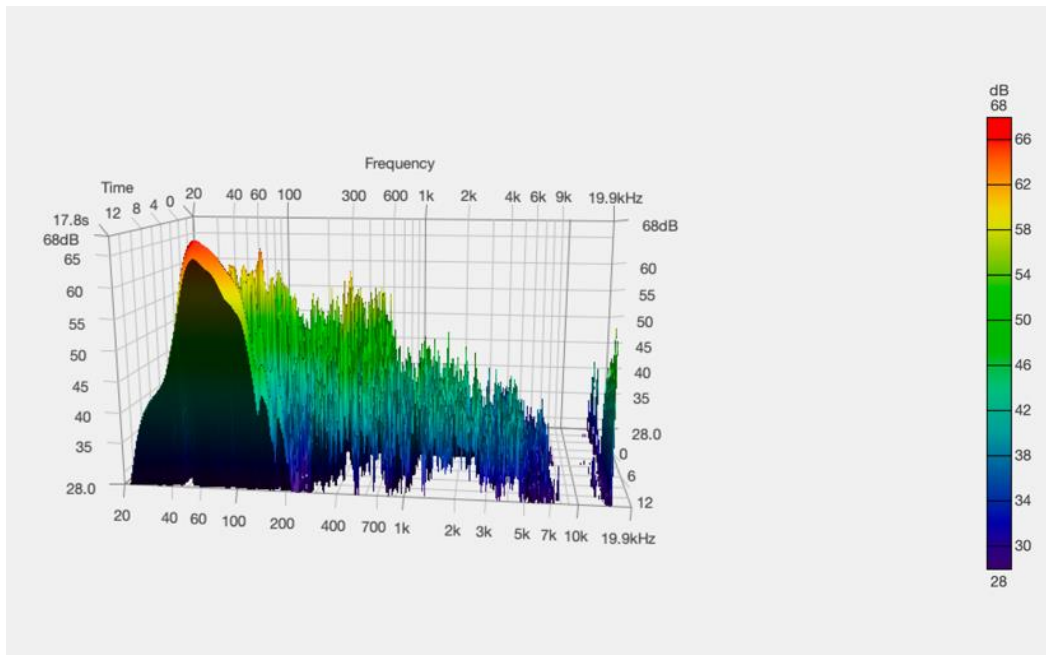


Εικόνα 4.64 Απεικόνιση του ρομποτικού οχήματος κατά τον εντοπισμό της ηχητικής πηγής.

Ακολούθως, αναλύεται ο θόρυβος περιβάλλοντος, καθώς και εκείνος που παράγεται από τα μηχανικά μέρη του ίδιου του ρομποτικού συστήματος. Για την αντιμετώπισή τους πραγματοποιήθηκε σχετική μελέτη και εφαρμόστηκαν φίλτρα άλφα. Η χρήση των φίλτρων κρίθηκε απαραίτητη για τον περιορισμό των παρεμβολών στη λειτουργία του οχήματος.



Εικόνα 4.65 Απεικόνιση των σημάτων περιβαλλοντικού και μηχανικού θορύβου του ρομποτικού οχήματος στο λογισμικό Audacity.



Εικόνα 4.66 Απεικόνιση του φασματογραφήματος των σημάτων περιβαλλοντικού και μηχανικού θορύβου του ρομποτικού οχήματος στο λογισμικό REW.

4.7 Μαθηματική μελέτη ηχητικού εντοπισμού

Η μαθηματική μελέτη για την εκτίμηση της θέσης της ηχητικής πηγής και την επακόλουθη κίνηση του ρομποτικού οχήματος προς αυτή, δύναται να διακριθεί σε τρία επιμέρους μέρη

1. Γεωμετρία του χώρου και του ρομποτικού οχήματος.
2. Ακουστική μοντελοποίηση.
3. Αλγόριθμος πλοήγησης

Στη συνέχεια, πραγματοποιείται μια συνοπτική ανάλυση για τα τρία αυτά μέρη.

1. Γεωμετρία του χώρου και του ίδιου του ρομποτικού οχήματος

Για την αναπαράσταση της γεωμετρίας του χώρου, ορίζεται ένα δισδιάστατο καρτεσιανό σύστημα συντεταγμένων (x,y). Το σύστημα αυτό περιλαμβάνει τον χώρο διεξαγωγής του πειράματος, καθώς

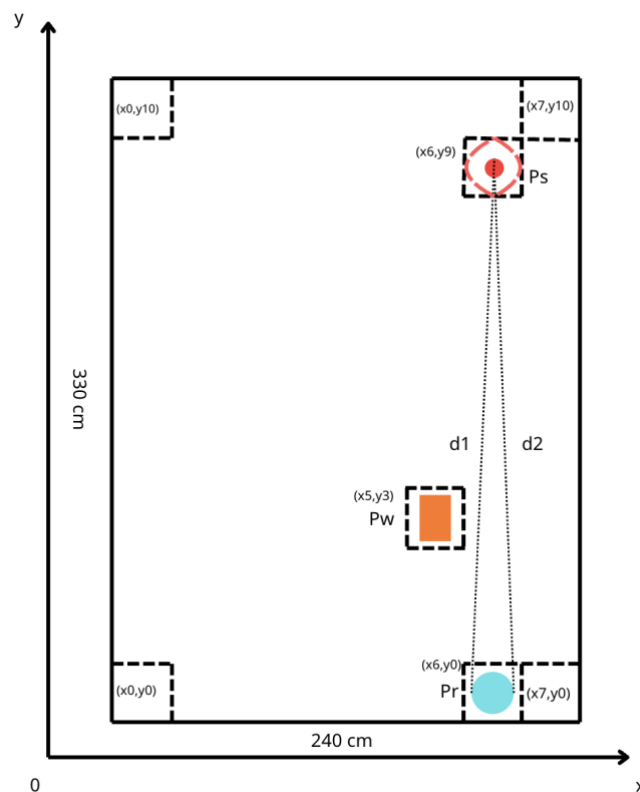
και τα επιμέρους στοιχεία που συμμετέχουν σε αυτό, όπως το ρομποτικό όχημα, το χωρικό εμπόδιο και την ηχητική πηγή.

Οι διαστάσεις και τα γεωμετρικά χαρακτηριστικά των στοιχείων του πειράματος ορίζονται στη συνέχεια ως εξής:

- Χώρος διεξαγωγής του πειράματος διαθέτει διαστάσεις 240 εκατοστά x 330 εκατοστά.
- Το ρομποτικό όχημα προσεγγίζεται ως κυκλικό σχήμα με διάμετρο 23 εκατοστά.
- Το εμπόδιο αποτελείται από διατεταγμένα βιβλία με διαστάσεις 16 εκατοστά x 24 εκατοστά.
- Η ηχητική πηγή έχει κυκλικό σχήμα με αρχική διάμετρο 10 εκατοστά.

Λόγω του ότι το ρομποτικό όχημα δεν μπορεί να προσεγγίσει την ηχητική πηγή σε απόσταση μικρότερη των 10 εκατοστών, κρίνεται απαραίτητο να συνυπολογιστεί ο συγκεκριμένος περιορισμός στη γεωμετρία του προβλήματος. Για το σκοπό αυτό, κατά τη θεωρητική ανάλυση και τη διεξαγωγή του πειράματος, η αρχική διάμετρος της πηγής επαυξάνεται περιμετρικά κατά 10 εκατοστά από κάθε πλευρά. Ως εκ τούτου, η ενεργός διάμετρος της ηχητικής πηγής διαμορφώνεται στα 30 εκατοστά.

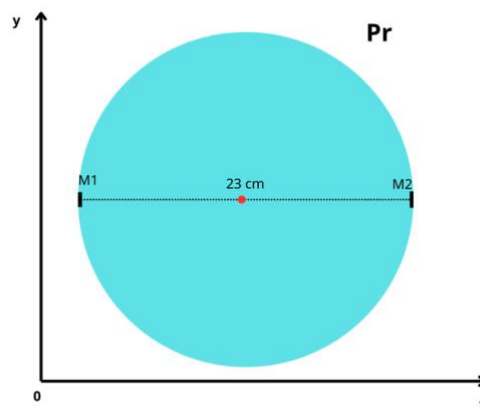
Ο χώρος διεξαγωγής του πειράματος έχει διαστάσεις 240 εκατοστά x 330 εκατοστά. Μεταξύ των τριών αντικειμένων που συμμετέχουν στη διαδικασία, το μεγαλύτερο μέγεθος, βάσει τις πειραματικές υποθέσεις, παρουσιάζει η ηχητική πηγή, η οποία έχει διάμετρο 30 εκατοστά. Εξαιτίας αυτού, ο συνολικός χώρος διαιρείται σε ένα πλέγμα τετραγωνικών κελιών, με την πλευρά κάθε τετραγώνου να ισούται με 30 εκατοστά αντιστοιχώντας στη διάμετρο της ηχητικής πηγής. Η συγκεκριμένη προσέγγιση οδηγεί στη δημιουργία περίπου 88 τετραγώνων. Με βάση τα δεδομένα αυτά, η γεωμετρία του χώρου αποτυπώνεται στο παρακάτω σχήμα.



Εικόνα 4.67 Απεικόνιση της κάτοψης του χώρου διεξαγωγής της πειραματικής διαδικασίας, με τον προσδιορισμό των ενδεικτικών θέσεων του ρομποτικού οχήματος, του εμποδίου και της ηχητικής πηγής.

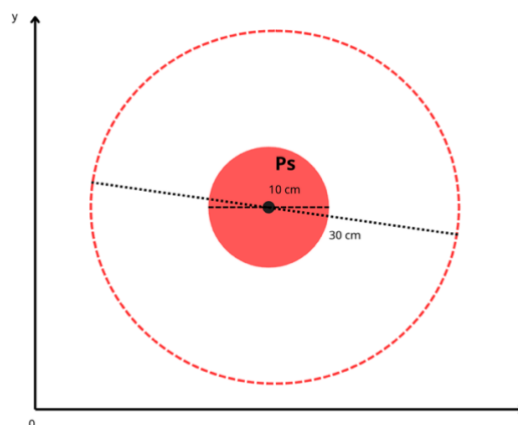
Κάθε ορισμένο τετράγωνο στον χώρο προσδιορίζεται μέσω των συντεταγμένων x και y . Το ρομποτικό όχημα βρίσκεται στη θέση (x_6, y_6) , το εμπόδιο στη θέση (x_5, y_5) και η ηχητική πηγή στη θέση (x_6, y_9) . Κατά την πειραματική διαδικασία, χρησιμοποιείται ένα πλέγμα 88 τετραγώνων, τα οποία αποτελούν τις δυνητικές θέσεις για τα τρία αντικείμενα. Από τις 88 αυτές θέσεις, κάθε αντικείμενο δύναται να καταλάβει μια από τις 86 τυχαίες θέσεις, καθώς δύο θέσεις είναι πάντοτε καλυμμένες από τα υπόλοιπα δύο αντικείμενα. Συνεπώς, κάθε αντικείμενο μπορεί να εντοπίζεται τυχαία σε μια από τις 86 ελεύθερες θέσεις του χώρου.

Το ρομποτικό όχημα διαθέτει ημισφαιρικό σχήμα και η δισδιάστατη προβολή του σχηματίζει κύκλο με διάμετρο 23 εκατοστά. Τα δύο μικρόφωνα, με τιμές M_1 και M_2 , είναι τοποθετημένα σε ίση απόσταση κατά μήκος της διαμέτρου του κύκλου. Τέλος, η μεταβλητή θέση του ρομπότ ορίζεται ως $P_r = (x_r, y_r)$, όπου $0 \leq x_r \leq 7$ και $0 \leq y_r \leq 10$.



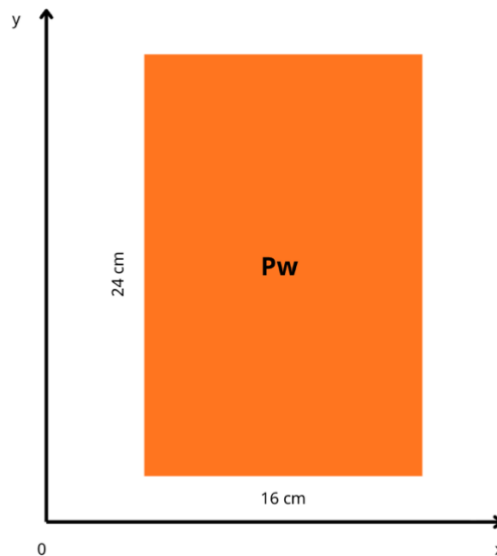
Εικόνα 4.68 Απεικόνιση της κάτοψης του ρομποτικού οχήματος.

