



ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΣΧΟΛΗ ΜΟΥΣΙΚΗΣ & ΟΠΤΙΚΟΑΚΟΥΣΤΙΚΩΝ ΤΕΧΝΩΝ  
ΤΜΗΜΑ ΤΕΧΝΩΝ ΗΧΟΥ & ΕΙΚΟΝΑΣ

Προσαρμοστικά Περιβάλλοντα Επαυξημένης  
Ηχητικής Πραγματικότητας με χρήση  
Δυναμικά Μεταβαλλόμενης Ηχητικής Μείξης

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

του

Μουστάκα Νικόλαου

Κέρκυρα, Μάρτιος 2021



# Περίληψη

Το πεδίο της Επαυξημένης Πραγματικότητας Ήχου, έχει εισάγει τον μείκτη Επαυξημένης πραγματικότητας ήχου ως μια μέθοδο για τη μείξη πραγματικού με εικονικό περιβάλλον. Σκοπός της παρούσας διατριβής είναι η αναβάθμιση της λειτουργίας του παραδοσιακού μοντέλου επαυξημένης μείξης ήχου με την ενσωμάτωση προσαρμοστικότητας για τον έλεγχο του απρόβλεπτου και μεταβλητού πραγματικού περιβάλλοντος σε τυπικές εφαρμογές επαυξημένης πραγματικότητας ήχου. Προτείνεται ένα νέο αντιληπτικό μοντέλο επαυξημένης μείξης ήχου το οποίο αναμιγνύει τα ακουστικά περιβάλλοντα με όρους αμφιωτικής ακουστότητας. Επιπλέον, μια νέα τεχνική πρόβλεψης αμφιωτικής ακουστότητας σε πραγματικό χρόνο για την εξόρυξη χαρακτηριστικών ακουστικής αντίληψης ενσωματώνεται στο νέο μοντέλο μείξης. Το προτεινόμενο μοντέλο δοκιμάστηκε και αξιολογήθηκε για την αποδοτικότητά του μέσα από ποιοτικές και ποσοτικές μεθόδους καθώς και με τη χρήση σε ένα ηχητικό παιχνίδι Επαυξημένης Πραγματικότητας Ήχου. Τα αποτελέσματα της ποιοτικής αξιολόγησης δείχνουν μια προτίμηση των υποκειμένων στο μοντέλο προσαρμοστικής μείξης. Η ανάλυση της απόδοσης των παικτών του ηχητικού παιχνιδιού στη διάρκεια της αξιολόγησης έδειξε μεγαλύτερα ποσοστά επιτυχίας για το μοντέλο προσαρμοστικής μείξης σε σύγκριση με το παραδοσιακό μοντέλο μείξης επαυξημένης πραγματικότητας ήχου. Συνεπώς, φαίνεται πως η νέα προσέγγιση της προσαρμοστικής επαυξημένης μείξης ήχου, είναι αποδοτικότερη για τους παίκτες του ηχητικού παιχνιδιού καθώς βελτιώνεται η εμπειρία διάδρασης μέσα σε μια τυπική εφαρμογή επαυξημένης πραγματικότητας ήχου όπως είναι το ηχητικό παιχνίδι επαυξημένης πραγματικότητας.

## Λέξεις Κλειδιά

Επαυξημένη Πραγματικότητα Ήχου, Προσαρμοστική μείξη, Αμφιωτική Ακουστότητα, Ηχητικά Παιχνίδια Επαυξημένης πραγματικότητας.



# Abstract

The field of Augmented Reality Audio (ARA) has introduced the ARA mixer as a method to mix the real with the virtual environment. The goal of this dissertation is to upgrade the function of ARA mix legacy model through the integration of adaptation to control the auditory fluctuations of unpredictable real environment in typical ARA applications. We propose a new perceptual ARA mixer model which adapts the blending of auditory environments in terms of binaural loudness. Consequently, a new technique of real-time binaural loudness prediction through feature extraction was integrated to ARA mixer. The effectiveness of proposed model of ARA mixer was demonstrated and evaluated through qualitative and quantitative methods; thus, it was applied and evaluated in a typical ARA game. Results show that among ARA mix quality tests, subjects prefer adaptive ARA mix. Analysis of players' performance under ARA game task shows that better success rates can be realized in adaptive mode of ARA mixer instead of legacy ARA mix model. We show that this approach to ARA mixer is significantly more efficient than legacy ARA mixer model and that it provides an improved interactive experience through the tasks of a typical ARA application like the ARA game.

## Keywords

Augmented Reality Audio, Binaural Loudness, Adaptive Mix System, Augmented Reality Audio Game.



# Ευχαριστίες

Ένας κύκλος κλείνει. Ένα κύκλος γεμάτος εμπειρίες και προκλήσεις. Γι' αυτό δεν θα μπορούσα να μην ευχαριστήσω εκείνους που με βοήθησαν και μου έδωσαν τη δύναμη για να τα καταφέρω.

Αρχικά, ένα μεγάλο ευχαριστώ είναι το λιγότερο για τον επιβλέπων Καθηγητή Ανδρέα Φλώρο, που με εισήγαγε στη διαδικασία της έρευνας. Η παρουσία του ήταν καταλυτική ως μέντορας, φίλος και συνεργάτης σε όλη την πορεία του διδακτορικού. Ακολούθως, ευχαριστώ τα υπόλοιπα μέλη της συμβουλευτικής επιτροπής, Ομότιμο Καθηγητή Νικόλαο Γρηγόριο Κανελλόπουλο και Αναπληρωτή Καθηγητή Νικόλαο Αλέξανδρο Τάτλα, των οποίων οι παρατηρήσεις ήταν πολύτιμες για την ολοκλήρωση της διδακτορικής διατριβής.

Την επταμελή επιτροπή, εκτός από τους παραπάνω, πλην του κ. Νικόλαου Γρηγόριου Κανελλόπουλου (όπου λόγω αφυπηρέτησης δεν μπορεί να συμμετέχει), συμπληρώνουν ο Καθηγητής Γεώργιος Καλλίρης, ο Καθηγητής Ιωάννης Ζάννος, ο Καθηγητής Ανδρέας Μνιέστρης, ο Αναπληρωτής Καθηγητής Ιωάννης Δεληγιάννης, ο Αναπληρωτής Καθηγητής Απόστολος Λουφόπουλος. Τους ευχαριστώ όλους για τη συμμετοχή τους και τη συμβολή τους στην παρούσα διατριβή.

Στη συνέχεια θα ήθελα να ευχαριστήσω την Ιωάννα, η οποία στάθηκε όλο αυτό τον καιρό δίπλα μου σε καλές και κακές στιγμές, και ανέχτηκε τις ιδιαιτερότητές μου ως υποψήφιος διδάκτορας. Επίσης, ευχαριστώ τους φίλους και συνεργάτες από το εργαστήριο «Επεξεργασίας Οπτικοακουστικών Σημάτων - ΕΠΟΑΣΗ» του τμήματος Τεχνών Ήχου και Εικόνας του Ιόνιου Πανεπιστημίου, Δρ. Κωνσταντίνo Βόγκλη και Δρ. Εμμανουήλ Ροβίθη για την επιστημονική τους βοήθεια. Ακολούθως, ευχαριστώ τον Σωτήρη Βλάχο για τον ξεχωριστό χώρο που μου παραχώρησε για τη συγγραφή της διατριβής, τον Γιώργο Ράδο και την Άντζελα Ρέγγη για όλες τις φορές που χρειάστηκα τη βοήθειά τους, όπως και την Ιωάννα Σπανού με τα ταξίδια της. Πολλές ευχαριστίες σε όλους όσους συνέβαλαν στην διαδικασία αξιολόγησης, αφιερώνοντας το χρόνο τους. Τέλος, ευχαριστώ την οικογένειά μου, για τη σημαντική υποστήριξη που μου πρόσφεραν όλα αυτά τα χρόνια.



# Περιεχόμενα

|                                                                                               |           |
|-----------------------------------------------------------------------------------------------|-----------|
| Περίληψη                                                                                      | i         |
| Abstract                                                                                      | iii       |
| Ευχαριστίες                                                                                   | v         |
| Περιεχόμενα                                                                                   | ix        |
| Κατάλογος Σχημάτων                                                                            | xiv       |
| Κατάλογος Πινάκων                                                                             | xv        |
| <b>1 Εισαγωγή</b>                                                                             | <b>1</b>  |
| 1.1 Ήχος και Επαυξημένα Περιβάλλοντα . . . . .                                                | 1         |
| 1.2 Παράγοντες διαμόρφωσης ηχητικού ερεθίσματος σε ARA περιβάλλοντα . . . . .                 | 3         |
| 1.3 Προσαρμοστικότητα στα Ηχητικά Συστήματα τύπου ARA . . . . .                               | 9         |
| 1.4 Επιστημονική Πρωτότυπη συμβολή της εργασίας . . . . .                                     | 13        |
| 1.5 Δομή Εργασίας . . . . .                                                                   | 13        |
| <b>2 Θεωρία και εφαρμογές ARA τεχνολογίας</b>                                                 | <b>15</b> |
| 2.1 Εισαγωγή . . . . .                                                                        | 15        |
| 2.2 Επισκόπηση τεχνολογίας σε ARA περιβάλλοντα . . . . .                                      | 15        |
| 2.2.1 Τεχνολογίες αναπαραγωγής πραγματικού περιβάλλοντος στα περιβάλλοντα τύπου ARA . . . . . | 16        |
| 2.2.2 Χωρική τοποθέτηση εικονικού περιβάλλοντος . . . . .                                     | 21        |
| 2.2.3 Τεχνολογίες διάδρασης σε ARA περιβάλλοντα . . . . .                                     | 24        |
| 2.3 Εφαρμογές ARA τεχνολογίας . . . . .                                                       | 25        |
| 2.4 Σύνοψη Κεφαλαίου . . . . .                                                                | 35        |
| <b>3 Αναλυτική μελέτη ARA μείξης</b>                                                          | <b>37</b> |
| 3.1 Εισαγωγή . . . . .                                                                        | 37        |
| 3.2 Αρχιτεκτονική παραδοσιακού μοντέλου ARA μείξης . . . . .                                  | 37        |
| 3.3 Περιορισμοί παραδοσιακού μοντέλου ARA μείξης . . . . .                                    | 39        |

|          |                                                                                                      |           |
|----------|------------------------------------------------------------------------------------------------------|-----------|
| 3.4      | Σύνοψη Κεφαλαίου . . . . .                                                                           | 41        |
| <b>4</b> | <b>Μεθοδολογία Ενσωμάτωσης Προσαρμοστικότητας στην ARA Μείξη</b>                                     | <b>43</b> |
| 4.1      | Εισαγωγή . . . . .                                                                                   | 43        |
| 4.2      | Γενικευμένο μοντέλο προσαρμοστικής ARA μείξης . . . . .                                              | 44        |
| 4.3      | Τεχνική Αμφιωτικής Επεξεργασίας Σημάτων Εισόδου ARA μοντέλου . . . . .                               | 46        |
| 4.4      | Εισαγωγή δεδομένων ακουστικής αντίληψης στην ARA μείξη . . . . .                                     | 48        |
| 4.4.1    | Υφιστάμενα μοντέλα πρόβλεψης Αμφιωτικής Ακουστότητας . . . . .                                       | 49        |
| 4.4.2    | Προτεινόμενη μέθοδος δημιουργίας προϋπολογισμένης μήτρας με βάση την Αμφιωτική Ακουστότητα . . . . . | 57        |
| 4.5      | Μεθοδολογία εκτίμησης Αμφιωτικής Ακουστότητας των σημάτων εισόδου . . . . .                          | 59        |
| 4.5.1    | Βαθμονόμηση Σημάτων Εισόδου . . . . .                                                                | 59        |
| 4.5.2    | Τεχνική υπολογισμού ενέργειας διέγερσης ηχητικών ροών . . . . .                                      | 61        |
| 4.5.3    | Αντιμετώπιση ακουστών παραμορφώσεων διεργασίας υπολογισμού καμύλης διέγερσης . . . . .               | 63        |
| 4.5.4    | Τεχνική Εκτίμησης Αμφιωτικής Ακουστότητας Σημάτων Εισόδου . . . . .                                  | 64        |
| 4.6      | Ανάλυση και Σύνθεση εικονικού περιβάλλοντος . . . . .                                                | 66        |
| 4.7      | Αλγόριθμος προσαρμογής περιβαλλόντων τύπου ARA . . . . .                                             | 69        |
| 4.7.1    | Αρχιτεκτονική αλγορίθμου προσαρμογής Αμφιωτικής Ακουστότητας . . . . .                               | 69        |
| 4.8      | Σύνοψη Κεφαλαίου . . . . .                                                                           | 72        |
| <b>5</b> | <b>Υλοποίηση Προσαρμοστικής ARA Μείξης Πραγματικού Χρόνου</b>                                        | <b>75</b> |
| 5.1      | Εισαγωγή . . . . .                                                                                   | 75        |
| 5.2      | Πειραματική Υλοποίηση . . . . .                                                                      | 76        |
| 5.2.1    | Υλοποίηση Αμφιωτικής Αναπαραγωγής Πραγματικού Περιβάλλοντος . . . . .                                | 77        |
| 5.2.2    | Διεργασίες Βαθμονόμησης σημάτων εισόδου . . . . .                                                    | 82        |
| 5.2.3    | Μέτρηση ενέργειας διέγερσης σημάτων εισόδου . . . . .                                                | 84        |
| 5.2.4    | Υλοποίηση Ανάλυσης και Σύνθεσης σήματος Εικονικού περιβάλλοντος . . . . .                            | 86        |
| 5.2.5    | Υλοποίηση Πίνακα μετατροπής Αμφιωτικής Ακουστότητας . . . . .                                        | 90        |
| 5.2.6    | Εκτίμηση Αμφιωτικής Ακουστότητας Σημάτων Εισόδου . . . . .                                           | 91        |
| 5.2.7    | Υλοποίηση Αλγορίθμου προσαρμογής . . . . .                                                           | 96        |
| 5.3      | Σύνοψη Κεφαλαίου . . . . .                                                                           | 97        |
| <b>6</b> | <b>Αξιολόγηση μοντέλου Προσαρμοστικής ARA Μείξης</b>                                                 | <b>99</b> |
| 6.1      | Εισαγωγή . . . . .                                                                                   | 99        |
| 6.2      | Μεθοδολογία Αξιολόγησης . . . . .                                                                    | 99        |
| 6.2.1    | Αντικειμενική Αξιολόγηση . . . . .                                                                   | 100       |
| 6.2.2    | Υποκειμενική Αξιολόγηση . . . . .                                                                    | 103       |
| 6.2.3    | Περίπτωση Ηχητικού Παιχνιδιού τύπου ARA . . . . .                                                    | 108       |
| 6.3      | Σύνοψη Κεφαλαίου . . . . .                                                                           | 110       |

|           |                                                                                                |            |
|-----------|------------------------------------------------------------------------------------------------|------------|
| <b>7</b>  | <b>Αποτελέσματα</b>                                                                            | <b>111</b> |
| 7.1       | Εισαγωγή . . . . .                                                                             | 111        |
| 7.2       | Αποτελέσματα αντικειμενικών μετρήσεων . . . . .                                                | 111        |
| 7.3       | Αποτελέσματα υποκειμενικής αξιολόγησης ARA μείξης . . . . .                                    | 116        |
| 7.4       | Αποτελέσματα υποκειμενικής αξιολόγησης Ηχητικού Παιχνιδιού τύπου ARA . . . . .                 | 124        |
| 7.5       | Σύνοψη Κεφαλαίου . . . . .                                                                     | 130        |
| <b>8</b>  | <b>Συμπεράσματα</b>                                                                            | <b>133</b> |
|           | <b>Βιβλιογραφία</b>                                                                            | <b>137</b> |
| <b>A'</b> | <b>Πίνακας Μετατροπής Αμφιωτικής Ακουστότητας - Συχνότητας -<br/>Στάθμης Ακουστικής Πίεσης</b> | <b>149</b> |
| <b>B'</b> | <b>Πηγαίος Κώδικας Υλοποίησης</b>                                                              | <b>153</b> |
| <b>Γ'</b> | <b>Γραφική Παρουσίαση Ήχων Αξιολόγησης</b>                                                     | <b>157</b> |
| <b>Δ'</b> | <b>Έγγραφα Ερευνητικής Μελέτης</b>                                                             | <b>163</b> |
| <b>Ε'</b> | <b>Συντομογραφίες - Αρκτικόλεξα -<br/>- Ακρωνύμια</b>                                          | <b>169</b> |
| <b>Γ'</b> | <b>Απόδοση Ξενόγλωσσων Όρων</b>                                                                | <b>171</b> |



# Κατάλογος Σχημάτων

|     |                                                                                                                                                                                              |    |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Συνεχές Εικονικότητας - Πραγματικότητας του Milgram (Τροποποιημένη Εικόνα από τη δημοσίευση του Paul Milgram και Fumio Kishino σελ.3) [83]                                                   | 2  |
| 1.2 | Ακουστική αντίληψη σε επαυξημένο περιβάλλον τύπου ARA                                                                                                                                        | 3  |
| 1.3 | Φυσιολογία Αυτιού (Τροποποιημένη εικόνα από το Βιβλίο Hole's Human Anatomy του D. N. Shier, J. Butler, R Lewis σελ.457) [117]                                                                | 5  |
| 1.4 | Εξωτερικό Αυτί (Τροποποιημένη εικόνα από το Βιβλίο του D. Hallowell σελ.46) [27]                                                                                                             | 6  |
| 1.5 | Μηχανισμός ακοής: Μετατροπή των δονήσεων αέρα σε υγρή μετατόπιση (Τροποποιημένη εικόνα από το βιβλίο του H. Lindsay, D. A. Norman σελ.133) [72]                                              | 6  |
| 1.6 | Καμπύλες Ίσης Ακουστότητας (Τροποποιημένη εικόνα από τη δημοσίευση του H. Fletcher, W. A. Munson σελ.394) [33]                                                                               | 9  |
| 1.7 | Αλυσίδα Ηχητικής Ψηφιοποίησης και Αναπαραγωγής                                                                                                                                               | 10 |
| 1.8 | Τυπική Αρχιτεκτονική ARA συστήματος                                                                                                                                                          | 11 |
| 1.9 | Γενική αρχιτεκτονική προσαρμοστικών ARA περιβαλλόντων                                                                                                                                        | 12 |
| 2.1 | Ακουστικά υποστηριζόμενα από τεχνολογία Hear-Through α)Τύπου In-Ear, β)Τύπου Closed-Back, γ)Τύπου Open-Back, δ)Τύπου Bone-Conduction                                                         | 20 |
| 3.1 | Γενικευμένο Μοντέλο ARA Μείζης (Τροποποιημένο σχήμα από το άρθρο του Aki Härmä σελ.625) [46]                                                                                                 | 38 |
| 4.1 | Γενικευμένο Μοντέλο Προσαρμοστικής ARA Μείζης                                                                                                                                                | 44 |
| 4.2 | Συχνοτική απόκριση ακουστικών τύπου In-Ear Philips SHN 2500 (Σχήμα από την εργασία του Jussi Rämö, σελ.43 [58])                                                                              | 48 |
| 4.3 | Διάγραμμα κλίμακας CB, ERB και 1/3 Οκτάβας (Τροποποιημένη εικόνα από τη δημοσίευση του Jacobsen F., Poulsen T., Rindel J. H., Gade A. C., Ohlrich, M. σελ.72 [56])                           | 52 |
| 4.4 | Διάγραμμα εξασθένισης για τον παράγοντα μετάδοσης του ήχου σε ελεύθερο πεδίο ( $a_0$ ) και κλειστό πεδίο ( $a_{0D}$ ). (Εικόνα από το βιβλίο του Zwicker Eberhard, Fastl Hugo σελ.226 [149]) | 53 |
| 4.5 | Διάγραμμα εξασθένισης για τον παράγοντα μετάδοσης του ήχου. (Μη τροποποιημένη εικόνα από το άρθρο του Brian R. Glasberg, Brian C. J. Moore στο περιοδικό της JAES σελ.333 [42])              | 53 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.6  | Υπολογισμός διέγερσης, Μοντέλο Zwicker. (α) Διάγραμμα στάθμης διέγερσης ως προς τιμή κρίσιμης ζώνης για θορύβους στενής ζώνης με προκαθορισμένη κεντρική συχνότητα και στάθμη ακουστικής πίεσης 60 dB SPL (Μη τροποποιημένη εικόνα από το βιβλίο του Hugo Fastl, Eberhard Zwicker σελ.168 [32]). (β) Διάγραμμα στάθμης διέγερσης ως προς τιμή κρίσιμης ζώνης για θορύβους ευρείας ζώνης με κεντρική συχνότητα 1 kHz και στάθμες ακουστικής πίεσης κρίσιμων ζωνών όπως καταγράφονται. (Μη τροποποιημένη εικόνα από το βιβλίο του Hugo Fastl, Eberhard Zwicker σελ.169 [32]). Η διακεκομμένη γραμμή απεικονίζει και στα δύο διαγράμματα το κατώφλι ησυχίας. | 54 |
| 4.7  | Διάγραμμα υπολογισμού καμπύλης διέγερσης ημιτόνου 1 KHz από το μοντέλο του Moore - Glasberg. (Μη τροποποιημένη εικόνα από τη δημοσίευση του Moore, B. C. J., Glasberg, B. R. (1983). σελ.752 [88])                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 55 |
| 4.8  | Μεθοδολογία Υπολογισμού Πίνακα Μετατροπής Ακουστικής Ενέργειας (dB SPL), Αμφιωτικής Ακουστότητας (Phons), Συχνότητας (Hz)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 59 |
| 4.9  | Διάγραμμα μεθοδολογίας υπολογισμού αμφιωτικής ακουστότητας                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 65 |
| 4.10 | Αρχιτεκτονική Τράπεζας Φίλτρων                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 67 |
| 4.11 | Συχνοτική Απόκριση Τράπεζας Φίλτρων                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 69 |
| 4.12 | Διάγραμμα Εισόδων-Εξόδων Συνάρτησης Προσαρμογής                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 70 |
| 5.1  | Αρχιτεκτονική μοντέλου προσαρμοστικής μείξης τύπου ARA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 76 |
| 5.2  | Επισκόπηση Υλοποίησης μοντέλου προσαρμοστικής μείξης τύπου ARA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 78 |
| 5.3  | Ακουστικά Αμφιωτικής Αναπαραγωγής - Philips SHN2500                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 79 |
| 5.4  | Μονάδα Προενίσχυσης Μικροφώνων Αμφιωτικής λήψης                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 79 |
| 5.5  | Μονάδα Ψηφιακής επεξεργασίας Ήχου Αμφιωτικής λήψης «Analog Devices EVAL ADAU1761»                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 80 |
| 5.6  | Περιβάλλον Υλοποίησης Φιλτραρίσματος Ακουστικού Καναλιού «Sigma Studio»                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 81 |
| 5.7  | Υλοποίηση φίλτρων αντιστάθμισης επίδρασης ακουστικού καναλιού: α) Υψηλοπερατό Φίλτρο, β) Φίλτρο εγκοπής, γ) Φίλτρο Κορυφής                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 81 |
| 5.8  | Διάγραμμα Αλγόριθμου Βαθμονόμησης Ηχητικού Σήματος                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 82 |
| 5.9  | Περιπτώσεις χρήσης Αλγορίθμου Βαθμονόμησης : α) Καθαρός Τόνος 1 kHz, β) Δοκιμαστικός Ήχος από Θόρυβο κίνησης σε δρόμο                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 83 |
| 5.10 | Διαγράμματα υπολογισμού Καμπύλης Διέγερσης για: α) Τόνο 1 kHz, β) Ηχητικό δείγμα Κίνησης Πόλης, γ) Ηχητικό δείγμα «Roboscream», δ) Ηχητικό δείγμα Λευκού θορύβου                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 85 |
| 5.11 | Συχνοτική Απόκριση Τράπεζας Φίλτρων                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 87 |
| 5.12 | Διαγράμματα υπολογισμού φασματικής απόκρισης της Τράπεζας Φίλτρων για: α) Τόνο 500 Hz, β) Τόνο 1 kHz, γ) Τόνο 4 kHz, δ) Τόνο 8 kHz, ε) Ηχητικό δείγμα «Roboscream», ζ) Ηχητικό δείγμα Λευκού θορύβου                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 88 |

|      |                                                                                                                                                                                                                                                                                              |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.13 | Διαγράμματα απόκρισης φάσης για καθυστέρηση ομάδας Τράπεζας Φίλτρων:<br>α)Χαμηλοδιαβατό φίλτρο 250 Hz, β)Ζωνοπερατό φίλτρο 500 Hz, γ)Ζωνοπερατό<br>φίλτρο 1 kHz, δ)Ζωνοπερατό φίλτρο 2 kHz, ε)Ζωνοπερατό φίλτρο 4 kHz, στ)<br>Ζωνοπερατό φίλτρο 8 kHz, ζ)Υψηλοπερατό φίλτρο 16 kHz . . . . . | 89  |
| 5.14 | Διαγράμματα εκτίμησης Αμφιωτικής Ακουστότητας μέσα από την Καμπύλη Δι-<br>έγερσης για μοντέλο AR-ION: α)Κίνηση Πόλης (69 dB SPL), β)Roboscream<br>(59 dB SPL), γ)Ομιλία (70 dB SPL) . . . . .                                                                                                | 92  |
| 5.15 | Διάγραμμα Μεθοδολογίας Σύγκρισης Μοντέλων AR-ION με GM . . . . .                                                                                                                                                                                                                             | 94  |
| 5.16 | Συγκριτικά Διαγράμματα Συσχέτισης μοντέλων AR-ION και GM: α) Κίνηση<br>Πόλης 69 dB SPL, β) Roboscream 59 dB SPL, γ) Ομιλία 70 dB SPL, δ)<br>Βιομηχανία 70 dB SPL . . . . .                                                                                                                   | 95  |
| 6.1  | Διάγραμμα Μεθόδου δημιουργίας δειγμάτων και υπολογισμός πίνακα διέγερσης<br>περιβαλλόντων για υλοποίηση μετρήσεων . . . . .                                                                                                                                                                  | 100 |
| 6.2  | Διάγραμμα Μεθόδου Εκτίμησης Ακουστότητας από πίνακα Διέγερσης . . . . .                                                                                                                                                                                                                      | 101 |
| 6.3  | Διάγραμμα Μεθόδου Μετρήσεων Αλγόριθμου Προσαρμογής . . . . .                                                                                                                                                                                                                                 | 102 |
| 6.4  | Διάγραμμα Μεθόδου Ανάλυσης μετρήσεων μοντέλου προσαρμογής . . . . .                                                                                                                                                                                                                          | 102 |
| 6.5  | Κύρια οθόνη ηλεκτρονικού ερωτηματολογίου . . . . .                                                                                                                                                                                                                                           | 104 |
| 6.6  | Οθόνη Τεχνογραφικών και Δημογραφικών στοιχείων ηλεκτρονικού ερωτημα-<br>τολογίου . . . . .                                                                                                                                                                                                   | 105 |
| 6.7  | Οθόνη Ψυχοακουστικών τεστ Ακουστικής Διάκρισης . . . . .                                                                                                                                                                                                                                     | 105 |
| 6.8  | Οθόνη Υποκειμενικών Τεστ τύπου MUSHRA . . . . .                                                                                                                                                                                                                                              | 106 |
| 6.9  | Βαθμονόμηση πραγματικού περιβάλλοντος αξιολόγησης . . . . .                                                                                                                                                                                                                                  | 107 |
| 6.10 | Διαδικασία Υποκειμενικής Αξιολόγησης . . . . .                                                                                                                                                                                                                                               | 108 |
| 6.11 | Οθόνη Ηχητικού παιχνιδιού Επαυξημένης Πραγματικότητας Ήχου . . . . .                                                                                                                                                                                                                         | 109 |
| 7.1  | Διαγράμματα Φασματικής Ανάλυσης περιβαλλόντων παραδειγμάτων εφαρμογής<br>αλγόριθμου προσαρμογής για τις περιπτώσεις 3.1, 3.2, 3.3, 3.4 με διαβάθμιση<br>μείξης 0, 3, 6 . . . . .                                                                                                             | 114 |
| 7.2  | Πεδίο ARA μείξης για σενάριο 3.1 . . . . .                                                                                                                                                                                                                                                   | 121 |
| 7.3  | Πεδίο ARA μείξης για σενάριο 3.2 . . . . .                                                                                                                                                                                                                                                   | 122 |
| 7.5  | Πεδίο ARA μείξης για σενάριο 3.3 . . . . .                                                                                                                                                                                                                                                   | 122 |
| 7.4  | Πεδίο ARA μείξης για σενάριο 3.4 . . . . .                                                                                                                                                                                                                                                   | 123 |
| 7.6  | Κατανομή Μέσης Τιμής Γενικού Συνόλου για κατηγορία προσαρμοστικής και<br>μη προσαρμοστικής μείξης σε Ηχητικό Παιχνίδι τύπου ARA . . . . .                                                                                                                                                    | 125 |
| 7.7  | Κατανομή Μέσων Τιμών Επιτυχίας Ηχητικού Παιχνιδιού τύπου ARA για κα-<br>τηγορίες με Διάκριση Εμπειρίας . . . . .                                                                                                                                                                             | 126 |
| 7.8  | Κατανομή Μέσων Τιμών Επιτυχίας Ηχητικού Παιχνιδιού τύπου ARA για κα-<br>τηγορίες Διαβαθμισμένης Μείξης . . . . .                                                                                                                                                                             | 127 |
| Γ'.1 | Κυματομορφή ήχου Κίνησης Πόλης . . . . .                                                                                                                                                                                                                                                     | 157 |
| Γ'.2 | Φασματική Ανάλυση ήχου Κίνησης Πόλης . . . . .                                                                                                                                                                                                                                               | 157 |