Επιπλέον, η ηχητική πηγή διαθέτει ελλειψοειδές σχήμα, με την κάτοψη της στον χώρο να διαμορφώνει έναν πλήρη κύκλο διαμέτρου 10 εκατοστών. Στο πλαίσιο της παρούσας μελέτης, θεωρείται ότι κάθε σημείο του κύκλου επεκτείνεται κατά επιπλέον δέκα εκατοστά περιμετρικά, καθώς το ρομποτικό όχημα αδυνατεί να προσεγγίσει την πηγή σε απόσταση μικρότερη των 10 εκατοστών. Συνεπώς, για τις ανάγκες του πειράματος, η πηγή προσομοιώνεται ως ένας κύκλος με συνολική διάμετρο 30 εκατοστών. Η μεταβλητή θέση της ηχητικής πηγής ορίζεται ως $P_s = (x_s, y_s)$, όπου με το $0 \leq x_s \leq 7$ και $0 \leq y_s \leq 10$.



Εικόνα 4.69 Απεικόνιση της κάτοψης της ηχητικής πηγής.

Τέλος, ως εμπόδιο καθορίζεται μια στοίβα βιβλίων διαστάσεων 24 εκατοστών x 16 εκατοστών, η οποία προσδιορίζει την αντίστοιχη κάτοψη της στον χώρο. Η μεταβλητή θέση του εμποδίου ορίζεται ως $P_w = (x_w, y_w)$, όπου $0 \leq x_w \leq 7$ και $0 \leq y_w \leq 10$.



Εικόνα 4.70 Απεικόνιση της κάτοψης του εμποδίου.

2. Ακουστική Μοντελοποίηση

Στο παρόν τμήμα της εργασίας αναλύεται η λειτουργία του ρομποτικού οχήματος. Η πλοήγηση του συγκεκριμένου οχήματος βασίζεται στη στάθμη της ηχητικής πίεσης που ανιχνεύεται μέσω των δύο ενσωματωμένων μικροφώνων.

Η ακουστική ενέργεια E που λαμβάνει το κάθε μικρόφωνο μειώνεται αντιστρόφως ανάλογα με το τετράγωνο της απόστασης d από την ηχητική πηγή, σύμφωνα με τον νόμο του αντιστρόφου τετραγώνου. Θεωρώντας ότι η ηχητική πηγή εκπέμπει ενέργεια E_0 μια δεδομένη χρονική στιγμή, η ενέργεια που καταγράφεται στο μικρόφωνο M_1 και αντίστοιχα στο μικρόφωνο M_2 θα είναι

$$E_1 = \frac{E_0}{d_1^2} \quad (4.3)$$

$$E_2 = \frac{E_0}{d_2^2} \quad (4.4)$$

όπου

d_1 : η απόσταση του αριστερού μικρόφωνου από την ηχητική πηγή

d_2 : η απόσταση του δεξιού μικρόφωνου από την ηχητική πηγή

E_0 : η παραγόμενη ενέργεια από την ηχητική πηγή

Εξαιτίας της συνεχούς κίνησης του ρομποτικού οχήματος, η απόσταση μεταξύ της ηχητικής πηγής και του εκάστοτε μικρόφωνου μεταβάλλεται διαρκώς. Για τον λόγο αυτό, η απόσταση υπολογίζεται ανά πάσα στιγμή μέσω της διαφοράς των συντεταγμένων μεταξύ των αντίστοιχων σημείων.

Ειδικότερα, οι αποστάσεις των δύο μικροφώνων από την ηχητική πηγή ορίζονται από τις παρακάτω εξισώσεις

$$d_1 = |M_1 - P_s| \quad (4.5)$$

$$d_2 = |M_2 - P_s| \quad (4.6)$$

όπου

P_s : το κέντρο της ηχητικής πηγής

Στην περίπτωση, ωστόσο, που ανάμεσα στην ακουστική πηγή και στα αντίστοιχα μικρόφωνα παρεμβάλλεται κάποιο τεχνητό εμπόδιο, το οποίο προκαλεί απορρόφηση και εξασθένηση των ηχητικών κυμάτων, κρίνεται απαραίτητο να εισαχθεί στους υπολογισμούς ένας συντελεστής εξασθένησης $\alpha \in [0,1]$. Στο πλαίσιο αυτό, οι σχετικές εξισώσεις διαμορφώνονται ως ακολούθως

$$E_1 = \alpha \frac{E_0}{d_1^2} \quad (4.7)$$

$$E_2 = \alpha \frac{E_0}{d_2^2} \quad (4.8)$$

όπου

α : ο συντελεστής εξασθένησης

3. Αλγόριθμος πλοήγησης

Για την πλοήγηση του ρομποτικού οχήματος, το οποίο διαθέτει αποκλειστικά δύο αισθητήρες, λαμβάνονται αποφάσεις μέσω της σύγκρισης των επιπέδων της λαμβανόμενης ενέργειας E_1 και E_2 από το εκάστοτε μικρόφωνο.

Για τον προσδιορισμό της θέσης της ηχητικής πηγής, το ρομποτικό όχημα υπολογίζει τη διαφορά ενέργειας ως εξής

$$\Delta E = E_1 - E_2 \quad (4.9)$$

Με βάση την παραπάνω εξίσωση, διακρίνονται οι εξής περιπτώσεις

1. Αν $\Delta E > 0$, η πηγή ήχου βρίσκεται στη δεξιά πλευρά του ρομποτικού οχήματος, με αποτέλεσμα αυτό να στρίβει προς τα δεξιά.
2. Αν $\Delta E < 0$, η πηγή ήχου βρίσκεται στην αριστερή πλευρά του ρομποτικού οχήματος, με αποτέλεσμα αυτό να στρίβει προς τα αριστερά.
3. Αν $\Delta E = 0$, η πηγή ήχου εντοπίζεται είτε ακριβώς μπροστά είτε ακριβώς πίσω από το ρομποτικό όχημα. Στην περίπτωση αυτή ισχύει $d_1 = d_2$, που σημαίνει ότι η απόσταση των δύο μικροφώνων από την ηχητική πηγή είναι ίση.

Ακολούθως, προκειμένου να προσδιοριστεί η ενέργεια που καταγράφεται σε κάθε μικρόφωνο, εφαρμόζεται ο νόμος του αντιστρόφου τετραγώνου. Σύμφωνα με αυτόν, η ενέργεια που λαμβάνεται σε ελεύθερο πεδίο είναι αντιστρόφως ανάλογη του τετραγώνου της απόστασης από την πηγή, όπως παρουσιάζεται με τις αντίστοιχες εξισώσεις στη συνέχεια

$$E_1 = \frac{k}{d_1^2} \Rightarrow d_1^2 = \frac{k}{E_1} \quad (4.10)$$

$$E_2 = \frac{k}{d_2^2} \Rightarrow d_2^2 = \frac{k}{E_2} \quad (4.11)$$

όπου

$k = \frac{P_s}{4\pi}$: είναι μια σταθερά, η οποία εξαρτάται από την ισχύ της ηχητικής πηγής

Με βάση το Πυθαγόρειο Θεώρημα, οι αποστάσεις d_1 και d_2 συνδέονται με τις συντεταγμένες της ηχητικής πηγής (x_s, y_s) για κάθε μικρόφωνο ως εξής

Σχετικά με το αριστερό μικρόφωνο ισχύει η παρακάτω σχέση

$$d_1^2 = (x_s + R)^2 + y_s^2 = \frac{k}{E_1} \quad (4.12)$$

Αντίστοιχα, για το δεξί μικρόφωνο ισχύει η παρακάτω σχέση

$$d_2^2 = (x_s - R)^2 + y_s^2 = \frac{k}{E_2} \quad (4.13)$$

Ακολούθως, προκειμένου να προσδιοριστούν οι συντεταγμένες του x_s της ηχητικής πηγής στο αζιμουθιακό επίπεδο, πραγματοποιείται η πράξη της αφαίρεσης μεταξύ των δύο προηγούμενων εξισώσεων

$$\begin{aligned} d_1^2 - d_2^2 &= [(x_s + R)^2 + y_s^2] - [(x_s - R)^2 + y_s^2] \\ &= (x_s^2 + 2x_sR + R^2 + y_s^2) - (x_s^2 - 2x_sR + R^2 + y_s^2) = 4x_sR \end{aligned}$$

Συνεπώς, η συντεταγμένη x_s υπολογίζεται από τη παρακάτω σχέση

$$x_s = \frac{d_1^2 - d_2^2}{4R} \quad (4.14)$$

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Αντικαθιστώντας στην εξίσωση 4.14 τις εκφράσεις των d_1^2 και d_2^2 από τις εξισώσεις 4.12 και 4.13, προκύπτει η παρακάτω εξίσωση

$$x_s = \frac{k}{4R} \left(\frac{1}{E_1} - \frac{1}{E_2} \right) \quad (4.15)$$

Μέσω της προηγούμενης σχέσης καθορίζεται η χωρική θέση της ηχητικής πηγής ως προς το αζιμουθιακό επίπεδο.

Στη συνέχεια, μελετάται η παράλληλη διαδικασία προσδιορισμού της απόστασης μεταξύ του ρομποτικού οχήματος και του εμποδίου. Ως εμπόδιο τοποθετείται μια στοίβα βιβλίων, καθώς και η ίδια η ηχητική πηγή.

Για τον προσδιορισμό της απόστασης χρησιμοποιείται ένας αισθητήρας υπερήχων. Ο συγκεκριμένος αισθητήρας εκπέμπει ένα ηχητικό κύμα υψηλής συχνότητας και καταγράφει τον χρόνο πτήσης. Πρόκειται για το χρονικό διάστημα Δt_{us} που μεσολαβεί μέχρι το κύμα να φτάσει στο εμπόδιο και να επιστρέψει.

Η απόσταση d_{us} από τους αισθητήρες υπερήχων, οι οποίοι είναι εγκατεστημένοι στο μπροστινό τμήμα του ρομποτικού οχήματος, μέχρι το πλησιέστερο αντικείμενο, δύναται να υπολογιστεί από την παρακάτω εξίσωση

$$d_{us} = \frac{U_{sound} \Delta t_{us}}{2} \quad (4.16)$$

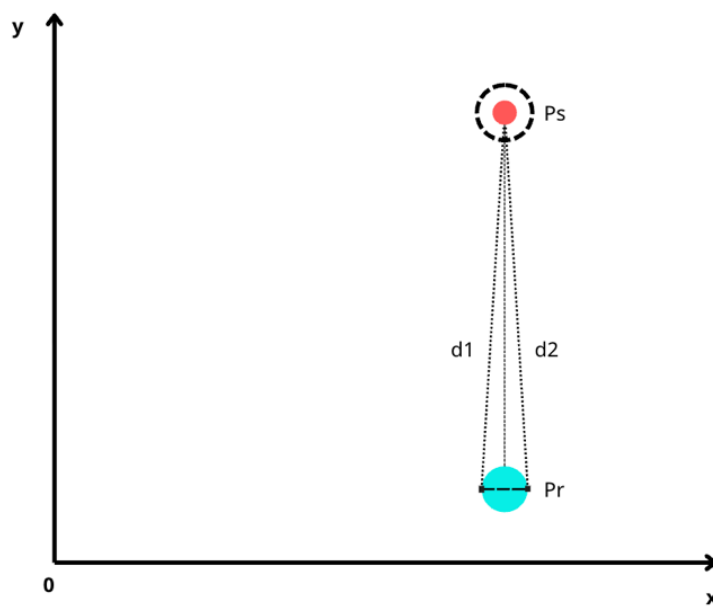
όπου

U_{sound} : η ταχύτητα του ήχου στο μέσο διάδοσης, της οποίας η τιμή εξαρτάται από τη θερμοκρασία T του μέσου και υπολογίζεται από την παρακάτω σχέση

$$U_{sound} = 331,3 + 0,606 T \quad (4.17)$$

Για τον χώρο διεξαγωγής του πειράματος και υπό την προϋπόθεση ιδανικών περιβαλλοντικών συνθηκών, με θερμοκρασία στους 25 °C, η ταχύτητα του ήχου ισούται με $U_{sound} = 346,45$ m/s.

Συμπερασματικά, το ρομποτικό όχημα καθορίζει την πορεία του με βάση τη διαφορά ενέργειας μεταξύ των μικροφώνων, μέχρι όπου $E_o = 0$, προκειμένου να κινείται σε ευθεία γραμμή με στόχο την ηχητική πηγή. Παράλληλα, ο αισθητήρας υπερήχων επιτηρεί διαρκώς την απόσταση από τον στόχο και όταν η απόσταση γίνει ίση ή μικρότερη από 10 εκατοστά, $d_{us} \leq 10$ cm, το ρομποτικό όχημα ακινητοποιείται.



Εικόνα 4.71 Απεικόνιση του ρομποτικού σχήματος και των ενδεικτικών αποστάσεων των δύο μικροφώνων από το κέντρο της ηχητικής πηγής.