|                                                                                              |     |
|----------------------------------------------------------------------------------------------|-----|
| Γ'.3 Κυματομορφή ήχου Καφέ . . . . .                                                         | 158 |
| Γ'.4 Φασματική Ανάλυση ήχου Καφέ . . . . .                                                   | 158 |
| Γ'.5 Κυματομορφή ήχου Βιομηχανίας . . . . .                                                  | 158 |
| Γ'.6 Φασματική Ανάλυση ήχου Βιομηχανίας . . . . .                                            | 159 |
| Γ'.7 Κυματομορφή ήχου Miss . . . . .                                                         | 159 |
| Γ'.8 Φασματική Ανάλυση ήχου Miss . . . . .                                                   | 159 |
| Γ'.9 Κυματομορφή ήχου Roboscream . . . . .                                                   | 160 |
| Γ'.10 Φασματική Ανάλυση ήχου Roboscream . . . . .                                            | 160 |
| Γ'.11 Κυματομορφή ήχου Score . . . . .                                                       | 160 |
| Γ'.12 Φασματική Ανάλυση ήχου Score . . . . .                                                 | 161 |
| Δ'.1 Αίτηση Εκπόνησης Ερευνητικής Μελέτης . . . . .                                          | 165 |
| Δ'.2 Δήλωση Συγκατάθεσης Υποκειμένου στην Πειραματική Διαδικασία - Ελληνική Έκδοση . . . . . | 166 |
| Δ'.3 Δήλωση Συγκατάθεσης Υποκειμένου στην Πειραματική Διαδικασία - Αγγλική Έκδοση . . . . .  | 167 |

# Κατάλογος Πινάκων

|      |                                                                                                                                                     |     |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.1  | Κρίσιμες Ζώνες συχνότητας της Κλίμακας Bark . . . . .                                                                                               | 50  |
| 4.2  | Κρίσιμες Ζώνες Συχνοτήτων σε διαστήματα Οκτάβας . . . . .                                                                                           | 58  |
| 4.3  | Ζώνες Φασματικής Ανάλυσης για παράθυρο 1024 δειγμάτων με ρυθμό δειγματοληψίας 48 kHz . . . . .                                                      | 62  |
| 4.4  | Πίνακας Ζώνης συχνοτήτων φίλτρων για Τράπεζα Φίλτρων με επικάλυψη 50% . . . . .                                                                     | 67  |
| 5.1  | Πίνακας χαρακτηριστικών Τράπεζας Φίλτρων . . . . .                                                                                                  | 86  |
| 5.2  | Συγκριτικός Πίνακας Συσχέτισης μοντέλων AR-ION και GM . . . . .                                                                                     | 95  |
| 7.1  | Πίνακας συσχέτισης φάσματος επαυξημένων περιβαλλόντων για παραδείγματα 3.1, 3.2, 3.3, 3.4 με διαβάθμιση μείξης 0, 3, 6 . . . . .                    | 115 |
| 7.2  | Πίνακας υπολογισμών ενέργειας και Ακουστότητας (Μοντέλο GM) για περιπτώσεις παραδειγμάτων 3.1, 3.2, 3.3, 3.4 με διαβάθμιση μείξης 0, 3, 6 . . . . . | 116 |
| 7.3  | MUSHRA Τέστ 3.1 . . . . .                                                                                                                           | 117 |
| 7.4  | MUSHRA Τέστ 3.2 . . . . .                                                                                                                           | 118 |
| 7.5  | MUSHRA Τέστ 3.3 . . . . .                                                                                                                           | 118 |
| 7.6  | MUSHRA Τέστ 3.4 . . . . .                                                                                                                           | 119 |
| 7.7  | Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών για κατηγορίες με Διάκριση Εμπειρίας . . . . .                                                             | 128 |
| 7.8  | Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών και Ποσοστιαία Διαφορά για κατηγορίες Εμπειρίας . . . . .                                                  | 128 |
| 7.9  | Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών για κατηγορίες Διαβαθμισμένης Μείξης . . . . .                                                             | 129 |
| 7.10 | Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών και Ποσοστιαία Διαφορά για κατηγορίες με Διαβάθμιση μείξης . . . . .                                       | 129 |
| 7.11 | Αποτελέσματα Χαμένων Ευκαιριών Ηχητικού Παιχνιδιού τύπου ARA . . . . .                                                                              | 130 |
| 7.12 | Αποτελέσματα Μέσης Τιμής Ηχητικού Παιχνιδιού - t-Test . . . . .                                                                                     | 130 |
| A'1  | Μήτρα Συσχέτισης Αμφιωτικής Ακουστότητας ( <i>Phons</i> ), Συχνότητας ( <i>Hz</i> ) και Ενέργειας ( <i>dB SPL</i> ) . . . . .                       | 152 |



# Κεφάλαιο 1

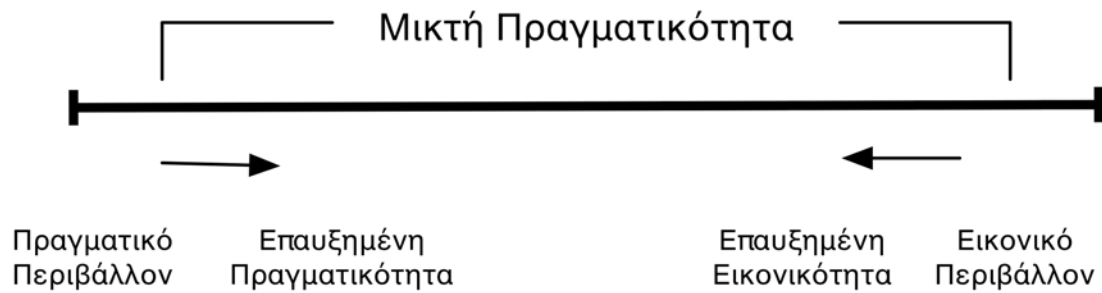
## Εισαγωγή

Στο εισαγωγικό κεφάλαιο της παρούσας διατριβής, παρουσιάζεται η σχέση του ήχου με τα επαυξημένα περιβάλλοντα έχοντας ως κοινή βάση την ακουστική αντίληψη για την δημιουργία του κατάλληλου ηχητικού ερεθίσματος. Ακολούθως, εισάγεται το ζήτημα της προσαρμοστικότητας μέσα στο ευρύτερο πεδίο των ηχητικών συστημάτων «Επαυξημένης Πραγματικότητας Ήχου» (Augmented Reality Audio - ARA). Τέλος δίνεται η επιστημονική πρωτότυπη συμβολή της εργασίας και συνολική δομή της διατριβής.

### 1.1 Ήχος και Επαυξημένα Περιβάλλοντα

Η προέλευση του ήχου στον πραγματικό κόσμο αποτελεί προϊόν φυσικής ή τεχνητής διεργασίας. Η εφεύρεση συσκευών καταγραφής και σύνθεσης του ήχου, επέκτεινε τις δυνατότητες χρήσης του. Ένα «προφανές» παράδειγμα είναι η εγγραφή ενός μουσικού κομματιού. Ως ηχητική έγγραφη, ορίζεται η αποθήκευσή των δονήσεων του αέρα που παράγονται από την ταλάντωση των μορίων του αέρα στο πεδίο του χρόνου σε κατάλληλο μέσο. Παράδειγμα μέσου ταλάντωσης των μορίων του αέρα, είναι ένα μουσικό όργανο. Παράδειγμα υλικού αποθήκευσης αποτελεί ο οπτικός δίσκος ή ο μηχανικός δίσκος του υπολογιστή. Η διαδικασία εγγραφής, δίνει τη δυνατότητα ανάκτησης της αποθηκευμένης πληροφορίας και αναπαραγωγής του. Μέσα από την αναπαραγωγή, το ηχητικό περιβάλλον που υπήρξε τη στιγμή της εγγραφής, μεταφέρεται στο περιβάλλον του ακροατή. Το παραπάνω παράδειγμα αποτελεί μια πρωταρχική προσέγγιση παρουσίας δύο διαφορετικών ηχητικών περιβαλλόντων στον ακροατή. Το εγγεγραμμένο ηχητικό περιβάλλον και το πραγματικό ηχητικό περιβάλλον μέσα στο οποίο βρίσκεται ο ακροατής.

Στα τέλη του 20ού αιώνα, άρχισε να καθορίζεται το πεδίο της επαυξημένης πραγματικότητας, με την εισαγωγή του όρου (Augmented Reality) από τον Caudell [125]. Στην πρωτότυπη εργασία του, ορίζει τις οθόνες επαυξημένης πραγματικότητας, οι οποίες ενσωματώνουν εικονική πληροφορία σε οθόνες κατάλληλα τοποθετημένες στο κεφάλι (Head-up Display - HUD). Την εικονική πληροφορία που ενσωματώνεται πάνω στον πραγματικό κόσμο, αποτελούν οδηγίες συναρμολόγησης ειδικού εξοπλισμού. Η θεμελίωση των όρων ωστόσο πραγματοποιήθηκε από τον Milgram λίγο αργότερα, με την παρουσίαση του συνεχούς «Μεικτής



Σχήμα 1.1: Συνεχές Εικονικότητας - Πραγματικότητας του Milgram (Τροποποιημένη Εικόνα από τη δημοσίευση του Paul Milgram και Fumio Kishino σελ.3) [83]