4.8 Σύνοψη 4^{ου} κεφαλαίου

Στο παρόν κεφάλαιο παρουσιάστηκε η διαδικασία σχεδιασμού, ανάπτυξης και υλοποίησης του ρομποτικού οχήματος. Αρχικά, πραγματοποιήθηκε αναλυτική παρουσίαση των επιμέρους ηλεκτρονικών και μηχανικών τμημάτων του, προσδιορίζοντας με ακρίβεια το λειτουργικό σκοπό αλλά και τη σκοπιμότητα ενσωμάτωσης του κάθε τμήματος στη τελικά κατασκευή. Επιπλέον, πραγματοποιήθηκε εκτενής αναφορά στη μελέτη του ενεργειακού τομέα και της διαχείρισης ισχύος, καθώς και στα άλλα τμήματα του ρομποτικού οχήματος. Ιδιαίτερη έμφαση δόθηκε στη μελέτη του υλικού που αφορά το ακουστικό κομμάτι. Στο πλαίσιο αυτό, εξετάστηκαν εκτενώς οι προδιαγραφές για την τοποθέτηση των μικροφώνων. Επίσης, παρουσιάστηκε το σύνολο του εξοπλισμού και αναλύθηκε η μεθοδολογία για την σύνδεση των περιφερειακών τμημάτων, των οποίων η συνδυαστική λειτουργία οδήγησε στη δημιουργία ενός πλήρως λειτουργικού ρομποτικού οχήματος. Παράλληλα με το υλικό μέρος, αναπτύχθηκε το λογισμικό που ελέγχει και συντονίζει τις διάφορες λειτουργίες που πραγματοποιεί το ρομποτικό όχημα. Παρουσιάστηκαν αναλυτικά όλα τα επιμέρους δομικά τμήματα του προγράμματος, επεξηγώντας τον σκοπό που επιτελούν για την πλήρη λειτουργία του. Στο σημείο αυτό, δόθηκε ιδιαίτερη βαρύτητα στα τμήματα του προγράμματος που είναι υπεύθυνα για τις ακουστικές λειτουργίες του ρομποτικού οχήματος.

Επιπλέον, προσδιορίστηκαν τα τεχνικά χαρακτηριστικά της ηχητικής πηγής που χρησιμοποιήθηκε, καθώς και το φάσμα των εκπεμπόμενων ηχητικών σημάτων της, ενώ παρατέθηκε μια λεπτομερής περιγραφή του χώρου, με εστίαση στα ακουστικά χαρακτηριστικά του.

Συνοψίζοντας, στο κεφάλαιο αυτό πραγματοποιήθηκε μια σφαιρική και ολοκληρωμένη ανασκόπηση της μελέτης, της σχεδίασης, της κατασκευής και του προγραμματισμού του ρομποτικού οχήματος. Επιπλέον, εξετάστηκαν διεξοδικά οι εσωτερικοί και οι εξωτερικοί παράγοντες που δύναται να επηρεάσουν τη σταθερότητα, την απόκριση και τη λειτουργικότητα του ρομποτικού οχήματος κατά την φάση πλοήγησής του.

Κεφάλαιο 5

5. Αξιολόγηση

5.1 Γενικά

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη των τεχνικών χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές, καθώς και η μελέτη και κατασκευή ενός ρομποτικού οχήματος, η λειτουργία του οποίου βασίζεται στην μέθοδο της διαφοράς ενέργειας του ηχητικού σήματος μεταξύ δύο μικροφώνων. Για την υλοποίηση της εργασίας ακολουθήθηκαν τέσσερα βασικά στάδια: η ανάλυση, η σχεδίαση, η ανάπτυξη και η αξιολόγηση. Αρχικά πραγματοποιήθηκε η φάση της ανάλυσης, την οποία διαδέχθηκαν η σχεδίαση και η ανάπτυξη. Στη συνέχεια, η διαδικασία ολοκληρώθηκε με την αξιολόγηση, προκειμένου να εξαχθούν τα προσδοκώμενα συμπεράσματα από την εφαρμογή και τη λειτουργία αλλά και τα αποτελέσματα από την υλοποίηση της ρομποτικής εφαρμογής. Επιπλέον, αναδεικνύονται οι περιορισμοί και οι δυσκολίες που προέκυψαν κατά τη διαδικασία της ανάπτυξης. Τέλος, διατυπώνονται συγκεκριμένες προτάσεις για την μελλοντική εξέλιξη της εφαρμογής.

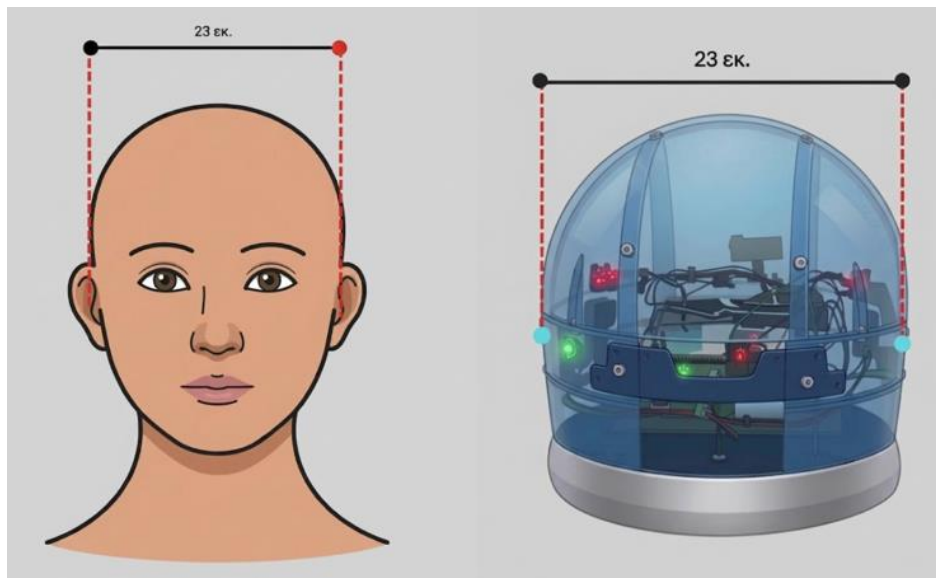
5.2 Αξιολόγηση αποτελεσμάτων

Σε αυτό το κεφάλαιο της εργασίας πραγματοποιείται η αξιολόγηση του ρομποτικού οχήματος, η οποία αποτελεί το τελευταίο στάδιο του δημιουργικού μέρους της εργασίας. Η διαδικασία της αξιολόγησης του ρομποτικού οχήματος και γενικά του συνόλου της εργασίας είναι καθοριστική για την εξαγωγή και την ανάδειξη των τελικών συμπερασμάτων. Συγκεκριμένα, εξετάζεται ο βαθμός στον οποίο το ρομποτικό όχημα πέτυχε τους αρχικούς στόχους. Επίσης, αναλύονται τα διάφορα ζητήματα που προέκυψαν κατά την διαδικασία της ανάπτυξης και της υλοποίησης, ενώ παράλληλα γίνεται παρουσίαση των τρόπων αντιμετώπισης τους. Η διαδικασία της αξιολόγησης έλαβε χώρα σε προκαθορισμένο περιβάλλον, με καθορισμένες ακουστικές και περιβαλλοντικές συνθήκες.

5.2.1 Αξιολόγηση των στόχων ανάπτυξης

Η βάση στην οποία βασίστηκε η παρούσα εργασία ήταν η κατασκευή ενός ρομποτικού οχήματος, το οποίο σε θεωρητικό επίπεδο εξομοιώνει τη χωρική ακουστική ικανότητα της ανθρώπινης κεφαλής. Για την επίτευξη αυτού του στόχου, χρησιμοποιήθηκε ο σκελετός και οι τροχοί με τους κινητήρες περιστροφής τους, από ένα ρομποτικό όχημα της εταιρείας De Agostini.

Στη συνέχεια, τοποθετήθηκαν τα δύο μικρόφωνα στο ίδιο ύψος από το έδαφος και η μεταξύ τους απόσταση μεταξύ τους, η οποία καθορίστηκε από τις διαστάσεις του ρομποτικού οχήματος ήταν 23 εκατοστά, προσεγγίζοντας τη μέση απόσταση μεταξύ των αυτιών του ανθρώπου. Με το τρόπο αυτό επιτεύχθηκε περίπου σε γενικές γραμμές η προσομοίωση της ανθρώπινης κεφαλής, το οποίο ήταν και ο αρχικός στόχος.



Εικόνα 5.1 Παρουσίαση σχηματικού της απόστασης μεταξύ των δύο αυτιών στο ανθρώπινο κεφάλι και της απόστασης μεταξύ των μικροφώνων στο ρομποτικό όχημα.

Στην συνέχεια, πραγματοποιήθηκε μελέτη και υλοποίηση για την τοποθέτηση των δύο μικροφώνων, ώστε θεωρητικά να δημιουργηθούν οι συναρτήσεις μεταφορά για την κεφαλή HRTF στο ρομποτικό όχημα. Μέσω της διαφοράς ενέργειας του ηχητικού σήματος που φτάνει σε κάθε μικρόφωνο, να επιτυγχάνεται ο εντοπισμός της ηχητικής πηγής.

Τα μικρόφωνα που χρησιμοποιήθηκαν ήταν τα KY-037, τα οποία αποτελούν μια αρκετά αξιόπιστη και οικονομική επιλογή. Διαθέτουν ρυθμιστή ευαισθησίας μέσω ενός ενσωματωμένου ποτενσιόμετρου, ωστόσο επιλέχθηκε εξ αρχής μια μέση τιμή ευαισθησίας και για τα δύο. Μετά την ολοκλήρωση της κατασκευής του ρομποτικού οχήματος, δεν υπήρχε δυνατότητα πρόσβασης για περαιτέρω ρύθμιση της ευαισθησίας. Εξετάστηκε επίσης η δυνατότητα χρήσης των μικροφώνων MAX9814, τα οποία είναι αρκετά πιο αξιόπιστα, αλλά η επιλογή αυτή απορρίφθηκε λόγω δυσκολιών στην τοποθέτηση.

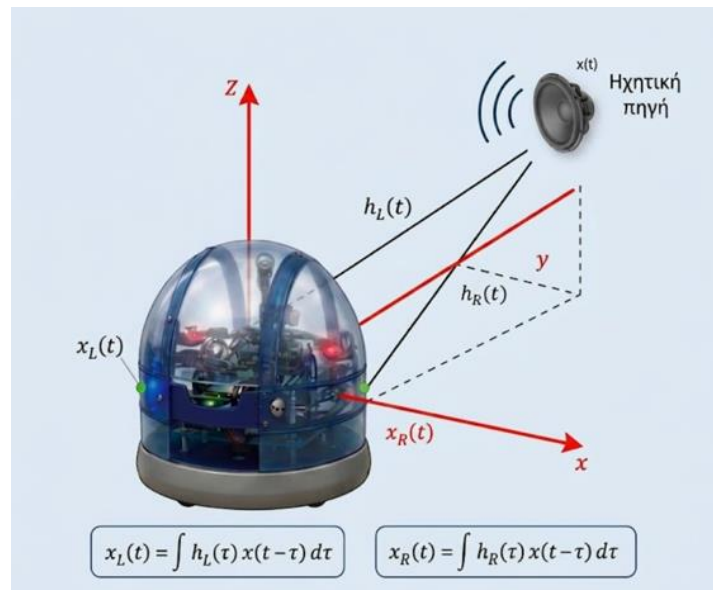
Τα δύο αυτά μικρόφωνα παρουσιάζουν σημαντικές διαφορές. Το μικρόφωνο KY-037 αποτελεί έναν απλό ανιχνευτή ήχου, κατάλληλο κυρίως για ανίχνευση θορύβου, όπως τα χειροκροτήματα. Αντίθετα, το μικρόφωνο MAX9814 διαθέτει ενσωματωμένο ενισχυτή με αυτόματο έλεγχο κέρδους, γεγονός που το καθιστά κατάλληλο για καταγραφή φωνής αλλά και ήχου. Στην συνέχεια, παρατίθεται ένας πίνακας σύγκρισης των δύο μικροφώνων που δοκιμάστηκαν στο ρομποτικό όχημα.

α/α	Μικροφώνο ΚΥ-037	Μικρόφωνο MAX9814
1	Ανίχνευση κατωφλίου θορύβου	Καταγραφή ήχου και φωνής
2	Ενίσχυση (Μέσω ποτενσιόμετρου)	Ενίσχυση (Αυτόματα)
3	Χαμηλή ποιότητα σήματος	Υψηλή ποιότητα με χαμηλό θόρυβο
4	Ψηφιακή και αναλογική έξοδο	Αναλογική έξοδο
5	Οικονομικό μικρόφωνο	Οικονομικό μικρόφωνο

Πίνακας 5.1 Παρουσίαση των χαρακτηριστικών των δύο μικροφώνων.

Σε γενικές γραμμές, το μικρόφωνο MAX9814 κρίνεται ως μια αρκετά κατάλληλη επιλογή για χρήση στο ρομποτικό όχημα. Ωστόσο, δεν χρησιμοποιήθηκε στην κατασκευή διότι υπήρχε αδυναμία τοποθέτησης. Το συγκεκριμένο μικρόφωνο θα επέτρεπε στο ρομποτικό όχημα να εντοπίζει την ηχητική πηγή σε μεγαλύτερη απόσταση, προσφέροντας παράλληλα υψηλότερο λόγο σήματος προς θόρυβο. Επιπλέον, για την ενσωμάτωση του θα απαιτούσε τροποποίηση του υπάρχοντος κυκλώματος, με την προσθήκη ενός αναλογικό σε ψηφιακό μετατροπέα (ADC) όπως πχ. το ADC ADS1015. Συνεπώς, οι δύο βασικοί λόγοι που απορρίφθηκε η χρήση του μικρόφωνου MAX9814 είναι η αδυναμία προσαρμογής τους στην κατασκευή και η ανάγκη ανασχεδιασμού του κυκλώματος με την προσθήκη ενός επιπλέον εξωτερικού αναλογικού σε ψηφιακό μετατροπέα.