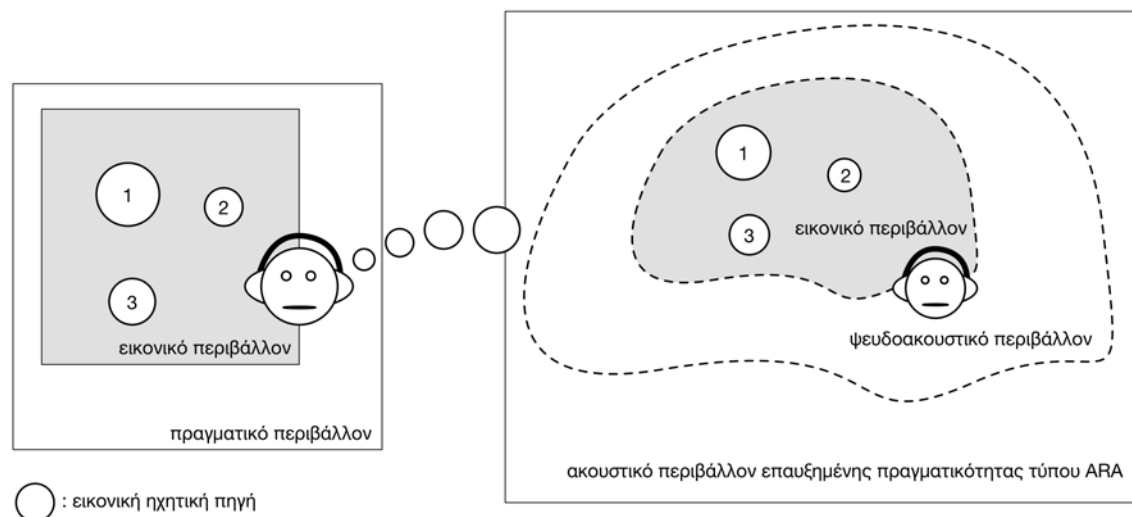
Πραγματικότητας» (βλπ Σχ. 1.1) [83]). Το συνεχές ορίζει την μικτή πραγματικότητα σαν μια αδιάσπαστη αλληλουχία «Εικονικότητας-Πραγματικότητας». Αναλυτικότερα, το συνεχές διαιρείται σε τέσσερα περιβάλλοντα: α) «Πραγματικό Περιβάλλον», β) «Εικονικό Περιβάλλον» τα οποία τοποθετούνται στα άκρα του συνεχούς, γ) «Επαυξημένη Πραγματικότητα» και δ) «Επαυξημένη Εικονικότητα» τα οποία βρίσκονται αντίστοιχα ενδιάμεσα στο συνεχές. Αναλύοντας το περιεχόμενο του κάθε περιβάλλοντος, το πραγματικό περιβάλλον αντικατοπτρίζει τον κόσμο που γίνεται αντιληπτός από τις αισθήσεις χωρίς κάποια άλλη απαίτηση. Η επαυξημένη πραγματικότητα, περιγράφει τον κόσμο τον οποίο για να γίνει αντιληπτός, χρειάζεται η χρήση βοηθητικού εξοπλισμού επέκτασης των αναγκαίων αισθήσεων στον εικονικό χώρο. Παράδειγμα αποτελεί η χρήση ειδικών γυαλιών με ενσωματωμένη οθόνη, η οποία παρέχει την εικονική πληροφορία τοποθετημένη χωροχρονικά στο πραγματικό περιβάλλον, χωρίς να απομονώνεται ο χρήστης από την πραγματικότητα. Στην δεξιά πλευρά του συνεχούς, υπάρχει το εικονικό περιβάλλον, όπως είναι για παράδειγμα η γραφική πληροφορία που παρουσιάζεται από μια οθόνη. Τέλος, η επαυξημένη εικονικότητα, ορίζει την ελάχιστη εμπλοκή του πραγματικού περιβάλλοντος μέσα στο εικονικό. Ως επαυξημένη εικονικότητα, μπορεί να χαρακτηριστεί το σενάριο στο οποίο το δάχτυλο του χρήστη ως πραγματικό περιβάλλον, εισέρχεται στο εικονικό περιβάλλον του κινητού μέσα από την οθόνη αφής.

Οι παραπάνω αναφορές, παρουσιάζουν τη γενική εικόνα της μεικτής πραγματικότητας. Από τη μεικτή πραγματικότητα, η παρούσα διατριβή θα εστιάσει στην περιοχή της «Επαυξημένης Πραγματικότητας» δίνοντας έμφαση στον ήχο και στο αισθητήριο όργανο της ακοής. Η «Επαυξημένη Πραγματικότητα Ήχου» (Augmented Reality Audio - ARA) όπως έχει καθιερωθεί, περιγράφει τη σύνδεση του πραγματικού περιβάλλοντος με το αντίστοιχο εικονικό ηχητικό του. Ο εικονικός κόσμος εκτείνεται στον τρισδιάστατο εικονικό χώρο, ο οποίος οριοθετείται σε σχέση με τον πραγματικό κόσμο χωρικά και χρονικά. Η «Επαυξημένη Πραγματικότητα Ήχου» δεν απομονώνει τον ακροατή από τον πραγματικό ηχητικό κόσμο, μετατρέποντας αυτόν σε ένα ηχητικό χαμβιά πάνω στον οποίο εισάγεται ο τρισδιάστατος εικονικός ηχητικός χώρος αποτελούμενος από ηχητικές πηγές, αναμιγνύεται με τους φυσικούς ήχους που προέρχονται από το πραγματικό περιβάλλον, και γίνεται αντιληπτό ως επαύξηση του πραγματικού κόσμου [128, 47]. Η εφαρμογή του παραπάνω μοντέλου λειτουργίας ARA χρησιμοποιείται σε ένα μεγάλο εύρος εφαρμογών όπως είναι οι εφαρμογές σοβαρού σκοπού, εκπαίδευσης, καλλιτε-

χνικές, πολιτισμού και διασκέδασης. Στη συγκεκριμένη διατριβή, παρουσιάζεται η ανάπτυξη ενός πρωτότυπου μοντέλου μείξης επαυξημένης πραγματικότητας ήχου με εφαρμογή στόχου, ένα ηχητικό παιχνίδι επαυξημένης πραγματικότητας ήχου.

## 1.2 Παράγοντες διαμόρφωσης ηχητικού ερεθίσματος σε ARA περιβάλλοντα

Η ακουστική αντίληψη μέσα στο ARA περιβάλλον, αποτελεί μια σύνθετη διεργασία η οποία επηρεάζεται από ποικίλους παράγοντες. Παρακάτω θα γίνει μια ανάλυση των παραγόντων οι οποίοι διαμορφώνουν το τελικό ηχητικό ερέθισμα, το οποίο είναι υπεύθυνο για την ακουστική αντίληψη. Επίσης, θα προσεγγιστεί το ζήτημα της ακουστικής αντίληψης μέσα από τη φυσιολογία του αισθητήριου οργάνου της ακοής και τη σχέση της με τα περιβάλλοντα τύπου ARA.



Σχήμα 1.2: Ακουστική αντίληψη σε επαυξημένο περιβάλλον τύπου ARA

Το περιβάλλον επαυξημένης πραγματικότητας τύπου ARA είναι το αποτέλεσμα χωροχρονικής αντιστοίχισης δύο περιβαλλόντων, του εικονικού και του πραγματικού, μέσα στο οποίο τοποθετείται ο ακροατής. Παρακάτω αναλύονται τα δύο περιβάλλοντα.

Το «Πραγματικό Περιβάλλον» ως ερέθισμα διαιρείται σε δύο μέρη, στο οπτικό και το ηχητικό. Αυτό το περιβάλλον γίνεται αντιληπτό με φυσικό τρόπο από τα αισθητήρια της ακοής και της όρασης. Ωστόσο λόγω της ανάγκης για ταυτόχρονη ύπαρξη του εικονικού ηχητικού περιβάλλοντος με το πραγματικό, έχει αναπτυχθεί η τεχνολογία τύπου ARA η οποία θα αναλυθεί στην Ενότητα 2.

Λόγω της αναγκαιότητας αναπαραγωγής των ηχητικών πηγών του εικονικού περιβάλλοντος με χρήση ακουστικού εξοπλισμού, το πραγματικό περιβάλλον μεταφέρεται στον ακουστικό πόρο διαμέσου ειδικού ακουστικού εξοπλισμού τύπου ARA (βλπ Σχ.1.2). Για να υπάρχει διαχωρισμός μεταξύ των εννοιών της αναπαραγωγής του πραγματικού περιβάλλοντος με χρήση ακουστικών τύπου ARA και χωρίς ειδικό εξοπλισμό (με ανοικτό ακουστικό πόρο), ο ήχος

του πραγματικού περιβάλλοντος στην πρώτη περίπτωση ορίζεται ως «Ψευδοακουστικό Περιβάλλον» [61, 47].

Το δεύτερο περιβάλλον ορίζεται ως εικονικό και μαζί με το ψευδοακουστικό, αποτελεί την επαύξηση του πραγματικού περιβάλλοντος. Το εικονικό περιβάλλον αποτελεί έναν τρισδιάστατο ηχητικό χώρο μέσα στον οποίο τα ηχητικά αντικείμενα μπορεί να αλληλεπιδρούν μεταξύ τους ή με τον θεατή εξαρτώμενα από το εκάστοτε σενάριο μέσα σε χρονικά και χωρικά πλαίσια.

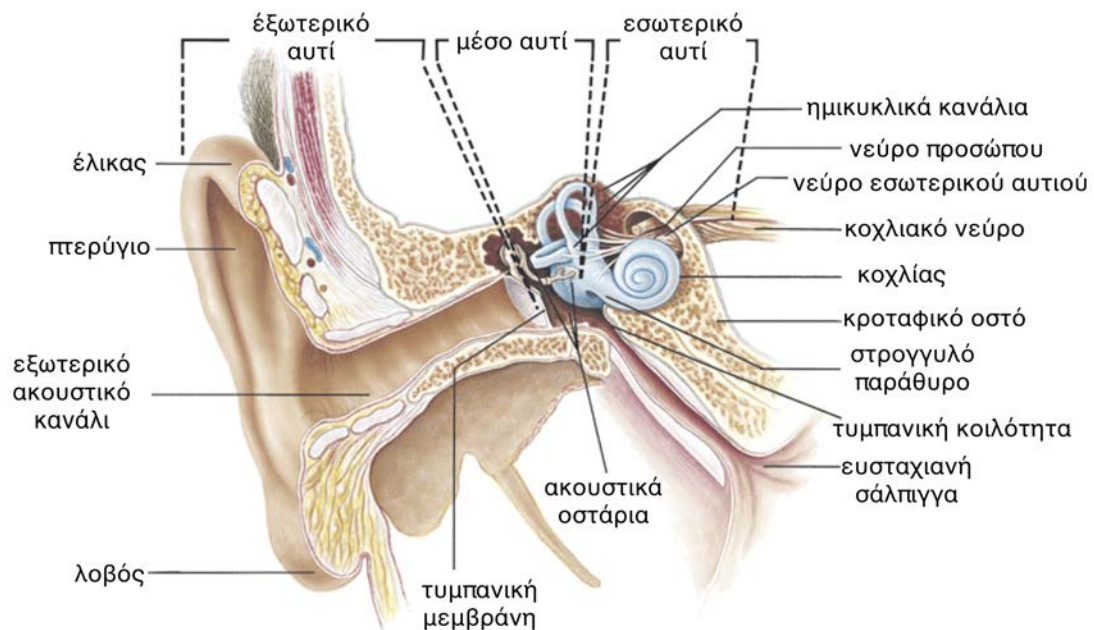
Καθώς η ακουστική αντίληψη του ARA περιβάλλοντος συνίσταται στο τελικό ηχητικό ερέθισμα που παράγεται από τη σύνδεση των δύο παραπάνω περιβαλλόντων, ο μηχανισμός της ακοής αποτελεί κομβικό σημείο για τον σχεδιασμό του ARA περιβάλλοντος. Η πρόβλεψη της επίδρασης του αισθητήριου της ακοής από το επαυξημένο ηχητικό σήμα διέγερσής του, ανάγει τη λειτουργία του αυτιού σε σημαντικό παράγοντα διαμόρφωσης του επαυξημένου ερεθίσματος.

Η μελέτη του αυτιού από άποψη κατασκευής, εντάσσεται στην επιστήμη της φυσιολογίας. Η αντίληψη του ήχου, αποτελεί κομμάτι της ψυχολογίας. Ωστόσο το διεπιστημονικό πεδίο της ψυχοακουστικής, είναι εκείνο που ενώνει τις δύο επιστήμες με σκοπό να δώσει ερμηνείες για την ανθρώπινη αντίληψη του ήχου, λαμβάνοντας υπόψη τη φυσική του ήχου και την φυσιολογία του ανθρώπινου αυτιού.

Ο ήχος ορίζεται ως μηχανική διαταραχή η οποία διαδίδεται με ορισμένη ταχύτητα μέσα σε ένα υλικό μέσο που μπορεί να αναπτύξει εσωτερικές δυνάμεις (π.χ. ελαστικότητας, εσωτερικής τριβής) και έχει τέτοιο χαρακτήρα, ώστε μπορεί να διεγείρει το αισθητήριο της ακοής και να προκαλέσει ακουστικό αίσθημα (ΕΛΟΤ 263.1).

Το αυτί αποτελεί έναν πολύπλοκο μηχανισμό ο οποίος ακόμα και στις μέρες μας δεν έχει αποκρυπτογραφηθεί πλήρως. Παρόλα αυτά, σε αυτό το κεφάλαιο παρουσιάζονται τα πιο σημαντικά μέρη του ανθρώπινου αυτιού. Η διαίρεσή του γίνεται σε τρία μέρη α) το εξωτερικό αυτί, β) το μέσο αυτί γ) το εσωτερικό αυτί. Παρακάτω παρουσιάζονται οι λειτουργίες του κάθε μέρους.

Το εξωτερικό αυτί αποτελείται από το περύγιο και τον έξω ακουστικό πόρο ο οποίος καταλήγει στο τύμπανο του αυτιού. Το περύγιο του αυτιού (Pinna) περικλείεται από το σκέλος της έλικας. Οι ακουστικές πτυχώσεις του μπορούν να αυξήσουν την ευαισθησία στους ήχους υψηλής συχνότητας όταν στρέφονται προς την πηγή τους, ελέγχοντας την ηχητική πληροφορία που προσπίπτει στο τύμπανο, και βοηθώντας στον εντοπισμό της πηγής ήχου [142]. Ακολούθως, μετά από τον ακουστικό μετασχηματισμό που παράγεται από τις ανακλάσεις στο ακουστικό περύγιο, ο ήχος διέρχεται μέσω του ακουστικού καναλιού στο τύμπανο. Το μήκος του εξωτερικού ακουστικού πόρου είναι στην περιοχή των 2,5 εκατοστών καθώς δεν είναι το ίδιο σε κάθε άνθρωπο, ξεκινάει από το περύγιο και καταλήγει στον τυμπανικό υμένα. Το εξωτερικό μέρος του ακουστικού πόρου αποτελείται από χόνδρο μεγέθους 8 χιλιοστών το μέγεθος του οποίου ομοίως με το προηγούμενο, είναι διαφορετικό σε κάθε άνθρωπο, και το υπόλοιπο μέρος του από οστό. Η συμπεριφορά του είναι αντίστοιχη με εκείνη ενός ηχητικού σωλήνα ανοικτού από τη μια πλευρά. Η αντήχηση που δημιουργείται μέσα σε αυτόν έχει σαν αποτέλεσμα την ενίσχυση των συχνοτήτων από 2 kHz έως 5 kHz για 5 dB ή και περισσότερο με μέγιστη ενίσχυση περίπου 11 dB να συμβαίνει γύρω από τα 4 kHz [143]. Η αλλαγή της πίεσης

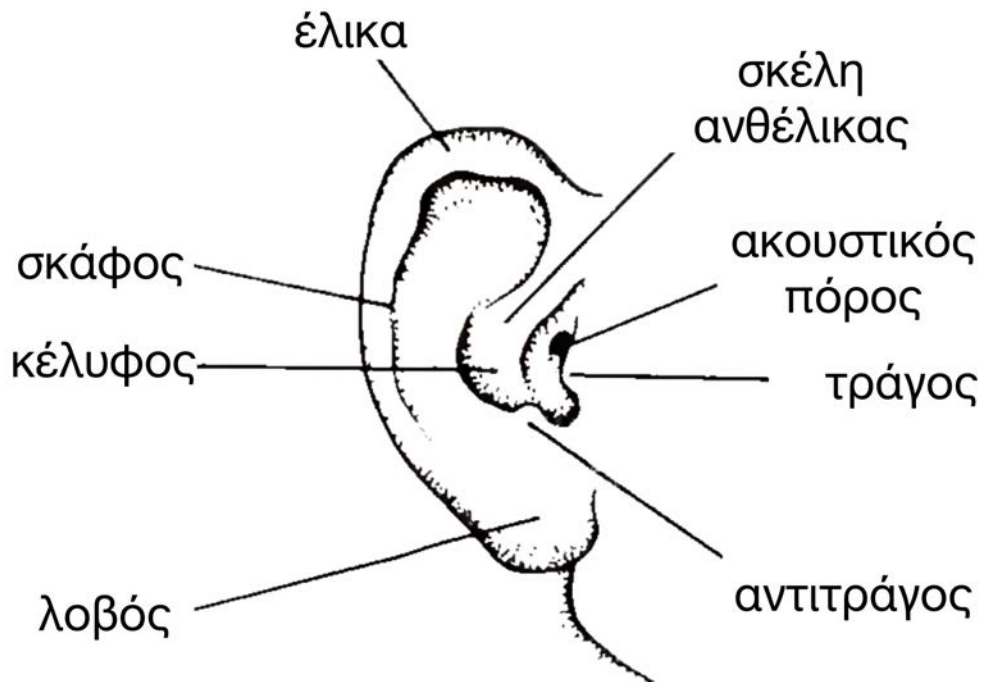


Σχήμα 1.3: Φυσιολογία Αυτιού (Τροποποιημένη εικόνα από το Βιβλίο Hole's Human Anatomy του D. N. Shier, J. Butler, R Lewis σελ.457) [117]

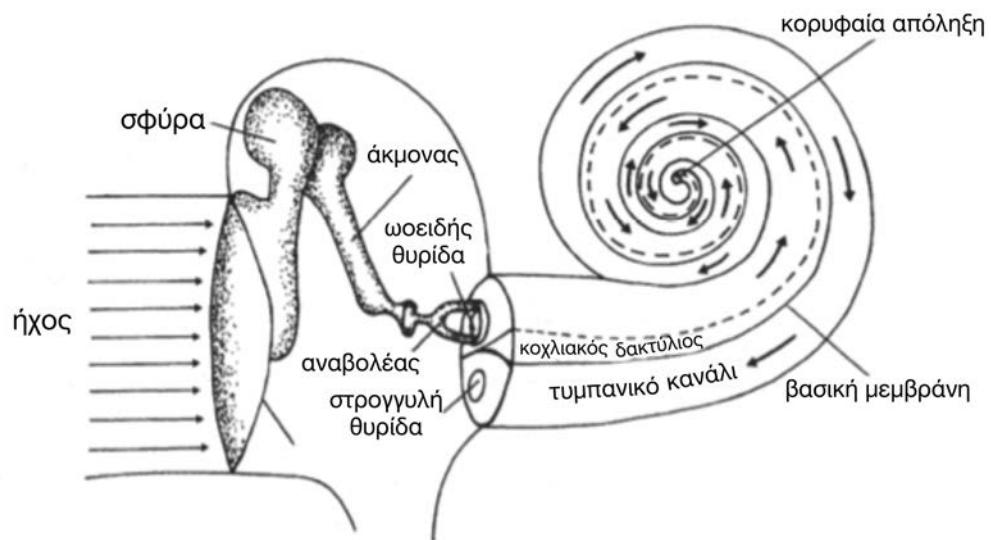
στο τέλος του καναλιού προκαλεί τη δόνηση στο τύμπανο. Αυτός ο κραδασμός συλλέγεται και μεταδίδεται από μια αλυσίδα τριών μικρών οστών που βρίσκονται στο μέσο αυτί.

Το μέσο αυτί περιλαμβάνει το μηχανικό σύστημα το οποίο μεταδίδει την ηχητική πληροφορία προερχόμενη από το εξωτερικό στο εσωτερικό αυτί διεγείροντας τα τριχωτά κύτταρα που βρίσκονται στον κοχλία. Το μηχανικό σύστημα αποτελείται από τρία μικρά οστά. Το πρώτο από αυτά τα οστά, η σφύρα είναι προσαρτημένη στο τύμπανο και η μετακίνησή της μεταδίδεται στον άκμονα και από εκεί στον αναβολέα. Τα οστάρια του αναβολέα συνδέονται με το ωοειδές παράθυρο στη βάση του κοχλία. Αυτό το παράθυρο αποτελεί ένα όριο που χωρίζει το μέσο από το εσωτερικό αυτί.

Το εσωτερικό αυτί αποτελεί μια πολύπλοκη κατασκευή εξαιτίας της οποίας έχει ονομαστεί λαβύρινθος. Μέσα στο λαβύρινθο κυκλοφορεί το υγρό που λέγεται λέμφος. Τα κυρίως τμήματα του εσωτερικού αυτιού είναι ο κοχλίας, η αίθουσα και οι τρεις ημικυκλικοί σωλήνες. Στο εσωτερικό τοίχωμα του τυμπάνου στο αιθουσαίο τμήμα, υπάρχουν η ωοειδής και η στρογγυλή θυρίδα οι οποίες βρίσκονται στην αρχή του κοχλία. Ο κοχλίας αποτελείται από έναν ελικοειδή σωλήνα μέσα στον οποίο βρίσκεται το αισθητήριο όργανο της ακοής ή όργανο του «Corti». Επίσης περιέχει τον βασικό υμένα (Basilar Membrane) ο οποίος είναι υπεύθυνος για τον διαχωρισμό των ήχων στο πεδίο της συχνότητας. Τα οστάρια μεταδίδουν το μηχανικό ερέθισμα του ηχητικού κύματος στον κοχλία στον οποίο, μέσω των υδραυλικών κυμάτων που δημιουργούνται στο εσωτερικό υγρό του (λέμφο), και την επίδραση του βασικού υμένα, δημιουργείται



Σχήμα 1.4: Εξωτερικό Αυτί (Τροποποιημένη εικόνα από το Βιβλίο του D. Hallowell σελ.46) [27]



Σχήμα 1.5: Μηχανισμός ακοής: Μετατροπή των δονήσεων αέρα σε υγρή μετατόπιση (Τροποποιημένη εικόνα από το βιβλίο του H. Lindsay, D. A. Norman σελ.133) [72]

διέγερση του οργάνου του «Corti», με αποτέλεσμα τον ηλεκτρικό παλμό ο οποίος μεταφέρεται μέσω του ακουστικού νεύρου στα αντίστοιχα εγκεφαλικά κέντρα του ακουστικού φλοιού του

εγκεφάλου.

Σύμφωνα με τα παραπάνω, η φυσιολογία της ακοής αποτελεί κοινό παρονομαστή για όλους τους ανθρώπους σε αντίθεση με την ακουστική αντίληψη. Η ακουστική αντίληψη αποτελεί ένα πολύπλοκο και υποκειμενικό στοιχείο της ακοής το οποίο καθορίζεται από ακουστικούς και μη ακουστικούς παράγοντες [17], όπως, για παράδειγμα, είναι η διαφορετικότητα του ακουστικού οργάνου σε κάθε άνθρωπο, το περιβάλλον του κάθε ακροατή, η εμπειρία η ηλικία ή άλλοι προσωπικοί παράγοντες [4]. Ο εισερχόμενος ήχος φιλτράρεται διαφορετικά μέσα στο αυτί του κάθε ανθρώπου καθώς δεν είναι πανομοιότυπα όλα τα αυτιά. Η ακουστική αντίληψη έχει προσαρμοστεί αντίστοιχα στα αισθητήρια του κάθε ανθρώπου. Η χρήση προσωποποιημένων συστημάτων βέβαια αποτελεί μια πολύπλοκη λύση.

Η φυσιολογία της ακοής, στο πεδίο της ARA, αποτελεί παράγοντα διαμόρφωσης της τεχνολογίας και της μεθόδου που χρησιμοποιείται για την αναπαραγωγή του πραγματικού και του εικονικού περιβάλλοντος. Στην περίπτωση χρήσης ARA τεχνολογίας με μέθοδο αμφιωτικής εγγραφής (βλπ Εν. 2.2.1), για την πιστότερη αναπαραγωγή του ψευδοακουστικού περιβάλλοντος, χρειάζεται να ληφθούν υπόψιν οι ανακλάσεις του εξωτερικού αυτιού, κατά συνέπεια θα πρέπει να καταγραφούν και να εφαρμοστούν στη σύνθεση του ψευδοακουστικού περιβάλλοντος. Επίσης η μεταβολή της ακουστικής του μέσου αυτιού από τη χρήση του αναγκαίου εξοπλισμού, θα πρέπει να ισοσταθμιστεί κατάλληλα με την εισαγωγή ή αφαίρεση των κατάλληλων συχνοτικών αντηχήσεων. Γενικά, οποιαδήποτε μεταβολή δημιουργηθεί η οποία επηρεάζει την φυσιολογία του αυτιού, θα πρέπει να αντισταθμίζεται. Λαμβάνοντας υπόψιν τους παραπάνω παράγοντες, είναι δυνατή η πιστή αμφιωτική αναπαραγωγή του πραγματικού περιβάλλοντος ως ψευδοακουστικό.

Η αναπαραγωγή του εικονικού περιβάλλοντος, χαρακτηρίζεται από την τοποθέτηση των ηχητικών πηγών στον τρισδιάστατο εικονικό ηχητικό χώρο. Κατά συνέπεια, ο εικονικός ήχος, λαμβάνει όλα εκείνα τα χαρακτηριστικά που χρειάζεται το ανθρώπινο αυτί ώστε να αντιληφθεί ο ανθρώπινος εγκεφάλος τη χωρική θέση της εκάστοτε πηγής στον τρισδιάστατο εικονικό χώρο. Για να επιτευχθεί το παραπάνω, έχουν δημιουργηθεί αμφιωτικές τεχνολογίες αναπαράστασης ηχητικών πηγών οι οποίες θα αναλυθούν στην Ενότητα 2.2.2.

Ένας επιπλέον παράγοντας διαμόρφωσης του ηχητικού ερεθίσματος στα περιβάλλοντα τύπου ARA, αποτελεί η ηχητική σχεδίαση του εικονικού περιβάλλοντος στην οποία λαμβάνεται υπόψιν και η διάδραση ακροατή με το επαυξημένο περιβάλλον. Σε αυτή την περίπτωση, χρησιμοποιείται το πεδίο της ηχητικής σχεδίασης της διάδρασης (Sonic Interaction Design), το οποίο καθορίζει το ηχητικό ερέθισμα και την πληροφορία που φέρει για να επιτευχθεί το αντίστοιχο ερέθισμα στον ακροατή και η αλληλεπίδρασή του με το σύστημα. Εργαλείο του ηχητικού σχεδιασμού αποτελούν οι μέθοδοι ηχοποίησης δεδομένων (Audification). Οι μέθοδοι ηχοποίησης, είναι αρκετά διαδεδομένοι με εκτενή βιβλιογραφία [111, 66, 52, 112].

Η ακουστική αντίληψη σε περιβάλλοντα τύπου ARA αποτελεί τον επόμενο παράγοντα όπου επηρεάζει τη διαμόρφωση του ηχητικού ερεθίσματος. Η δημιουργία του ακουστικού ερεθίσματος/διέγερσης του αισθητήριου οργάνου της ακοής, χρειάζεται να προσαρμόζεται μέσα στα όρια τις ακουστικής αντίληψης. Το τελικό ηχητικό ερέθισμα, διέρχεται μέσα από αισθητήριο όργανο της ακοής μέχρι να ενεργοποιήσει την αντίστοιχη περιοχή του εγκεφάλου.

Παρακάτω γίνεται μια σύντομη ανάλυση του αντιληπτικού μηχανισμού της ακοής, ο οποίος αποτελεί αφετηρία για τον σχεδιασμό επαυξημένων περιβαλλόντων τύπου ARA.