Στην συνέχεια, παρουσιάζεται σχηματικά ο υπολογισμός της εκτιμώμενης θέσης της ηχητικής πηγής. Από το σύνολο των μεθόδων που αναλύθηκαν στο θεωρητικό μέρος της εργασίας, η μόνη μέθοδος που μπορούσε να εφαρμοστεί και να αποδώσει ικανοποιητικά αποτελέσματα ήταν η μέθοδος εντοπισμού με βάση τη διαφορά ενέργειας του ηχητικού σήματος που φτάνει στα δύο μικρόφωνα. Στην περίπτωση που επιλεγόταν η μέθοδος της διαφοράς χρόνου άφιξης (TDOA), ο υπολογισμός δεν θα ήταν εφικτός. Αυτό οφείλεται στο γεγονός ότι τα χρησιμοποιούμενα μικρόφωνα ανιχνεύουν αποκλειστικά το επίπεδο του θορύβου, επιτρέποντας μόνο την πραγματοποίηση του εντοπισμού με βάση τη ενέργεια του σήματος. Αντίθετα, με τη χρήση του μικροφώνου MAX9814 η εφαρμογή της μεθόδου της διαφοράς χρόνου άφιξης θα ήταν εφικτή. Το συγκεκριμένο μικρόφωνο αναπαράγει με ακρίβεια το ηχητικό σήμα που λαμβάνει στην είσοδο του, παρέχοντας τη δυνατότητα της εφαρμογής μαθηματικών μεθόδων όπως πχ. η ετεροσυσχέτιση, για τον εντοπισμό της διαφοράς φάσης του ηχητικού σήματος.



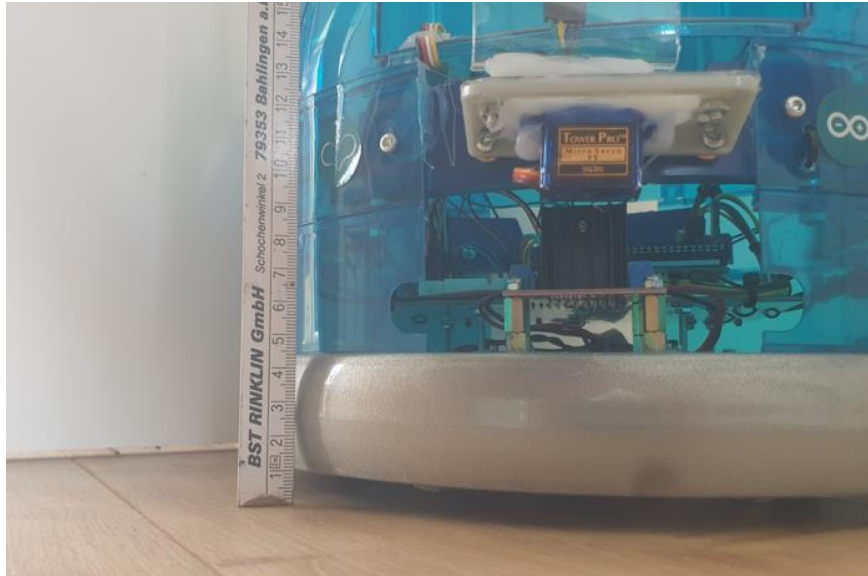
Εικόνα 5.2 Παρουσίαση γραφικά τις απόστασης των δύο μικροφώνων από την ηχητική πηγή.

Ένα εναλλακτικό αρχικό σχέδιο προέβλεπε τη χρήση του ρομποτικού οχήματος για τη διεξαγωγή δοκιμών με εναλλακτικά επιφανειακά υλικά, τα οποία θα διέθεταν διαφορετικό συντελεστή ηχοαπορρόφησης. Οι δοκιμές αυτές θα λάμβαναν χώρα σε διαφορετικές περιβαλλοντικές συνθήκες και επίπεδα στάθμης του θορύβου. Συγκεκριμένα, οι δοκιμές θα πραγματοποιούνταν σε συνθήκες θερμοκρασίας από -12 °C κατά τους μήνες Δεκέμβριο και Ιανουάριο έως και 32 °C τον Μάιο και τον Ιούνιο, με αντίστοιχα επίπεδα υγρασίας περίπου στο 90% έως τον Μάρτιο και 5% κατά τους μήνες Μάιο και Ιούνιο. Ο απαιτούμενος θόρυβος θα αναπαραγόταν τεχνικά με τη χρήση ηχείου και φορητού υπολογιστή. Στην συνέχεια, θα δημιουργούταν μια βάση δεδομένων για την αξιολόγηση του βαθμού στον οποίο τα διαφορετικά υλικά και οι μεταβολές των συνθηκών επηρέαζαν τον εντοπισμό της ηχητικής πηγής, ενώ παράλληλα θα πραγματοποιούνταν δοκιμές σε διάφορες αποστάσεις. Ωστόσο, αυτό το σχέδιο εγκαταλείφθηκε λόγω σημαντικής καθυστέρησης της ανάπτυξης και της υλοποίησης του ρομποτικού οχήματος και δεν ήταν εφικτό να δημιουργηθούν μεγάλες διακυμάνσεις στις περιβαλλοντικές συνθήκες. Ωστόσο, παρατηρήθηκε ότι οι μικρές μεταβολές των περιβαλλοντικών συνθηκών δεν προκαλούσαν αξιοσημείωτη διαφορά στον εντοπισμό της ηχητικής πηγής. Επιπλέον, η αλλαγή των επιφανειακών υλικών επέφερε σημαντικό πρόσθετο βάρος στο ρομποτικό όχημα, επηρεάζοντας την λειτουργία του, γεγονός που καθιστούσε αναγκαίο τον επανασχεδιασμό του συστήματος τροφοδοσίας του.

Στην συνέχεια, εξετάστηκαν οι προτεινόμενοι χώροι για τη διεξαγωγή των δοκιμών, οι οποίες δοκιμές θα ήταν δυνατόν να πραγματοποιηθούν είτε σε εξωτερικό είτε σε εσωτερικό περιβάλλον. Αρχικά, επιχειρήθηκε η εκτέλεση του πειράματος σε εξωτερικό χώρο, ωστόσο η επιλογή αυτή κρίθηκε μη βιώσιμη για δύο βασικούς λόγους.

Ο πρώτος λόγος σχετίζεται με το ακουστικό περιβάλλον. Τα επίπεδα του περιβαλλοντικού θορύβου ήταν ίσα ή και υψηλότερα από τη στάθμη ακουστικής πίεσης που παρήγαγε η ηχητική πηγή, δυσχεραίνοντας σημαντικά τον εντοπισμό της. Η μόνη λύση σε αυτό το ζήτημα ήταν η τοποθέτηση του ρομποτικού οχήματος σε απόσταση μικρότερη από 120 εκατοστών από την ηχητική πηγή, όπου η σχετική μείωση του περιβαλλοντικού θορύβου επέτρεπε τον επιτυχή εντοπισμό της. Επιπλέον, ο περιβαλλοντικός θόρυβος παρουσίαζε τυχαίες μεταβολές και μη γραμμικότητα, καθιστώντας αδύνατον τον έλεγχο του. Παρά τη προσπάθεια να γίνει έλεγχος μέσω σχεδιασμού και εφαρμογής ψηφιακών φίλτρων, ωστόσο τα αποτελέσματα δεν υπήρξαν ικανοποιητικά.

Ο δεύτερος λόγος δεν σχετίζεται με την ακουστική λειτουργία, αλλά με την κατασκευή του ρομποτικού οχήματος. Λόγω της μικρής απόστασης μεταξύ του σασί και του εδάφους, ήταν εξαιρετικά δύσκολο να εντοπιστεί εξωτερικός χώρος που να επιτρέπει την απρόσκοπτη κίνηση του ρομποτικού οχήματος. Παράλληλα, εμφανίστηκε ενεργειακό πρόβλημα, καθώς το σύστημα τροφοδοσίας δεν διέθετε την απαιτούμενη ισχύ για την κίνηση των κινητήρων σε μη επίπεδες επιφάνειες. Επιπλέον, πραγματοποιήθηκε μια απόπειρα διεξαγωγής της πειραματικής διαδικασίας στο εθνικό στάδιο της πόλης, η οποία όμως δεν τελεσφόρησε.



Εικόνα 5.3 Παρουσίαση της απεικόνισης της απόστασης του σασί του ρομποτικού οχήματος από το έδαφος.

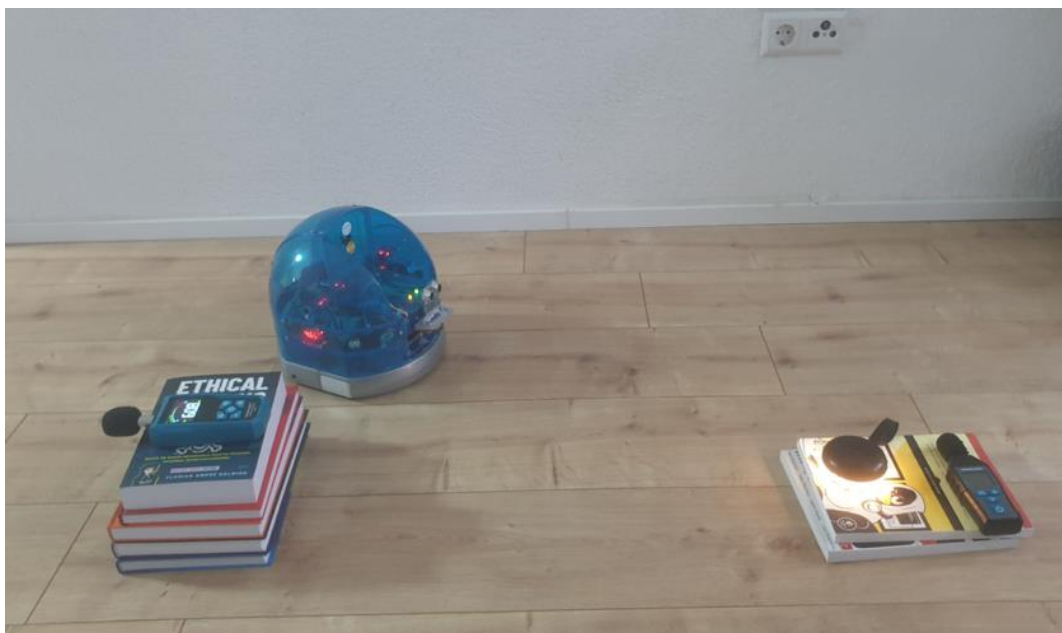


Εικόνα 5.5 Παρουσίαση στιγμιότυπου από την προσπάθεια διεξαγωγής του πειράματος σε εξωτερικό χώρο.



Εικόνα 5.6 Παρουσίαση στιγμιότυπου από την προσπάθεια διεξαγωγής του πειράματος σε εξωτερικό χώρο.

Λόγω της έλλειψης πρόσβασης σε εσωτερικό χώρο μεγάλων διαστάσεων, η πειραματική διαδικασία διεξήχθη σε εσωτερικό χώρο οικίας και συγκεκριμένα σε καθορισμένο τμήμα σαλονιού. Εξαιτίας των περιορισμένων διαστάσεων του χώρου, δεν κατέστη εφικτό να ληφθούν υπόψη συγκεκριμένες αποστάσεις μεταξύ του ρομποτικού οχήματος, του εμποδίου και της ηχητικής πηγής. Κατά συνέπεια, η επιλογή των αποστάσεων πραγματοποιήθηκε με τυχαίο τρόπο. Παράλληλα, ο περιβαλλοντικός θόρυβος κυμαινόταν σε επίπεδα που δεν επηρέαζαν τη διεξαγωγή του πειράματος. Επιπλέον, επισημαίνεται ότι η πειραματική διαδικασία διεξαγόταν πάντα με τη χρήση μόνο ενός εμποδίου.



Εικόνα 5.7 Παρουσίαση στιγμιότυπου από τη διεξαγωγή του πειράματος σε εσωτερικό χώρο.

Στην συνέχεια, εξετάζονται τα διάφορα σενάρια σχετικά με τον εντοπισμό της ηχητικής πηγής. Το πρώτο σενάριο προβλέπει ότι τόσο το ρομποτικό όχημα όσο και η ηχητική πηγή παραμένουν σταθερά. Στην περίπτωση αυτή, μόλις η ηχητική πηγή έκπεμπε κάποιον ηχητικό σήμα, το ρομποτικό όχημα την εντοπίζει και κινείται προς το μέρος της. Το δεύτερο σενάριο περιλαμβάνει ότι το ρομποτικό όχημα βρίσκεται σε κίνηση και μόλις λάβει το ηχητικό σήμα κατευθύνεται προς την σταθερή ηχητική πηγή. Το τρίτο σενάριο προβλέπει σταθερό ρομποτικό όχημα και κινούμενη ηχητική πηγή, με το ρομποτικό όχημα να σταματά ανά χρονικές περιόδους προκειμένου να επαναπροσδιορίσει τη θέση της. Το τέταρτο και τελευταίο σενάριο περιλαμβάνει την ταυτόχρονη κίνηση τόσο του ρομποτικού οχήματος όσο και της ηχητικής πηγής. Ακολουθεί συνοπτικός πίνακας με την παρουσίαση των τεσσάρων σεναρίων.

α/α	Ρομποτικό όχημα	Ηχητική πηγή	Τρόπος εντοπισμού
1	Ακίνητο	Ακίνητη	Σταθερός εντοπισμός στόχου
2	Κινούμενο	Ακίνητη	Σταθερός εντοπισμός στόχου
3	Ακίνητο	Κινούμενη	Συνεχής ανίχνευση ηχητικού σήματος
4	Κινούμενο	Κινούμενη	Συνεχής ανίχνευση ηχητικού σήματος

Πίνακας 5.2 Παρουσίαση των σεναρίων διεξαγωγής του πειράματος.

Αρχικά, εξετάζεται η περίπτωση της κινούμενης ηχητικής πηγής. Για την υλοποίηση αυτού του σεναρίου, θα απαιτούνταν η χρήση ενός δεύτερου ρομποτικού οχήματος, το οποίο να παράγαγε ήχο επαρκούς στάθμης ακουστικής πίεσης ώστε να ανιχνεύεται από τα μικρόφωνα. Ο χειρισμός του δεύτερου ρομποτικού οχήματος θα μπορούσε να πραγματοποιηθεί είτε χειροκίνητα μέσω Bluetooth είτε αυτόνομα, βάση προγραμματισμού για κίνηση στο χώρο. Μια εναλλακτική προσέγγιση θα μπορούσε να προβλέπει τη χωρική αναπαραγωγή του ήχου μέσω ενός συστήματος ηχείων, το οποίο με την εφαρμογή κατάλληλων ακουστικών εφέ θα προσομοίωνε κινούμενες ηχητικές πηγές. Επιπλέον μια τρίτη προσέγγιση, θα ήταν η χειροκίνητη μετακίνηση της ηχητικής πηγής, διαδικασία που προϋποθέτει την ανθρώπινη παρουσία και συμμετοχή στην διαδικασία διεξαγωγής του πειράματος. Πέραν των ανωτέρων περιορισμών ο εντοπισμός μια κινούμενης ηχητικής πηγής παρουσίαζε επιπλέον τεχνικές δυσκολίες. Συγκεκριμένα, η χρήση μόνο δύο μικρόφωνων, τα τεχνικά χαρακτηριστικά των μικροφώνων καθώς και η χαμηλή ταχύτητα επεξεργασίας από τον μικροελεγκτή της πλατφόρμας του Arduino καθιστούσαν την διαδικασία ανέφικτη. Για την επιτυχή υλοποίηση του συγκεκριμένου σεναρίου, θα απαιτούνταν η ενσωμάτωση τεσσάρων μικρόφωνων, όπως τα MAX9814, σε συνδυασμό με ένα εξωτερικό μετατροπέα αναλογικού σε ψηφιακό σήμα όπως για παράδειγμα ο ADS1115. Η προσθήκη αυτή θα επέτρεπε τον ακριβή χωρικό εντοπισμό του ηχητικού σήματος καθώς και τη ταχύτερη επεξεργασία των δεδομένων μέσω του αναλογικού σε ψηφιακό μετατροπέα.

Στην συνέχεια, εξετάζεται ο εντοπισμός της ηχητικής πηγής από κινούμενο ρομποτικό όχημα. Το κύριο πρόβλημα εντοπίστηκε στα υψηλά επίπεδα θορύβου που προκαλούσε η κίνηση του ίδιου του ρομποτικού οχήματος, τα οποία ήταν συγκρίσιμα με τη στάθμη ακουστικής πίεσης της ηχητικής πηγής. Στην περίπτωση αυτή, μια πιθανή λύση αποτελεί η τοποθέτηση των κινητήρων του ρομποτικού οχήματος σε ηχομονωτικούς κλωβούς, με σκοπό τη μείωση του παραγόμενου θορύβου. Επιπλέον, προτείνεται η αύξηση της απόστασης μεταξύ κινητήρων και μικροφώνων ή και η εφαρμογή ακουστικής μόνωσης με χρήση αφρώδους υλικού ανάμεσα στην βάση του μικροφώνου και το κελύφους του ρομποτικού οχήματος. Εναλλακτικά, προτείνεται η χρήση τεσσάρων μικροφώνων, όπως τα MAX9814, σε συνδυασμό με ένα μετατροπέα αναλογικού σε ψηφιακό σήμα όπως για παράδειγμα το ADS1115. Με την αλλαγή αυτή στο κύκλωμα διασφαλίζεται ο ακριβής χωρικός

εντοπισμός του ηχητικού σήματος και η ταχύτερη επεξεργασία των δεδομένων, επιτρέποντας τον αποτελεσματικό διαχωρισμό του θορύβου του ρομποτικού οχήματος από το σήμα της ηχητικής πηγής.

Η τελευταία μέθοδος, η οποία τελικά υλοποιήθηκε, αφορά τον εντοπισμό ακίνητης ηχητικής πηγής από ακίνητο ρομποτικό όχημα. Στην περίπτωση αυτή, δεν παρουσιάστηκαν τα προβλήματα των άλλων μεθόδων, ενώ η υλοποίηση πραγματοποιήθηκε με χρήση δύο μικροφώνων, χωρίς την ανάγκη πρόσθετου εξοπλισμού ή τροποποίησης της αρχικής σχεδίασης. Συγκεκριμένα το ρομποτικό όχημα χωρίς να κινείται, ανέμενε μέχρι τη λήψη του ηχητικού σήματος, μόλις το λάμβανε, κλείδωνε τον στόχο και ξεκινούσε την πλοήγηση προς αυτόν.

Ως αποτέλεσμα, κατασκευάστηκε ένα ρομποτικό όχημα ηχητικού εντοπισμού με χρήση δύο μικροφώνων, το οποίο μπορούσε να εντοπίζει ικανοποιητικά την ηχητική πηγή σε προκαθορισμένο χώρο υπό συγκεκριμένες ακουστικές συνθήκες. Η πειραματική διαδικασία πραγματοποιήθηκε υπό αρκετούς περιορισμούς, όπως ο εντοπισμός ηχητικής πηγής με ακίνητο ρομπότ σε εσωτερικό χώρο χωρίς περιβαλλοντικό θόρυβο. Με αυτές τις συνθήκες είχαμε επιτυχία 100% στον εντοπισμό της ηχητικής πηγής. Ωστόσο, αναδείχθηκε ένα σημαντικό σφάλμα, στην περίπτωση έπειτα από εμπόδιο απαιτείτο επαναπροσδιορισμός της ηχητικής πηγής. Επειδή οι κινητήρες παρέμεναν σε λειτουργία, το ρομποτικό όχημα εγκλωβιζόταν σε έναν ατέρμονα βρόχο, εκλαμβάνοντας τον θόρυβο που παρήγαγε το ίδιο ως την εξωτερική ηχητική πηγή, με αποτέλεσμα να εκτελεί κυκλική κίνηση χωρίς να κατευθύνεται προς τον στόχο.

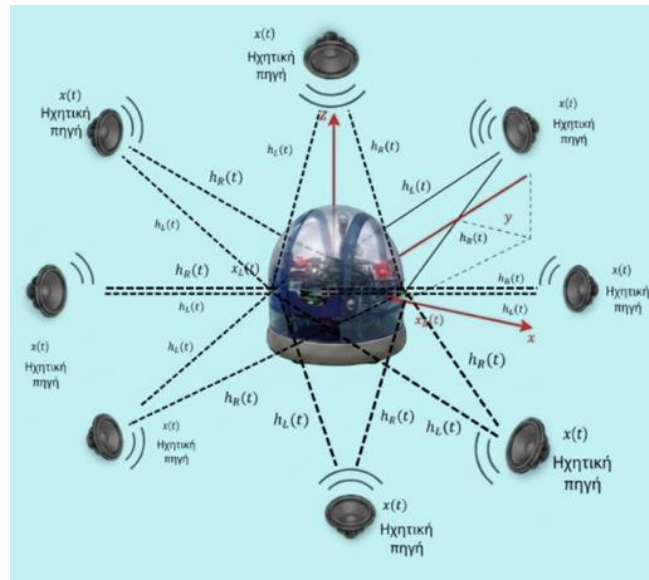
Για την αξιολόγηση της υλοποίησης της παρούσας εργασίας, κρίνεται αρχικά απαραίτητος ο προσδιορισμός των στόχων της, εκ των οποίων ο βασικότερος είναι η αξιολόγηση του εντοπισμού της ηχητικής πηγής.

Σε γενικές γραμμές, ο στόχος αυτός επιτεύχθηκε. Συγκεκριμένα, μελετήθηκε, σχεδιάστηκε, αναπτύχθηκε και υλοποιήθηκε ένα ρομποτικό όχημα, το οποίο έχει τη δυνατότητα να εκτιμά και να εντοπίζει τη θέση της ηχητικής πηγής. Στη βασική έκδοση, όπου ο εντοπισμός πραγματοποιούνταν με το ρομποτικό όχημα ακίνητο και ακίνητη ηχητική πηγή, χωρίς ιδιαίτερο περιβαλλοντικό θόρυβο και με ηχητικό σήμα από ένα επίπεδο ενέργειας και πάνω, τα αποτελέσματα υπήρξαν ιδιαίτερα ενθαρρυντικά. Αντιθέτως, στην περίπτωση που ο εντοπισμός ηχητικής πηγής επιχειρήθηκε με κινούμενο ρομποτικό όχημα ή κινούμενη ηχητική πηγή, τα αποτελέσματα παρουσίασαν αποκλίσεις από τα αναμενόμενα. Για την ομαλή λειτουργία σε αυτή την περίπτωση, απαιτούνταν περιβάλλον χωρίς θόρυβο και σημαντικά υψηλότερο επίπεδο εκπεμπόμενου σήματος από την ηχητική πηγή.

Συμπερασματικά, τα τελικά αποτελέσματα κρίνονται επαρκή και σε μεγάλο βαθμό ανταποκρίνονται στις αρχικές προσδοκίες.

5.2.2 Αξιολόγηση χωρικού εντοπισμού

Σχετικά με τον χωρικό εντοπισμό της ηχητικής πηγής και τα αποτελέσματα αυτού, σημειώνεται ότι η πειραματική διαδικασία διεξήχθη σε εσωτερικό χώρο περιορισμένων διαστάσεων και υπό ελεγχόμενες ακουστικές συνθήκες. Η εύρεση της ηχητικής πηγής πραγματοποιούνταν επιτυχώς από οποιαδήποτε γωνία ως προς το ρομποτικό όχημα. Ο μοναδικός περιορισμός ήταν η σχέση της στάθμης του ηχητικού σήματος από την πηγή με το περιβαλλοντικό θόρυβο.



Εικόνα 5.8 Παρουσίαση του ρομποτικού οχήματος με χωρικά κατανεμημένες εικονικές ηχητικές πηγές.

5.3 Προβλήματα υλοποίησης

Ως αρχική ιδέα προτάθηκε η μελέτη και η υλοποίηση ενός κινούμενου ρομποτικού οχήματος με σκοπό τον εντοπισμό της ηχητικής πηγής. Στο ρομποτικό όχημα αυτό θα εφαρμόζονταν τέσσερα μικρόφωνα, επιτρέποντας τον τρισδιάστατο ηχητικό εντοπισμό τόσο σε αζιμουθιακό όσο και σε κατακόρυφο επίπεδο. Παράλληλα, σχεδιαζόταν μια συνολική μελέτη για τους συνδυασμούς ακίνητου ή κινούμενου ρομποτικού οχήματος με ακίνητη ή κινούμενη ηχητική πηγή. Η μελέτη θα περιλάμβανε, επίσης, τη μεταβολή των υλικών της επιφάνειας του ρομποτικού οχήματος, το οποίο αρχικά ήταν πλαστικό, και θα καλυπτόταν, για παράδειγμα, με αλουμινόχαρτο, με χαρτί ή με άλλα υλικά. Την μεταβολή των περιβαλλοντικών συνθηκών, υγρασίας και θερμοκρασίας, και του επιπέδου του περιβαλλοντικού θορύβου. Στόχος των παραπάνω ήταν να εξεταστεί ο βαθμός στον οποίο αυτές οι διαφορετικές καταστάσεις θα επηρέαζαν την ακρίβεια και την ταχύτητα εντοπισμού της ηχητικής πηγής. Ωστόσο, αυτή η ιδέα εγκαταλείφθηκε λόγω χρονικών περιορισμών και δυσκολιών στην προσομοίωση των διαφορετικών περιβαλλοντικών συνθηκών.