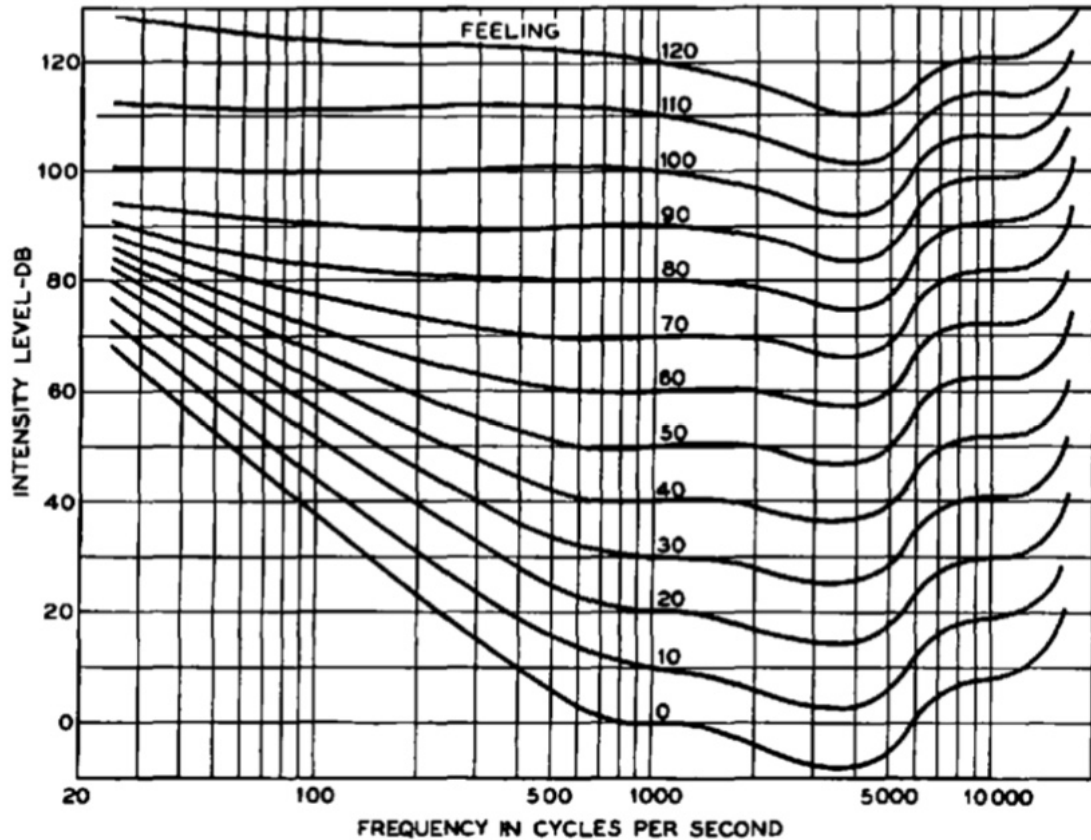
Είναι γνωστό πως η ανθρώπινη ακοή έχει όρια τα οποία βρίσκονται μέσα στο συχνοτικό εύρος μεταξύ των 20 Hz - 20 kHz. Το συγκεκριμένο όριο είναι γενικά αποδεκτό σε μικρές ηλικίες, ωστόσο σε μεγαλύτερες, παρατηρείται το φαινόμενο της απώλειας της ακοής, με αποτέλεσμα, να περιορίζεται το συχνοτικό εύρος και ειδικότερα να χάνονται υψηλές συχνότητες [22]. Οι παράγοντες απώλειας της ακοής, αποτελούν ένα ανοικτό ερευνητικό πεδίο, καθώς δεν έχουν σχέση μόνο με ηλικία, αλλά μπορεί να είναι ψυχολογικοί [64], θέμα φύλου [48], ασθένειας [124], γενετικοί [51] ή να οφείλονται και σε κάποια άγνωστη μέχρι στιγμής αιτία. Συνεπώς, λόγω του υποκειμενικού χαρακτήρα αντίληψης της ακοής, έχουν καθορισθεί μέθοδοι εύρεσης των υποκειμενικών ορίων με βάση την ηχητική στάθμη που αντιλαμβάνεται ο άνθρωπος μέσα στο παραπάνω συχνοτικό εύρος. Χαρακτηριστικό παράδειγμα αποτελεί το ελάχιστο κατώφλι ακουστότητας το οποίο περιγράφει την ελάχιστη ποσότητα ενέργειας σε συνάρτηση της συχνότητας που χρειάζεται για να γίνει αντιληπτός ο ήχος από τον άνθρωπο. Ο κάθε άνθρωπος έχει το δικό του ελάχιστο κατώφλι ακουστότητας.

Εκτός από το παραπάνω κατώφλι ακουστότητας, υπάρχει και το ανώτερο όριο πόνου το οποίο ορίζεται από τις μέγιστες στάθμες σε συγκεκριμένες συχνότητες που μπορεί να αντέξει ο άνθρωπος. Ο μέσος όρος των απόλυτων κατωφλίων ακουστότητας από ένα μεγάλο αριθμό ατόμων, έχει ως αποτέλεσμα το κανονικοποιημένο κατώφλι ακουστότητας. Το ακουστικό πεδίο του ανθρώπου τοποθετείται μεταξύ του ανώτερου ορίου πόνου και του ελάχιστου κατωφλίου ακουστότητας μέσα στο συχνοτικό φάσμα των 20 Hz - 20 kHz.

Η στάθμη ακουστικής πίεσης ενός ηχητικού συμβάντος είναι διαφορετική από την υποκειμενική στάθμη ακουστικής αντίληψης του ακροατή. Για να περιγραφεί το πόσο δυνατά ή πόσο χαμηλά ακούγεται ένας ήχος, έχει δημιουργηθεί το μέτρο της ακουστότητας. Η ακουστότητα είναι ένας ψυχολογικός όρος που χρησιμοποιείται για να περιγραφεί το μέγεθος της ακουστικής αίσθησης [33]. Η ακουστότητα εξαρτάται από την ένταση, το συχνοτικό φάσμα και το χρονικό παράθυρο του ηχητικού ερεθίσματος και συνήθως χρησιμοποιείται για να περιγράψει τη συνολική ακουστότητα ενός ήχου. Παράδειγμα ανάλυσης της ακουστότητας αποτελεί η στιγμιαία ακουστότητα (Specific Loudness), η οποία αποτελεί την ακουστότητα για καθορισμένες ακουστικές ζώνες σε μια δεδομένη χρονική στιγμή.

Ο υπολογισμός της ακουστότητας παράγει υποκειμενικά αποτελέσματα, καθώς αυτά αποτελούν υποθέσεις για το πως γίνεται αισθητός ο ήχος και επιδέχονται ερμηνείας. Ένα παράδειγμα εκτίμησης της ακουστότητας, αποτελεί οι καμπύλες ίσης ακουστότητας (βλπ Σχ. 1.6) οι οποίες απεικονίζουν τους καθαρούς τόνους που παράγουν την ίδια ακουστική αίσθηση σε συνάρτηση της συχνότητας [55] σύμφωνα με στατιστικά στοιχεία.

Ωστόσο για τον υπολογισμό της ακουστότητας υπάρχουν διαφορετικοί τύποι τεστ ανάλογα με τον τρόπο που γίνεται η ακρόαση για την εκάστοτε μέτρηση. Τέτοια παραδείγματα είναι η μονοωτική ακρόαση η οποία γίνεται με το ένα αυτί, η διωτική στην οποία γίνεται ακρόαση κοινού σήματος με χρήση και των δύο αυτιών και η διχωτική στην οποία ο ακροατής λαμβάνει διαφορετικό ηχητικό σήμα στο κάθε αυτί [87]. Στα περιβάλλοντα τύπου ARA υπάρχει ενδιαφέρον για την διχωτική ακρόαση.



Σχήμα 1.6: Καμπύλες Ίσης Ακουστότητας (Τροποποιημένη εικόνα από τη δημοσίευση του H. Fletcher, W. A. Munson σελ.394) [33]

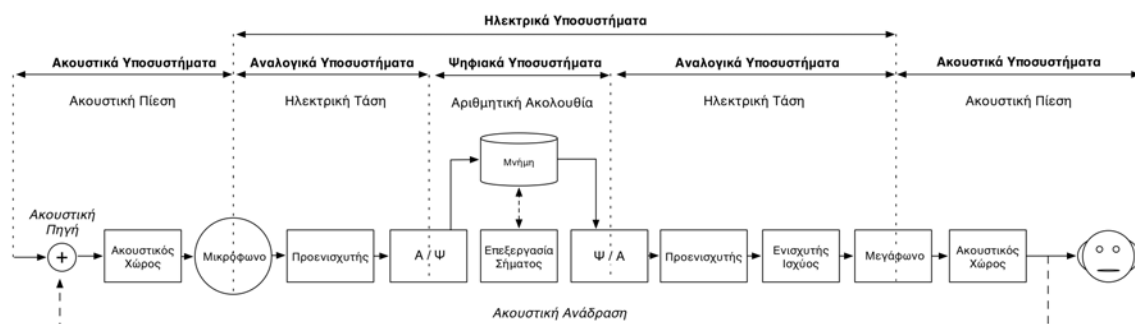
Είναι σημαντική η γνώση της λειτουργίας του ακουστικού οργάνου και της ακουστικής αντίληψης στην ακουστική διέγερση που υποβάλλεται μέσα στα περιβάλλοντα τύπου ARA ώστε να υπάρχει το επιθυμητό εμβυθιστικό αποτέλεσμα καθώς και η βέλτιστη διατήρηση της αυθεντικότητας. Η επεξεργασία του εκάστοτε ηχητικού περιβάλλοντος μπορεί να θέσει σε κίνδυνο την αυθεντικότητα του ηχητικού περιεχομένου στο εικονικό περιβάλλον [40] αλλά και στο πραγματικό [11] με επίδραση και στην εμβύθιση του ακροατή. Επιπροσθέτως, πρέπει να σημειωθεί πως υπάρχουν μοντέλα πρόβλεψης της αμφιωτικής ακουστότητας για τα σήματα διέγερσης του ακουστικού οργάνου τα οποία θα παρουσιαστούν στην Ενότητα 4.4.1.

### 1.3 Προσαρμοστικότητα στα Ηχητικά Συστήματα τύπου ARA

Η έννοια «Ηχητικό Σύστημα» περιλαμβάνει το ευρύ φάσμα συστημάτων εγγραφής, επεξεργασίας και αναπαραγωγής ήχου, καθώς και τον συνδυασμό αυτών. Μια τυπική αλυσίδα ηχητικής ψηφιοποίησης και αναπαραγωγής ως παράδειγμα ηχητικού συστήματος, αποτελείται από ένα σύνολο ηχητικών υποσυστημάτων (βλπ Σχ. 1.7). Αναλυτικά, δομείται σε δυο κατηγορίες

υποσυστημάτων, α) το ακουστικό υποσύστημα, β) τα ηλεκτρικά υποσυστήματα.

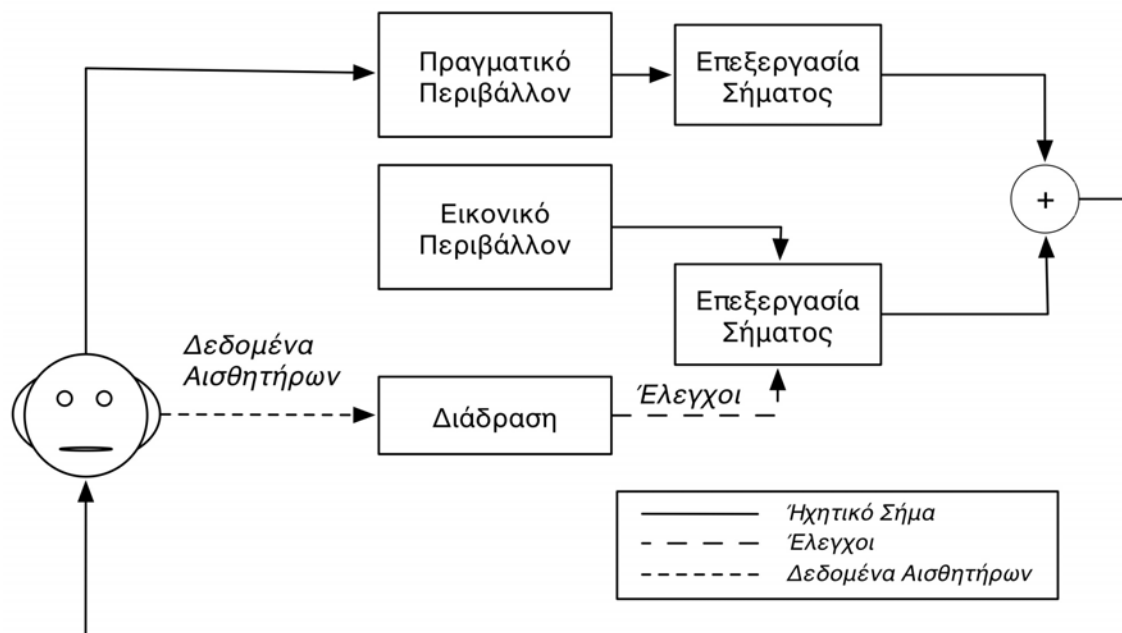
Το ακουστικό υποσύστημα αποτελεί την πύλη εισόδου και εξόδου του ήχου. Κατά την είσοδο, η ακουστική πηγή διέρχεται από τον ακουστικό χώρο και μέσω του μικροφώνου, η ακουστική πίεση μετατρέπεται σε ηλεκτρική τάση. Η παραγόμενη τάση διέρχεται κατόπιν από τα ηλεκτρικά υποσυστήματα. Τα ηλεκτρικά υποσυστήματα διαιρούνται σε αναλογικά και ψηφιακά. Το αναλογικά υποσύστημα εισόδου, μεταφέρουν την ηλεκτρική τάση στη μονάδα μετατροπής αναλογικού σήματος σε ψηφιακό (Α/Ψ). Το αποτέλεσμα της ψηφιοποίησης, μέσω της δειγματοληψίας και του κβαντισμού, είναι η αριθμητική ακολουθία η οποία αντιπροσωπεύει το σήμα στον ψηφιακό κόσμο και αποθηκεύεται στη μνήμη του συστήματος. Κατά την επεξεργασία του σήματος, η ανάκτηση και αποθήκευση της ψηφιακής πληροφορίας του ήχου γίνεται από τη μονάδα μνήμης. Η μονάδα μετατροπής ψηφιακής πληροφορίας σε αναλογική (Ψ/Α) μετατρέπει την αποθηκευμένη αριθμητική ακολουθία του ηχητικού σήματος σε ηλεκτρική τάση. Τα αναλογικά υποσυστήματα εξόδου, μεταφέρουν την ηλεκτρική τάση στη μονάδα του μεγαφώνου μέσω του οποίου μετατρέπεται σε ακουστική πίεση. Το τελικό ηχητικό σήμα διέρχεται από τον ακουστικό χώρο για να φτάσει στον ακροατή.