Στην συνέχεια, η μελέτη επικεντρώθηκε στην απλή έκδοση του ρομποτικού οχήματος που δημιουργήθηκε. Το πρώτο ζήτημα που ανέκυψε ήταν το ενεργειακό, το οποίο σχετίζεται άμεσα με το βάρος του ρομποτικού οχήματος και το πλήθος των περιφερικών εξαρτημάτων που διαθέτει. Ως λύση επιλέχθηκε η διεξαγωγή δοκιμών σύντομης διάρκειας ή απευθείας τροφοδοσία από τον υπολογιστή μέσω του καλωδίου προγραμματισμού.

Ένα επιπλέον σημαντικό πρόβλημα αφορούσε τον ακριβή εντοπισμό την ηχητικής πηγής, από το ρομποτικό όχημα. Ο ακριβής εντοπισμός της ηχητικής πηγής επηρεαζόταν αρνητικά από τον περιβαλλοντικό θόρυβο. Για την αντιμετώπιση αυτών των ζητημάτων αναπτύχθηκαν και ενσωματώθηκαν στο λογισμικό φίλτρα άλφα. Εντούτοις, ένα βασικό πρόβλημα που παρέμεινε άλυτο, ήταν όταν έγινε προσπάθεια εντοπισμού της ηχητικής πηγής με κινούμενο ρομποτικό όχημα, ήταν ο θόρυβος που παράγαν τα ίδια τα μηχανικά μέρη του ρομποτικού οχήματος, ο οποίος επηρέαζε την εκτίμηση εντοπισμού της ηχητικής πηγής.

Στα θετικά στοιχεία για την ανάπτυξη της εργασίας συγκαταλέγεται η ύπαρξη πολυάριθμων διαδικτυακών τόπων συζητήσεων, τα οποία παρείχαν χρήσιμες πληροφορίες, υποστήριξη και συμβουλές ακόμα και στην ελληνική γλώσσα.

5.4 Μελλοντικά ανοιχτά ζητήματα

Ως πρώτο άμεσο ζήτημα που παραμένει ανοιχτό, είναι η μελέτη και η υλοποίηση του εντοπισμού ηχητικής πηγής υπό διάφορους συνδυασμούς όπως μεταξύ ακίνητου ή κινούμενου ρομποτικού οχήματος και ακίνητης ή κινούμενης ηχητικής πηγής. Επίσης, κρίνεται σκόπιμη η ένταξη στην έρευνα του εντοπισμού συγκεκριμένων ήχων όπως η ανθρώπινη ομιλία, καθώς και η δημιουργία φίλτρων για την αφαίρεση του περιβαλλοντικού θορύβου ή των ηχητικών φαινομένων που σχετίζονται με τους κλειστούς χώρους όπως πχ. η αντήχηση. Παράλληλα, προτείνεται ο εγκλεισμός του συστήματος οδήγησης των τροχών και του βηματικού κινητήρα με ηχοαπορροφητικό υλικό, ώστε να μειωθεί ο θόρυβος που παράγουν τα ίδια τα μηχανικά μέρη και ο οποίος επηρεάζει αρνητικά την διαδικασία εντοπισμού της ηχητικής πηγής.

Ένα άλλο βασικό πεδίο βελτίωσης αποτελεί η αντικατάσταση των μπαταριών μόλυβδου με μπαταρίες λιθίου μεγαλύτερης χωρητικότητας. Η αλλαγή αυτή θα επιτρέψει την παροχή ενέργειας σε ολόκληρο το ρομποτικό όχημα, εξαλείφοντας την ανάγκη για δύο ανεξάρτητες τροφοδοσίες. Στην περίπτωση αυτή, κρίνεται απαραίτητη η ενσωμάτωση ενός σταθεροποιητή τάσης στην τροφοδοσία των μικροφώνων, ώστε να διασφαλιστεί η προστασία τους από τυχόν διακυμάνσεις της τάσης. Επιπλέον, στο πλαίσιο αυτό, καθοριστικό ρόλο θα διαδραματίσει ο ήδη υπάρχον μετατροπέας συνεχούς τάσης, διασφαλίζοντας την κατάλληλη πτώση τάσης για ορισμένα περιφερειακά. Οι μπαταρίες λιθίου με ενσωματωμένο κύκλωμα φόρτισης θα παρέχουν την απαραίτητη ενέργεια, μειώνοντας ταυτόχρονα το συνολικό βάρος του ρομποτικού οχήματος, γεγονός που θα επιδράσει θετικά στην ενεργειακή αυτονομία του ρομποτικού οχήματος.

Μια επιπλέον πρόταση αφορά την υλοποίηση της αρχικής ιδέας για ενσωμάτωση τεσσάρων μικροφώνων στο ρομποτικό όχημα. Η επέκταση αυτή είναι εφικτή στην παρούσα διάταξη με την τοποθέτηση και σύνδεση των επιπλέον μικροφώνων, ώστε να επιτευχθεί τρισδιάστατος εντοπισμός της ηχητικής πηγής. Η μελέτη, επιπλέον, θα περιλάμβανε τη μεταβολή των υλικών της επιφάνειας του ρομποτικού οχήματος που είναι από πλαστικό, και θα μπορούσε να καλυφθεί με αλουμινόχαρτο, με χαρτί ή με άλλα υλικά, καθώς και τη μεταβολή των περιβαλλοντικών συνθηκών, δηλαδή της υγρασίας και της θερμοκρασίας, και του επιπέδου του περιβαλλοντικού θορύβου. Στόχος των ανωτέρω δοκιμών είναι να εξεταστεί ο βαθμός στον οποίο οι μεταβαλλόμενες αυτές συνθήκες επηρέαζαν την ακρίβεια και την ταχύτητα εντοπισμού της ηχητικής πηγής. Συνοπτικά, προτείνεται η διεξαγωγή μιας στατιστικής μελέτης σχετικά με την επίδραση των διαφόρων αυτών καταστάσεων στην ικανότητα του εντοπισμού της ηχητικής πηγής.

Επιπλέον μπορεί να διεξαχθεί μελέτη για εναλλακτικές μεθόδους προσέγγισης για τον εντοπισμό της ηχητικής πηγής, πέρα αυτών που αναπτύχθηκαν κατά την διάρκεια της παρούσας εργασίας.

6. Βιογραφικές αναφορές

6.1 Ελληνόγλωσσες Βιογραφικές αναφορές

Αλεξανδράκη, Χ. (2020). *Τεχνολογίες ήχου και μουσικής* [Σημειώσεις]. Ελληνικό Μεσογειακό Πανεπιστήμιο.

Βλαχόπουλος, Μ. Γ. (2022). *Ήχος και θόρυβος* [Σημειώσεις].

Ευταξίας, Κ. (2015). *Εισαγωγή στην Κυματική – Ερωτήσεις και Ασκήσεις* [Σημειώσεις].

Καραϊσκος, Π. (2015). *Υπέρηχοι*. Πανεπιστήμιο Αθηνών.

Κοντογιάννης, Α. (2019). *Μελέτη, επεξεργασία και μετάδοση ραδιοφωνικού σήματος* [Σημειώσεις].

Λουτρίδης, Σ. Ι. (2024). *Ηλεκτροακουστική*. Εκδόσεις Τζιόλα.

Μουρτζόπουλος, Γ. (2016). *Σημειώσεις Ηλεκτροακουστικής*. Πανεπιστημίο Πατρών.

Φλώρος, Α. (2013). *Ήχος και Τέχνες*. Ελληνικό Ανοικτό Πανεπιστήμιο.

6.2 Ξενόγλωσσες Βιογραφικές αναφορές

Aditya, S. (2018). *Characterizing the Impact of SNR Heterogeneity on Time-of-Arrival based localization outage probability*.

Ahmed, M. (2012). *Adaptive sub band GSC beam forming using linear microphone-array for noise reduction/speech enhancement*.

Amann-Vesti, B., Aschwanden, M., Gassmann, B., Huck, B., Huck, K., Husmann, M., Jeanneret-Gris, C., Meier, T. O., Thalhammer, C. (2012). *Kursbuch doppler und duplexsonografie*. Georg Thieme Verlag KG.

Andreopoulou, A., Katz, B. FG. (2016). *Investigation on subjective HRTF rating repeatability*.

Asoh, H., Asano, F., Yoshimura, T., Motomura, Y., Ichimura, N., Hara, I., & Ogata, J. (2013). *An application of a particle filter to Bayesian multiple sound source tracking with audio and video information fusion*.

Brüel & Kjaer. (2019). *Microphone handbook*.

Chen, J. C., Yao, K., Hudson, R. E. (2003). *Acoustic source localization and beamforming: Theory and practice*.

Correia, S. D., Tomic, S., Beko, M. (2021). *A feed-forward neural network approach for energy-based acoustic source localization*.

Deleforge, A., Horaud, R., Schechner, Y. Y., Girin, L. (2015). *Co-localization of audio sources in images using binaural features and locally-linear regression*.

Dostalek, P., Vasek, V., Dolinay, J. (2009). *Direction of sound wave arrival determination using time-delay estimation and beamforming methods*.

Doucet, A., Johansen, A. M. (2008). *A tutorial on particle filtering and smoothing: Fifteen years later*.

Edge, A. (2019). *Principles of geolocation techniques*.

Flynn, A. M., Brooks, R. A., Wells III, W. M., Barrett, D. S. (1989). *Squirt: The prototypical mobile robot for autonomous graduate students*. Massachusetts Institute of Technology.

Griffin, R. (2018). *Analysis and prediction of protected-ear localization*.

Hu, D., Huang, Z., Zhang, S., Lu, J. (2018). *Joint TDOA, FDOA and differential Doppler rate estimation: Method and its performance analysis*.

Huang, J., Supaongprapa, T., Terakura, I., Wang, F., Ohnishi, N., Sugie, N. (1999). *A model-based sound localization system and its application to robot navigation*.

Hwang, S., Park, Y., Park, Y. S. (2011). *Sound direction estimation using an artificial ear for robots*.

Kim, H.-D., Komatani, K., Ogata, T., Okuno, H. G. (2009). *Human tracking system integrating sound and face localization using an expectation-maximization algorithm in real environments*.

Kim, U.-H., Mizumoto, T., Ogata, T., Okuno, H. G. (2011). *Improvement of speaker localization by considering multipath interference of sound wave for binaural robot audition*.

Knapp, C., Carter, C. (1976). *The generalized correlation method for estimation of time delay*.

Koka, K., Read, H. L., Tollin, D. J. (2013). *Distance-dependent head-related transfer functions and interaural level differences for the rabbit*.

Krol, D., Lorenc, A., Swiecinski, R. (2015). *Detecting laterality and nasality in speech with the use of a multi-channel recorder*.

Liang, Z., Ma, X., Dai, X. (2008). *Robust tracking of moving sound source using multiple model Kalman filter*.

Liao, W. (2015). *MUSIC for multidimensional spectral estimation: Stability and super-resolution*.

Lim, H., Yoo, I.-C., Cho, Y., Yook, D. (2015). *Speaker localization in noisy environments using steered response voice power*.

Liu, Y., Yang, Z., Wang, X. (2010). *Location, localization and localizability*.

Lu, Y. C., Cooke, M. (2010). *Binaural estimation of sound source distance via the direct-to-reverberant energy ratio for static and moving sources*.

Martinson, E., Schultz, A. (2007). *Robotic discovery of the auditory scene*.

Montalbano, V. (2012). *An empirical study on a learning path on wave physics focused on energy*.

Murray, J. C., Wermter, S., Erwin, H. R. (2009). *Robotic sound-source localisation architecture using cross-correlation and recurrent neural networks*.

Murray, J. C., Wermter, S., Erwin, H. R. (2009). *Bioinspired auditory sound localisation for improving the signal to noise ratio of socially interactive robots*. University of Sunderland.

Nakamura, K., Nakadai, K., Asano, F., Hasegawa, Y., Tsujino, H. (2009). *Intelligent sound source localization for dynamic environments*.

Ogiso, S., Kawagishi, T., Mizutani, K., Wakatsuki, N., Zempo, K. (2015). *Self-localization method for mobile robot using acoustic beacons*.

Pourmehrer, S., Bruce, J., Wawerla, J., Vaughan, R. (2015). *A sensor fusion framework for finding an HRI partner in crowd*. Simon Fraser University.

Rascon, C., Meza, I. (2016). *Localization of sound sources in robotics*.

Rodemann, T. (2010). *A study on distance estimation in binaural sound localization*.

Sasaki, Y., Kagami, S., Mizoguchi, H. (2006). *Multiplesound source mapping for a mobile robot by self-motion triangulation*.

Scharine, A. A., Cave, K. D., Letowski, T. R. (2009). *Sound pressure and sound pressure level*.

Sheng, X., Hu, Y.-H. (n.d.). *Energy based acoustic source localization*.

Staab, W. (2023). *Hearing with two ears*.

Stamp, E. M., Liesche-Starnecker, F., Schaller, T., Baumgarten, P., Lilla, N., Mielke, D., Collins, D. (2026). *Acoustic sensing as a tool for brain tumor diagnostics*.

Thoen, B., Wielandt, S., De Strycker, L. (2019). *Improving AoA localization accuracy in wireless acoustic sensor networks with angular probability density functions*.

Tortosa, D. D. (2021). *Acoustic positioning system for 3D localization of sound sources based on the time of arrival of a signal for a low-cost system*.

Valin, J.-M., Michaud, F., Hadjou, B., Rouat, J. (2004). *Localization of simultaneous moving sound sources for mobile robot using a frequency-domain steered beamformer approach*.

Van Velzen, J. (2012). *Spatial perception in real-life acoustics*.

Vincent, E., Sini, A., Charpillat, F. (2015). *Audio source localization by optimal control of a mobile robot*.

Youssef, K., Argentieri, S., Zarader, J. L. (2013). *A learning-based approach to robust binaural sound localization*.

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Yun, J., Kim, M. (2020). *SybilEye: Observer-assisted privacy-preserving sybil attack detection on mobile crowdsensing*.

Zhao, Y., Li, Z., Huang, L., Hao, B. (2023). *On the performance analysis of TDOA localization with uncommon reference nodes*.

Zhong, X. (2015). *Dynamic spatial hearing by human and robot listeners*.

7. Ιστογραφία

Πανελλήνιο σχολικό δίκτυο. (χ.χ.). *Το δίκτυο στην υπηρεσία της εκπαίδευσης*. users.sch.gr

Brainkart. (n.d.). *Education*. www.brainkart.com

Wikiwand. (n.d.). *Wikipedia*. www.wikiwand.com

Orion strom. (χ.χ.). *Στρώματα ύπνου για όλους*. www.orionstrom.gr

Amazona. (n.d.). *Music can change the world*. www.amazona.de

Teufel Blog. (n.d.). *Die schönsten Momente hörst du laut*. blog.teufel.de

Mikrofon-tests. (n.d.). *Mikrofon test & vergleich*. mikrofon-tests.de

Mikrofon. (n.d.). *Mikrofon*. www.mikrofon.guru

Arrow Electronics. (n.d.). *Electronic components*. www.arrow.de

Turanis. (n.d.). *Elektronik-Projekte und Experimente*. elektro.turanis.de

Hollyland. (n.d.). *Hollyland wireless audio & video solutions*. www.hollyland.com

Stageaid. (n.d.). *Stageaid: Veranstaltungstechnik & Live-Sound*. www.stageaid.de

Microfono Rocks. (n.d.). *Microfono Rocks: Equipamiento de audio*. microfono.rocks

TOA Electronics. (n.d.). *Sound TOA: Commercial audio systems*. www.sound-toa.com

Hygeia. (n.d.). *Υπηρεσίες υγείας*. www.hygeia.gr

Ακουστικά Θεοδώρου. (χ.χ.). *Ακουστικά βαρηκοΐας Θεοδώρου*. www.akoustikatheodorou.gr

Naver. (n.d.). *Naver blog platform*. blog.naver.com

Alamy. (n.d.). *Alamy stock photos*. www.alamy.de

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Uacgr. (χ.χ.). *United Acoustical Consultants*. uacgr.blogspot.com

Web Urbanist. (n.d.). *WebUrbanist: Design, architecture and art*. weburbanist.com

Grokipedia. (n.d.). *Grokipedia reference wiki*. grokipedia.com

Zhihu. (n.d.). *Zhuanlan Zhihu knowledge platform*. zhuanlan.zhihu.com

Essel Acozstics. (n.d.). *Essel Acoustics: Sound and audio engineering*. esselacoustics.com

Gfaitech. (n.d.). *Gfaitech: Acoustic cameras and sound engineering*. www.gfaitech.com

Az-delivery. (n.d.). *AZ-Delivery: Mikroelektronik und Zubehör*. www.az-delivery.de

HandsOn Tech. (n.d.). *HandsOn Technology: Electronics specs and tutorials*. www.handsontec.com

Amazon. (n.d.). *Amazon: Online-Shopping für Elektronik, Bücher & mehr*. www.amazon.de

Ardumotive. (χ.χ.). *Ardumotive: DIY electronics and Arduino projects*. www.ardumotive.com

Components101. (n.d.). *Electronic component datasheets & pinouts*. components101.com

Electronic Wings. (n.d.). *Hardware and software tutorials*. www.electronicwings.com

Mouser Electronics. (n.d.). *Mouser Electronics: Electronic components distributor*. www.mouser.com

Alldatasheet. (n.d.). *Alldatasheet: Electronic components datasheet search*. www.alldatasheet.com

Arduino. (n.d.). *Arduino official homepage*. www.arduino.cc

8. Video

Φυσική Γ Λυκείου. (2020, 15 Μαρτίου). *Περίθλαση & Αρχή του Huygens* [Video]. YouTube.
https://www.youtube.com/watch?v=-gHoHAz_pC8&t=97s

Rohde & Schwarz. (2023, 18 Απριλίου). *Understanding I2C* [Video]. YouTube.
<https://www.youtube.com/watch?v=CAvawEcxoPU>

9. Data sheets

Texas Instruments. (2023). *LM2596 SIMPLE SWITCHER® Power Converter 150-kHz 3-A Step-Down Voltage Regulator*: Data sheet.
<https://www.ti.com/lit/ds/symlink/lm2596.pdf#page=37.16>

Handson Technology. (n.d.). *L298N Dual H-Bridge Motor Driver*: Data sheet.
<https://www.handsontec.com/dataspecs/L298N%20Motor%20Driver.pdf>

Elec Freaks. (n.d.). *Ultrasonic Ranging Module HC - SR04*: Data sheet.
<https://cdn.sparkfun.com/datasheets/Sensors/Proximity/HCSR04.pdf>

Handson Technology. (n.d.). *SG90 TowerPro Micro Servo Motor*: Data sheet.
https://handsontec.com/dataspecs/motor_fan/SG90-Servo.pdf

Hitachi. (1999). *HD44780U (LCD-II)*: Data sheet.
<https://cdn.sparkfun.com/assets/9/5/f/7/b/HD44780.pdf>

Mouser Electronics. (n.d.). *DHT11 Humidity & Temperature Sensor*: Data sheet.
https://www.mouser.com/datasheet/2/758/DHT11-Technical-Data-Sheet-Translated-Version-1143054.pdf?srltid=AfmBOoqsl2SioIXAHBpbk7uWUuXrB_TOFNLswjuKMWtsSKjgnGaRXM6a#page=2.00

Joy-it. (n.d.). *KY-006 Passiv Piezo-Buzzer module*: Data sheet.
https://eclass.uth.gr/modules/document/file.php/E-CE_U_269/Sensors/Sensors_%20Datasheets/KY-006-Joy-IT.pdf#page=1.00

Same sky. (2024). *CMA-6542PF Electret condenser microphone*: Data sheet.
https://www.sameskydevices.com/product/resource/cma-6542pf.pdf?srltid=AfmBOoqbVlrOKUabNJwVRe3JmZ7wN5TbBG0XH3Uufm2_wNEzyaUQ3Y0Q

Arduino. (2026). *Arduino® Mega 2560 Rev3*: Data sheet.
<https://docs.arduino.cc/resources/datasheets/A000067-datasheet.pdf>

Παράρτημα Α

9.1 Αλγόριθμος 1

Στην συνέχεια παρουσιάζεται η έκδοση α του προγράμματος, η οποία είναι αυτή με την οποία έγιναν οι δοκιμές για την λειτουργία του ρομποτικού οχήματος.

```
/* Arduino Robot with servo motor and ultrasonic sensor HC-SR04
   Hellenic Open University
   Sound Design and Digital Sound
   Dev. Petros Amorgianos
   Date: 30.06.2026    */

// Library
#include "Ultrasonic.h"
#include <Servo.h>
#include <Wire.h>
#include <LiquidCrystal_I2C.h>
#include <avr/wdt.h>
#include <DHT.h>

//=== PINS =====

// Motors
const int motorR1 = 6; // Motor A positive (+) to pin 6 (PWM)
const int motorR2 = 5; // Motor A positive (-) to pin 5 (PWM)
const int motorL1 = 11; // Motor B positive (+) to pin 11 (PWM)
const int motorL2 = 10; // Motor B positive (-) to pin 10 (PWM)

// Button & buzzer
const int button = 4; //Botton for reset
const int buzzer = 12; //Buzzer

// Leds
const int ledGreen = 7; //Led green
```

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

```
const int ledYellow = 8; //Led yellow
const int ledRed = 9; //Led red

// Microphones
const int micLeft = A4; //Left microphone
const int micRight = A3; //Right microphone

// Servo
Servo myServo;
const int servoPin = 3; //Servo

// Ultrasonic
Ultrasonic ultrasonic(A0,A1); // Create Ultrasonic object ultrasonic (trig pin, echo pin)

// DHT11
#define DHTPIN 13 //Sensor temperatur and humanity
#define DHTTYPE DHT11
DHT dht(DHTPIN, DHTTYPE);

// LCD
LiquidCrystal_I2C lcd(0x3F, 16, 2); // set the LCD address to 0x3F for a 16 chars and 2 line display

//=== Variables =====

//Variables
int distance; // Variable to store distance from an object
int threshold = 20;

int soundLevel = 0;

// Microphones calibration
int baseLeft = 0;
int baseRight = 0;

// Filter
float alpha = 0.7; //Filter to reduce environmental noise
int smoothLeft = 0;
int smoothRight = 0;
```

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

```
// FSM
enum State {
    WAIT_SOUND, MOVE_TO_SOUND, AVOID_OBSTACLE, TARGET_REACHED
};

State currentState = WAIT_SOUND;
int lockedDirection = 0; // -1 left, 0 forward, 1 right

//=== Functions =====

int readMicLevel(int pin);
void stop();
void forward();
void backward();
void left();
void right();
void resetArduino();

//=== Setup =====

void setup(){

    pinMode(motorR1, OUTPUT);
    pinMode(motorR2, OUTPUT);
    pinMode(motorL1, OUTPUT);
    pinMode(motorL2, OUTPUT);

    pinMode(ledGreen, OUTPUT);
    pinMode(ledYellow, OUTPUT);
    pinMode(ledRed, OUTPUT);

    digitalWrite(ledGreen, HIGH);

    pinMode(buzzer, OUTPUT);
    pinMode(button, INPUT_PULLUP);

    Serial.begin(9600);
```

```
myServo.attach(servoPin); // servo fixed
myServo.write(70);

//=== Microphones =====

delay(2000);
for (int i=0; i<50; i++) {
  baseLeft += analogRead(micLeft);
  baseRight += analogRead(micRight);
  delay(10);
}

baseLeft /= 50;
baseRight /= 50;

//=== DHT sensor =====

dht.begin();
delay (1500);

//=== LCD =====

lcd.init();      // initialize the lcd
lcd.backlight(); // Turn on the LCD screen backlight

lcd.setCursor(0, 0);
lcd.print(" Hellenic Open ");
lcd.setCursor(0, 1);
lcd.print(" University ");
delay(3000);
lcd.clear();

lcd.setCursor(1, 0);
lcd.print("Acoustic design");
delay(3000);
lcd.clear();

lcd.setCursor(1, 0);
lcd.print(" and ");
delay(3000);
lcd.clear();
```

```
lcd.setCursor(2, 0);
```

```
lcd.print("Digital Sound");
```

```
delay(3000);
```

```
lcd.clear();
```

```
lcd.setCursor(0, 0);
```

```
lcd.print("  Petros  ");
```

```
lcd.setCursor(0, 1);
```

```
lcd.print("  Amorgianos  ");
```

```
lcd.clear();
```

```
float h = dht.readHumidity();
```

```
float t = dht.readTemperature();
```

```
lcd.setCursor(0,0);
```

```
lcd.print("Temp: ");
```

```
lcd.print(t);
```

```
lcd.print(" C");
```

```
lcd.setCursor(0,1);
```

```
lcd.print("Hum: ");
```

```
lcd.print(h);
```

```
lcd.print(" %");
```

```
delay(3000);
```

```
lcd.clear();
```

```
lcd.setCursor(0, 0);
```

```
lcd.print("  Petros  ");
```

```
lcd.setCursor(0, 1);
```

```
lcd.print("  Amorgianos  ");
```

```
}
```

```
//=== LOOP =====
```

```
void loop() {
```

```
//=== RESET BUTTON =====
```