Σχήμα 1.7: Αλυσίδα Ηχητικής Ψηφιοποίησης και Αναπαραγωγής

Το παραπάνω σύστημα, χρησιμοποιείται ως γενική αρχή από αρκετές εφαρμογές όπως είναι και εκείνες του τύπου ARA. Ωστόσο στο σύστημα τύπου ARA, συνυπάρχει με το ηχητικό σύστημα και ο παράγοντας της διάδρασης μεταξύ του ακροατή και του επαυξημένου περιβάλλοντος (βλπ Σχ. 1.8). Η διάδραση συναντάται μέσα στα επαυξημένα ηχητικά περιβάλλοντα με ποικίλους τρόπους. Χαρακτηριστικό παράδειγμα είναι η χωρική τοποθέτηση του ακροατή στο τρισδιάστατο επαυξημένο περιβάλλον ήχου. Μέσω της χωρικής τοποθέτησης, ο χρήστης αλληλεπιδρά με το εικονικό περιβάλλον καθώς μεταβάλλει τη θέση του στον πραγματικό χώρο και, ακολούθως, την θέση του στον εικονικό κόσμο, επηρεάζοντας παραμέτρους της ηχητικής σύνθεσης των εικονικών ηχητικών πηγών όπως είναι η ηχητική στάθμη τους ή η κατευθυντικότητα τους, καθώς και την ακουστική του εικονικού κόσμου.

Η προσαρμοστικότητα ως ιδιότητα αποτελεί έναν πολύπλευρο όρο μέσα σε ένα ηχητικό σύστημα. Μπορεί να συναντηθεί σε ηχητικά συστήματα επεξεργασίας ηχητικού σήματος. Παραδείγματα αποτελούν τα προσαρμοστικά ηχητικά εφέ (Adaptive Digital Audio Effects - ADAFx) [134], στα οποία ο μετασχηματισμός του ηχητικού σήματος βασίζεται σε εξόρυξη αντιληπτικών χαρακτηριστικών από το σήμα εισόδου, με σκοπό την εισαγωγή και επεξεργασία τους από ειδικές συναρτήσεις αντιστοίχισης για την δημιουργία προσαρμοστικών ψηφιακών



Σχήμα 1.8: Τυπική Αρχιτεκτονική ARA συστήματος

ηχητικών εφέ. Μια άλλη κατηγορία είναι η πολυκάναλη ηχητική μείξη [109], στην οποία ακολουθείται η ίδια μεθοδολογία με το προηγούμενο παράδειγμα προσαρμοστικής επεξεργασίας, αλλά λαμβάνοντας υπόψιν επιπλέον παραμέτρους με αυξημένη πολυπλοκότητα, λόγω επεξεργασίας πολυκαναλικών ηχητικών σημάτων.

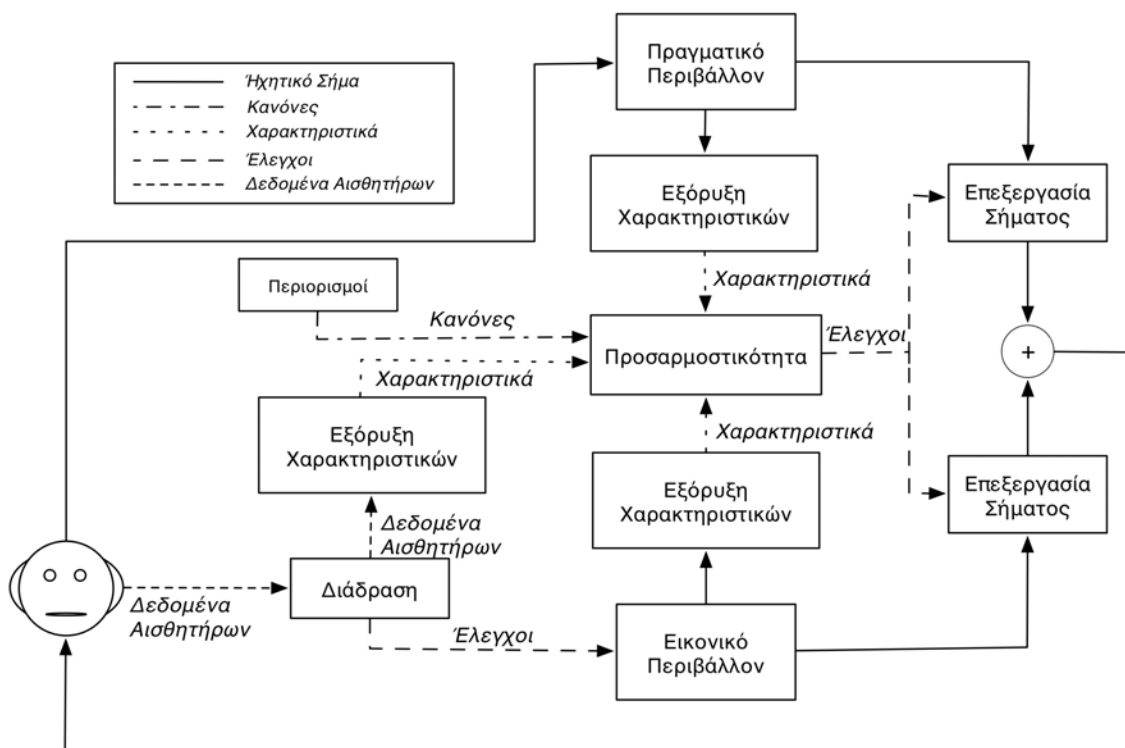
Στο πεδίο της ARA, η προσαρμοστικότητα εμφανίζεται συχνά στο περιεχόμενο του σεναρίου, σύμφωνα με δεδομένα που προέρχονται από τον χρήστη. Χαρακτηριστικό παράδειγμα είναι το σύστημα που χρησιμοποιήθηκε σε έκθεση [147] το οποίο προσαρμόζε το σενάριο στους στόχους του θεατή, στη γνώση του και σε ενδιαφέροντά του. Παρόμοια λειτουργία με σκοπό την αύξηση της εμπύθισης συναντάται και από τον Eckel [30], ο οποίος προσαρμόζει το περιεχόμενο ενός περιβάλλοντος ARA σύμφωνα με την χωρική συμπεριφορά του χρήστη. Επιπλέον, αρκετές εργασίες εστιάζουν στο ηχητικό περιεχόμενο [23, 69], το οποίο προσαρμόζεται στις επιθυμίες του ακροατή και σε χρήση σε μουσείο [136].

Εκτός από το περιεχόμενο, υπάρχουν και κάποιες επιπλέον τεχνολογικές προσεγγίσεις εισαγωγής της προσαρμοστικότητας στα ARA περιβάλλοντα. Αυτές εστιάζουν στην φυσική αναπαραγωγή του πραγματικού περιβάλλοντος με τη χρήση προσαρμοστικών φίλτρων για την ισοστάθμιση ειδικού ζεύγους ακουστικών [106], καθώς και αντίστοιχα την προσαρμοστική ισοστάθμιση εικονικών πηγών με τη χρήση μεθόδου ενεργητικής ακύρωσης θορύβου (Active Noise Cancellation) και βοηθούμενη ακρόαση με τη χρήση ακουστικών [133, 107].

Στα ARA περιβάλλοντα, σύμφωνα με τα παραπάνω, είναι πιθανόν να υπάρχει σύγχυση μεταξύ διαδραστικότητας και προσαρμοστικότητας. Ωστόσο, μπορεί να ορισθεί ότι η διαδραστικότητα σε ARA σενάρια παρέχει αίσθηση διάθεσης και ατμόσφαιρας μέσα από την προσαρμοστικότητα του ARA περιβάλλοντος.

Στο Σχήμα 1.9 παρουσιάζεται η γενικευμένη αρχιτεκτονική για συστήματα ήχου ARA με

εισαγωγή της προσαρμοστικότητας παράλληλα με την υφιστάμενη διάδραση. Ένα προσαρμοστικό ηχητικό σύστημα ARA αποτελείται αρχικά από τα περιβάλλοντα της επαυξημένης πραγματικότητας τα οποία είναι το πραγματικό και το εικονικό. Ο θεατής λαμβάνει όμοια με την τυπική αρχιτεκτονική το ηχητικό σήμα του πραγματικού περιβάλλοντος σε μείξη με το εικονικό περιβάλλον. Η προσαρμοστικότητα συνδυάζει ένα σύνολο προδιαγεγραμμένων κανόνων με ένα ή και περισσότερα μέρη του ηχητικού συστήματος ARA. Τα μέρη αυτά χωρίζονται στο πραγματικό περιβάλλον, το εικονικό και τη διάδραση. Σε όλες τις περιπτώσεις, υπάρχει εξόρυξη καθορισμένων χαρακτηριστικών, τα οποία αποτελούν δεδομένα εισόδου στη διεργασία της προσαρμοστικότητας. Οι κανόνες μπορεί να στηρίζονται για παράδειγμα σε ψυχοακουστικές μεθόδους, σε μηχανική μάθηση, ή σε τεχνητή νοημοσύνη. Το αποτέλεσμα της προσαρμοστικής διεργασίας είναι ο έλεγχος της επεξεργασίας σήματος του εικονικού και του πραγματικού περιβάλλοντος.



Σχήμα 1.9: Γενική αρχιτεκτονική προσαρμοστικών ARA περιβαλλόντων

Η προσαρμοστικότητα στην τεχνολογία επαυξημένης ηχητικής πραγματικότητας σύμφωνα και με τα παραδείγματα που αναφέρθηκαν σε αυτή την ενότητα, φαίνεται να είναι περιορισμένη. Οι προσπάθειες εισαγωγής προσαρμοστικότητας διαιρούνται σε εκείνες τις οποίες υπάρχει προσαρμογή του περιεχομένου του σεναρίου και σε τεχνολογικό επίπεδο στην προσπάθεια βελτίωσης της αναπαραγωγής του πραγματικού περιβάλλοντος. Στην προτεινόμενη αρχιτεκτονική εφαρμογής της προσαρμοστικότητας μέσα στο γενικό μοντέλο ARA, φαίνεται η δυνατότητα δημιουργίας αντιληπτικών συστημάτων ήχου ARA τα οποία σκοπό θα έχουν την πρόβλεψη της αντίληψης του ανθρώπου μέσα από την δημιουργία κατάλληλων σημάτων διέγερσης για τα οποία λαμβάνεται υπόψιν η μεταβλητότητα του πραγματικού περιβάλλοντος.

με σκοπό τη μέγιστη εμπύθιση του θεατή μέσα από τη δημιουργία διάθεσης και ατμόσφαιρας στο ARA περιβάλλον.

## 1.4 Επιστημονική Πρωτότυπη συμβολή της εργασίας

Στο πλαίσιο της παρούσας εργασίας, αναπτύχθηκε το προσαρμοστικό μοντέλο ARA μείζης [96]. Η αναβάθμιση της λειτουργίας της παραδοσιακής ARA μείζης [61] με την εισαγωγή της προσαρμοστικότητας, δεν έχει αναφερθεί μέχρι σήμερα βιβλιογραφικά. Η μεθοδολογία που οδήγησε στην ανάπτυξη του παραπάνω μοντέλου, στηρίχθηκε στη μελέτη των περιπτώσεων επαυξημένων περιβαλλόντων ήχου υπό το πρίσμα της διάδρασης και ηχητικής σχεδίασης σε σύνθετα σενάρια τύπου ARA εστιάζοντας σε πιλοτικές εφαρμογές εκπαίδευσης και πολιτιστικής κληρονομιάς [110]. Η παραπάνω προσέγγιση μείζης, έχει βασικά προτερήματα ως προς την ευκολία ενσωμάτωσης σε πλατφόρμες ψηφιακής ανάπτυξης εφαρμογών τύπου ARA καθώς και σε συσκευές που ενσωματώνουν τεχνολογίες τύπου ARA.

Η χρήση του παραδοσιακού μοντέλου ARA μείζης μέσα από την υλοποίηση τυπικού σεναρίου ηχητικών παιχνιδιών τύπου ARA [111], κατέδειξε, ότι η δυναμική μεταβολή του πραγματικού περιβάλλοντος, επηρεάζει τη διάδραση ακροατή - συστήματος μέσα στο ARA περιβάλλον, δημιουργώντας ελάττωση της ακουστικής αντίληψης του [93] καθώς και της ευχρηστίας του ARA συστήματος. Το ζήτημα ευχρηστίας έγκειται σε ζητήματα ακουστικής αντίληψης του επαυξημένου περιβάλλοντος, καθώς και σε κατανόηση της χωρικής τοποθέτησης του εικονικού περιβάλλοντος που δημιουργούνται λόγω του ακουστικού φαινομένου της ηχητικής επικάλυψης μεταξύ των περιβαλλόντων, καθώς δε λαμβάνεται υπόψιν από το παραδοσιακό μοντέλο μείζης.