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

```
if (digitalRead(button) == LOW) {
    resetArduino();
}

//=== Distance =====

distance = ultrasonic.Ranging(CM);

//=== SOUND =====

int rawLeft = readMicLevel(micLeft);
int rawRight = readMicLevel(micRight);

//=== FILTER =====

smoothLeft = alpha * smoothLeft + (1 - alpha) * rawLeft;
smoothRight = alpha * smoothRight + (1 - alpha) * rawRight;

int leftSound = smoothLeft;
int rightSound = smoothRight;

soundLevel = (leftSound + rightSound) / 2;
int diff = leftSound - rightSound;
// int deadZone = 15;

// ===== CURRENT STATES =====

switch (currentState) {

// ===== 1.WAIT =====

case WAIT_SOUND:

    digitalWrite(ledYellow, LOW);
    digitalWrite(ledRed, LOW);
    stop();

    if (soundLevel > 20) {
```

```
if (diff > threshold) lockedDirection = -1;
else if (diff < -threshold) lockedDirection = 1;
else lockedDirection = 0;

currentState = MOVE_TO_SOUND;
}

break;

// ===== 2. MOVE =====

case MOVE_TO_SOUND:

digitalWrite(ledYellow, HIGH);

if (distance < 10 ) {
    currentState = TARGET_REACHED;
    break;
}

if (distance < 20 ) {
    currentState = AVOID_OBSTACLE;
    break;
}

if (lockedDirection == -1) left();
else if (lockedDirection == 1) right();
else forward();

break;

// ===== 3. AVOID =====

case AVOID_OBSTACLE:

digitalWrite(ledYellow, LOW);
digitalWrite(ledRed, HIGH);

backward();
delay(300);
```

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

```
right();
delay(400);

digitalWrite(ledRed, LOW);
digitalWrite(ledYellow, HIGH);

currentState = MOVE_TO_SOUND;

break;

// ==== 4. TARGET =====

case TARGET_REACHED:

stop();

digitalWrite(ledYellow, LOW);
digitalWrite(ledRed, HIGH);

tone(buzzer, 1000);
delay(3000);
noTone(buzzer);

delay(5000);

digitalWrite(ledRed, LOW);

currentState = WAIT_SOUND;

break;

}

delay(50);

}

//==== MICROPHONE =====
```

```
int readMicLevel(int pin) {
    long sum = 0;
    for(int i=0; i<15; i++) {
        int val = analogRead(pin);
        if(pin == micLeft)
            sum += abs(val - baseLeft);
        else
            sum += abs(val - baseRight);
    }
    return sum / 15;
}

//===  RESET  =====

void resetArduino() {
    wdt_enable(WDTO_15MS);
    while(1) {}
}

//===  MOVEMENT  =====

//Movement functions
void forward(){
    digitalWrite(motorR1, HIGH);
    digitalWrite(motorR2, LOW);
    digitalWrite(motorL1, HIGH);
    digitalWrite(motorL2, LOW);
}

void backward(){
    digitalWrite(motorR1, LOW);
    digitalWrite(motorR2, HIGH);
    digitalWrite(motorL1, LOW);
    digitalWrite(motorL2, HIGH);
}

void right(){
    digitalWrite(motorR1, LOW);
    digitalWrite(motorR2, HIGH);
    digitalWrite(motorL1, HIGH);
```

```
digitalWrite(motorL2, LOW);  
}
```

```
void left(){  
    digitalWrite(motorR1, HIGH);  
    digitalWrite(motorR2, LOW);  
    digitalWrite(motorL1, LOW);  
    digitalWrite(motorL2, HIGH);  
}
```

```
void stop(){  
    digitalWrite(motorR1, LOW);  
    digitalWrite(motorR2, LOW);  
    digitalWrite(motorL1, LOW);  
    digitalWrite(motorL2, LOW);  
}
```

Παράρτημα Β

9.2 Αλγόριθμος 2

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

Στην συνέχεια παρουσιάζεται η έκδοση β του προγράμματος, η οποία ωστόσο δεν ενσωματώθηκε στην τελική αναφορά, εξαιτίας των εμφανών σφαλμάτων που προκαλούνται λόγω του θορύβου κατά την διαδικασία κίνησης του του ρομπότ.

```
/* Arduino Robot with servo motor and ultrasonic sensor HC-SR04
```

```
    Hellenic Open University
```

```
    Sound Design and Digital Sound
```

```
    Dev. Petros Amorgianos
```

```
    Date: 30.06.2026    */
```

```
// Library
```

```
#include "Ultrasonic.h"
```

```
#include <Servo.h>
```

```
#include <Wire.h>
```

```
#include <LiquidCrystal_I2C.h>
```

```
#include <avr/wdt.h>
```

```
#include <DHT.h>
```

```
//=== PINS =====
```

```
// Motors
```

```
const int motorR1 = 6; // Motor A positive (+) to pin 6 (PWM)
```

```
const int motorR2 = 5; // Motor A positive (-) to pin 5 (PWM)
```

```
const int motorL1 = 11; // Motor B positive (+) to pin 11 (PWM)
```

```
const int motorL2 = 10; // Motor B positive (-) to pin 10 (PWM)
```

```
// Button & buzzer
```

```
const int button = 4; //Botton for reset
```

```
const int buzzer = 12; //Buzzer
```

```
// Leds
```

```
const int ledGreen = 7; //Led green
```

```
const int ledYellow = 8; //Led yellow
```

```
const int ledRed = 9; //Led red
```

```
// Microphones
```

```
const int micLeft = A4; //Left microphone
```

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

```
const int micRight = A3; //Right microphone

// Servo
Servo myServo;
const int servoPin = 3; //Servo

// Ultrasonic
Ultrasonic ultrasonic(A0,A1); // Create Ultrasonic object ultrasonic (trig pin, echo pin)

// DHT11
#define DHTPIN 13 //Sensor temperatur and humanity
#define DHTTYPE DHT11
DHT dht(DHTPIN, DHTTYPE);

// LCD
LiquidCrystal_I2C lcd(0x3F, 16, 2); // set the LCD address to 0x3F for a 16 chars and 2 line display

//=== Variables =====

//Variables
int distance; // Variable to store distance from an object
int threshold = 20;

int soundLevel = 0;

// Microphones calibration
int baseLeft = 0;
int baseRight = 0;

// Filter
float alpha = 0.7; //Filter to reduce environmental noise
int smoothLeft = 0;
int smoothRight = 0;

// FSM
enum State {
    WAIT_SOUND, MOVE_TO_SOUND, AVOID_OBSTACLE, TARGET_REACHED
};

State currentState = WAIT_SOUND;
```

```
int lockedDirection = 0; // -1 left, 0 forward, 1 right

//=== Functions =====

int readMicLevel(int pin);
void stop();
void forward();
void backward();
void left();
void right();
void resetArduino();

//==== Funktion soundTracking =====

int soundThreshold = 30;
int lastSoundLevel = 0;
int currentSoundTarget = 0;

//=== Check 10 sec =====

unsigned long previousMillis = 0;
const long interval = 10000;

//=== Setup =====

void setup(){

  pinMode(motorR1, OUTPUT);
  pinMode(motorR2, OUTPUT);
  pinMode(motorL1, OUTPUT);
  pinMode(motorL2, OUTPUT);

  pinMode(ledGreen, OUTPUT);
  pinMode(ledYellow, OUTPUT);
  pinMode(ledRed, OUTPUT);

  digitalWrite(ledGreen, HIGH);

  pinMode(buzzer, OUTPUT);
  pinMode(button, INPUT_PULLUP);

  Serial.begin(9600);
```

```
myServo.attach(servoPin); // servo fixed
myServo.write(70);

//=== Microphones =====

delay(2000);
for (int i=0; i<50; i++) {
  baseLeft += analogRead(micLeft);
  baseRight += analogRead(micRight);
  delay(10);
}

baseLeft /= 50;
baseRight /= 50;

//=== DHT sensor =====

dht.begin();
delay (1500);

//=== LCD =====

lcd.init();    // initialize the lcd
lcd.backlight(); // Turn on the LCD screen backlight

lcd.setCursor(0, 0);
lcd.print("  Hellenic Open ");
lcd.setCursor(0, 1);
lcd.print("  University ");
delay(3000);
lcd.clear();

lcd.setCursor(1, 0);
lcd.print("Acoustic design");
delay(3000);
lcd.clear();
lcd.setCursor(1, 0);
lcd.print("  and  ");
delay(3000);
```

```
lcd.clear();
lcd.setCursor(2, 0);
lcd.print("Digital Sound");
delay(3000);
lcd.clear();

lcd.setCursor(0, 0);
lcd.print("  Petros  ");
lcd.setCursor(0, 1);
lcd.print("  Amorgianos  ");
lcd.clear();

float h = dht.readHumidity();
float t = dht.readTemperature();

lcd.setCursor(0,0);
lcd.print("Temp: ");
lcd.print(t);
lcd.print(" C");

lcd.setCursor(0,1);
lcd.print("Hum: ");
lcd.print(h);
lcd.print(" %");

delay(3000);
lcd.clear();

lcd.setCursor(0, 0);
lcd.print("  Petros  ");
lcd.setCursor(0, 1);
lcd.print("  Amorgianos  ");

}

//===  LOOP  =====

void loop() {

//===  RESET BUTTON  =====
```

```
if (digitalRead(button) == LOW) {
    resetArduino();
}

//=== Distance =====

distance = ultrasonic.Ranging(CM);

//=== SOUND =====

int rawLeft = readMicLevel(micLeft);
int rawRight = readMicLevel(micRight);

//=== FILTER =====

smoothLeft = alpha * smoothLeft + (1 - alpha) * rawLeft;
smoothRight = alpha * smoothRight + (1 - alpha) * rawRight;

int leftSound = smoothLeft;
int rightSound = smoothRight;

soundLevel = (leftSound + rightSound) / 2;
int diff = leftSound - rightSound;
// int deadZone = 15;

// ===== CURRENT STATES =====

switch (currentState) {

// ===== 1.WAIT =====

case WAIT_SOUND:

    digitalWrite(ledYellow, LOW);
    digitalWrite(ledRed, LOW);
    stop();

    if (soundLevel > 20) {
```

```
if (diff > threshold) lockedDirection = -1;
else if (diff < -threshold) lockedDirection = 1;
else lockedDirection = 0;

currentState = MOVE_TO_SOUND;
}

break;

// ===== 2. MOVE =====

case MOVE_TO_SOUND:

digitalWrite(ledYellow, HIGH);

if (distance < 10 ) {
    currentState = TARGET_REACHED;
    break;
}

if (distance < 20 ) {
    currentState = AVOID_OBSTACLE;
    break;
}

if (lockedDirection == -1) left();
else if (lockedDirection == 1) right();
else forward();

break;

// ===== 3. AVOID =====

case AVOID_OBSTACLE:

digitalWrite(ledYellow, LOW);
digitalWrite(ledRed, HIGH);

backward();
delay(300);
```

```
right();
delay(400);

digitalWrite(ledRed, LOW);
digitalWrite(ledYellow, HIGH);

currentState = MOVE_TO_SOUND;

break;

// ==== 4. TARGET =====

case TARGET_REACHED:

stop();

digitalWrite(ledYellow, LOW);
digitalWrite(ledRed, HIGH);

tone(buzzer, 1000);
delay(3000);
noTone(buzzer);

delay(5000);

digitalWrite(ledRed, LOW);

currentState = WAIT_SOUND;

break;

}

delay(50);

}

// ===== Continuous sound detection function =====
```

```
void soundTracking() {

    // Read value for microphones
    int leftSound = analogRead(micLeft);
    int rightSound = analogRead(micRight);

    // Calculate incensity
    int maxSound = max(leftSound, rightSound);

    if (maxSound < soundThreshold) {
        return;
    }

    // Check new Sound
    if (maxSound > lastSoundLevel + 20) {

        lastSoundLevel = maxSound;

        // Lock target
        if (leftSound > rightSound) {
            currentSoundTarget = 1; // αριστερά
        } else {
            currentSoundTarget = 2; // δεξιά
        }
    }

    // ===== Continuous sound search =====
    soundTracking();
}

//=== MICROPHONE =====

int readMicLevel(int pin) {
    long sum = 0;
    for(int i=0; i<15; i++) {
        int val = analogRead(pin);
        if(pin == micLeft)
```

```
        sum += abs(val - baseLeft);
    else
        sum += abs(val - baseRight);
    }
    return sum / 15;
}

//===  RESET  =====

void resetArduino() {
    wdt_enable(WDTO_15MS);
    while(1) {}
}

//===  MOVEMENT  =====

//Movement functions
void forward(){
    digitalWrite(motorR1, HIGH);
    digitalWrite(motorR2, LOW);
    digitalWrite(motorL1, HIGH);
    digitalWrite(motorL2, LOW);
}

void backward(){
    digitalWrite(motorR1, LOW);
    digitalWrite(motorR2, HIGH);
    digitalWrite(motorL1, LOW);
    digitalWrite(motorL2, HIGH);
}

void right(){
    digitalWrite(motorR1, LOW);
    digitalWrite(motorR2, HIGH);
    digitalWrite(motorL1, HIGH);
    digitalWrite(motorL2, LOW);
}

void left(){
    digitalWrite(motorR1, HIGH);
```

Διπλωματική εργασία: Τεχνικές χωρικής επεξεργασίας ήχου σε ρομποτικές εφαρμογές.

```
digitalWrite(motorR2, LOW);  
digitalWrite(motorL1, LOW);  
digitalWrite(motorL2, HIGH);  
}
```

```
void stop(){  
    digitalWrite(motorR1, LOW);  
    digitalWrite(motorR2, LOW);  
    digitalWrite(motorL1, LOW);  
    digitalWrite(motorL2, LOW);  
}
```