Για την αντιμετώπιση του παραπάνω προβλήματος, αναπτύχθηκε μια πρωτότυπη τεχνική δυναμικά μεταβαλλόμενης μείζης [93] η οποία επιτυγχάνει την προσαρμογή της ARA μείζης στο εκάστοτε δυναμικά μεταβαλλόμενο επαυξημένο περιβάλλον ήχου λαμβάνοντας υπόψιν το θέμα της αισθητικής και της ευχρηστίας. Κύριο χαρακτηριστικό του προτεινόμενου μοντέλου, αποτελεί η πρωτότυπη τεχνική εξόρυξης αντιληπτικών χαρακτηριστικών ήχου και χρήση τους στη διεργασία προσαρμοστικής μείζης. Η αισθητική βελτίωση της μείζης τεκμηριώνεται ποσοτικά μέσα από μια αλληλουχία υποκειμενικών ακουστικών τεστ. Επιπροσθέτως, το πρωτότυπο μοντέλο ARA μείζης, εφαρμόστηκε κατά την υλοποίηση ενός ηχητικού παιχνιδιού τύπου ARA και αξιολογήθηκε υποκειμενικά σε σύγκριση με το παραδοσιακό μοντέλο ARA μείζης, καταφέροντας με τα αποτελέσματά του, να επιβεβαιώσει τη θεωρητική επίδραση της ARA μείζης στη διάδραση του παίκτη με το σύστημα και τη βελτίωση της ακουστικής αντίληψής του, σε σύγκριση με το παραδοσιακό μοντέλο APA μείζης.

## 1.5 Δομή Εργασίας

Η εργασία δομείται σε οκτώ μέρη. Αρχικά, γίνεται μια εισαγωγή στο θέμα της εργασίας μέσα από μια γενική επισκόπηση των ARA περιβαλλόντων. Στο δεύτερο Κεφάλαιο παρουσιάζεται η ARA τεχνολογία υπό το πρίσμα της θεωρητικής προσέγγισης μέσα από την

τεχνολογική επισκόπησή, την παρουσίαση της σχετικής στάθμης τεχνικής και συναφών παραδειγμάτων. Το τρίτο Κεφάλαιο εστιάζει στη μελέτη της ARA μείξης μέσα από την αρχιτεκτονική του υφιστάμενου παραδοσιακού μοντέλου και τους περιορισμούς του. Στο τέταρτο Κεφάλαιο, παρουσιάζεται η μεθοδολογία ενσωμάτωσης της προσαρμοστικότητας στην ARA μείξη μέσα από την παρουσίαση του γενικευμένου μοντέλου και αναλύοντας τα μέρη του. Το Κεφάλαιο 5 εμφανίζει την πειραματική υλοποίηση της προσαρμοστικής ARA μείξης σε πραγματικό χρόνο μέσα από τα διάφορα τμήματά της. Στο κεφάλαιο 6, περιγράφεται η αξιολόγηση του μοντέλου προσαρμοστικής ARA μείξης. Στο Κεφάλαιο 7 καταγράφονται και αναλύονται τα αποτελέσματα της αξιολόγησης. Τέλος στο Κεφάλαιο 8 αναπτύσσονται τα συμπεράσματα και ακολούθως παρατίθενται μελλοντικές ερευνητικές προεκτάσεις.

## Κεφάλαιο 2

# Θεωρία και εφαρμογές ARA τεχνολογίας

### 2.1 Εισαγωγή

Σε αυτό το κεφάλαιο πραγματοποιείται η θεωρητική επισκόπηση της ARA τεχνολογίας. Οι πρόσφατες τεχνολογικές εξελίξεις, ενδεικτικά παραδείγματα εφαρμογών και κατ' επέκταση η στάθμη τεχνικής της, οριοθετούν το σύγχρονο πεδίο της «Επαυξημένης Πραγματικότητας Ήχου». Κοινός στόχος για την εξέλιξη της τεχνολογίας τύπου ARA αποτελεί η συνεχής αναβάθμιση της ποιότητας ως προς την αυθεντικότητα της αναπαράστασης του ψευδοακουστικού περιβάλλοντος, την αναπαραγωγή του εικονικού περιβάλλοντος μέσα από την τοποθέτηση των ηχητικών πηγών στον τρισδιάστατο εικονικό ηχητικό χώρο και τις μεθόδους αλληλεπίδρασης του ακροατή με τον επαυξημένο κόσμο.

Για την επίτευξη των παραπάνω, χρησιμοποιείται ένα σύμπλεγμα τεχνολογιών το οποίο εμπλέκει την αναπαραγωγή του πραγματικού περιβάλλοντος μέσω εξοπλισμού ειδικών προδιαγραφών, μεθόδους χωρικής τοποθέτησης ηχητικών αντικειμένων στον τρισδιάστατο εικονικό χώρο, μηχανικές διάδρασης του ακροατή με τον επαυξημένο κόσμο και τον ηχητικό σχεδιασμό. Τα παραπάνω μέρη συνδέονται μέσα από το εκάστοτε σενάριο επαυξημένης ηχητικής τεχνολογίας. Παραδείγματα εφαρμογών παρουσιάζονται στην Ενότητα 2.3.

### 2.2 Επισκόπηση τεχνολογίας σε ARA περιβάλλοντα

Τα τελευταία χρόνια, η ανάπτυξη της τεχνολογίας είναι ραγδαία. Το μέγεθος των υπολογιστικών συσκευών έχει ελαττωθεί σημαντικά. Η υλοποίηση έξυπνων κινητών συσκευών οι οποίες ενσωματώνουν ένα πλήθος συστημάτων οπτικοακουστικής αναπαραγωγής, καθώς και διαφορετικών τύπων αισθητήρων, βρίσκονται πλέον σε ευρεία κυκλοφορία. Επιπλέον, ειδικά παρελκόμενα, όπως ακουστικά με ενσωματωμένους αισθητήρες ή γυαλιά επαυξημένης πραγματικότητας, είναι πλέον προσιτά για το ευρύ κοινό. Ακολουθώντας τις τεχνολογικές εξελίξεις, φαίνεται ότι και ο τομέας της επαυξημένης ηχητικής πραγματικότητας έχει επηρεαστεί από αυτή την τεχνολογική ανάπτυξη σε σύγκριση με το παρελθόν, μέσα από την παραγωγή

ειδικού εξοπλισμού κατάλληλου για χρήση και στην καθημερινή ζωή. Παρ' όλα τα ζητήματα ευχρηστίας του εξοπλισμού που έχουν τεθεί στο παρελθόν για την υποστήριξη σεναρίων ARA [127], πλέον φαίνεται πως όλο και περισσότερες εταιρείες έχουν επενδύσει σε υλοποίηση τέτοιου εξοπλισμού [57]. Παρακάτω, θα παρουσιαστεί η τεχνολογική επισκόπηση της τεχνολογίας ARA διαιρεμένη σε τρία διακριτά σκέλη:

Πρώτο σκέλος: Τεχνολογίες που χρησιμοποιούνται στη μετάδοση του πραγματικού ηχητικού περιβάλλοντος στον ακροατή. Σε αυτό το σκέλος γίνεται μια ανάλυση των υφιστάμενων μεθόδων και τεχνολογιών οι οποίες στοχεύουν στη σύνδεση του πραγματικού ηχητικού περιβάλλοντος με τον ακροατή.

Δεύτερο σκέλος: Τεχνικές χωρικής τοποθέτησης των ηχητικών αντικειμένων του εικονικού περιβάλλοντος. Σε αυτό το σκέλος, δίνεται μια επισκόπηση των βέλτιστων τεχνικών χωρικής τοποθέτησης ηχητικών πηγών στον τρισδιάστατο εικονικό ηχητικό κόσμο.

Τρίτο σκέλος: Το τελευταίο σκέλος εστιάζει στις σύγχρονες τεχνολογίες διάδρασης που χρησιμοποιούνται σε επαυξημένα ηχητικά περιβάλλοντα. Μέσα από αυτές τις τεχνολογίες πραγματοποιείται η αλληλεπίδραση του ακροατή με το περιβάλλον επαυξημένης πραγματικότητας ολοκληρώνοντας την επικοινωνία του ακροατή με το σύστημα.

### 2.2.1 Τεχνολογίες αναπαραγωγής πραγματικού περιβάλλοντος στα περιβάλλοντα τύπου ARA

Στα επαυξημένα περιβάλλοντα τύπου ARA, ο ακροατής εισάγεται στον εικονικό κόσμο χωρίς να απομονώνεται από τον αντίστοιχο πραγματικό. Για να επιτευχθεί αυτό, το πραγματικό περιβάλλον θα πρέπει να μεταφέρεται στον ακροατή με φυσικό τρόπο ή να ελέγχεται ως προς τη φυσικότητα της αναπαραγωγής του. Για το λόγο αυτό, γίνεται χρήση ακουστικών τα οποία μπορούν να υποστηρίξουν ειδικές τεχνικές, οι οποίες ορίζονται ως Hear-Through (HT), δίνοντας έμφαση στην ηχητική διαφάνεια που δύναται να παρέχει ο ακουστικός εξοπλισμός.

Ως HT ακουστικά στην ARA τεχνολογία ορίζονται τα ακουστικά τα οποία επιτρέπουν παράλληλα την αναπαραγωγή του πραγματικού περιβάλλοντος χωρίς να απομονώνουν τον ακροατή από αυτό. Συνηθισμένη πρακτική είναι η χρήση αμφιωτικής τεχνολογίας. Η συγκεκριμένη τεχνολογία λαμβάνει ως βάση της, την λειτουργία της αμφιωτικής ακοής η οποία περιγράφει την ακοή ως ένα σύμπλεγμα από δύο σήματα τα οποία μεταβαίνουν από το εξωτερικό αυτί στον τυμπανικό υμένα του κάθε αυτιού δημιουργώντας την αμφιωτική αίσθηση της ακοής με χρήση δύο ανεξάρτητων ακουστικών καναλιών.

Στην περίπτωση που αυτά τα σήματα καταγράφουν από την θέση του αυτιού, μεθοδολογία που ορίζεται ως αμφιωτική εγγραφή, και αναπαράχθουν αυτοτελή, διατηρούν τα χαρακτηριστικά του ήχου κατά την ώρα της ηχογράφησης, συμπεριλαμβάνοντας χωρικά χαρακτηριστικά του ήχου καθώς και την χροιά του [86]. Ως αμφιωτική θεωρείται η αναπαραγωγή των δύο εγγεγραμμένων σημάτων με τη χρήση εξοπλισμού που τοποθετείται στην είσοδο του ακουστικού πόρου. Στην πράξη, τοποθετούνται αμφιωτικά μικρόφωνα στην είσοδο του ακουστικού πόρου με συννηθέστερη θέση την εξωτερική πλευρά των ακουστικών για την καταγραφή του πραγματικού περιβάλλοντος. Η λειτουργική διαφάνεια του ακουστικού εξοπλισμού είναι σημαντική

για την επίτευξη μια πιστής αναπαραγωγής του ομοιώματος του πραγματικού ακουστικού περιβάλλοντος, το οποίο ορίζεται ως ψευδοακουστικό.

Ο τύπος των ακουστικών διαδραματίζει κυρίαρχο ρόλο στον τρόπο που εφαρμόζεται η τεχνική HT. Αναλυτικότερα, ο τρόπος που εφαρμόζονται τα εκάστοτε ακουστικά, η ευχρηστία τους [127], το σημείο του κεφαλιού στο οποίο τοποθετούνται και η απαραίτητη ισοστάθμιση η οποία μπορεί να διορθώσει θέματα ποιότητας στα ακουστικά [129], αποτελούν κοινές συνιστάμενες για την HT τεχνική και τα ακουστικά. Παρακάτω γίνεται μια επισκόπηση ακουστικών στα οποία εφαρμόζεται τεχνική HT, στοχεύοντας στη φυσική αναπαραγωγή του πραγματικού ηχητικού περιβάλλοντος σε επαυξημένα περιβάλλοντα τύπου ARA.

Ακουστικά εσωτερικής τοποθέτησης στην είσοδο του ακουστικού πόρου τύπου In-Ear: η αναπαραγωγή του πραγματικού περιβάλλοντος με την χρήση ακουστικών τύπου In-Ear, υποστηρίζει τη χρήση αμφιωτικής τεχνολογίας εγγραφής και αναπαραγωγής. Πιο αναλυτικά, τεχνικές ενσωμάτωσης μικροφώνου στην εξωτερική πλευρά κάθε ακουστικού, επιτρέπουν την εγγραφή του ηχητικού περιβάλλοντος και ακολούθως τη μεταφορά στον τυμπανικό υμένα μέσω των ακουστικών που εφαρμόζονται στην είσοδο του ακουστικού πόρου. Η τοποθέτησή τους στην είσοδο του ακουστικού πόρου, εγγυάται την καταγραφή από τα μικρόφωνα, των ανακλάσεων του ήχου που προσπίπτει στο κάθε αυτί και προέρχονται από το ανθρώπινο σώμα και το εξωτερικό αυτί. Αυτές οι ανακλάσεις, είναι υπεύθυνες για τη διατήρηση της κατευθυντικότητας, της απόστασης και των φασματικών χαρακτηριστικών που προσδίδουν για παράδειγμα τη χροιά του πραγματικού ήχου. Η μοντελοποίηση των παραπάνω ανακλάσεων γίνεται με την υλοποίηση των HRTFs - Head Related Transfer Functions (Συναρτήσεις Μεταφοράς Ανατομικών Χαρακτηριστικών Κεφαλής - Ωτός) [10] οι οποίες χρησιμοποιούνται στη χωρική τοποθέτηση του εικονικού περιβάλλοντος (βλπ Εν. 2.2.2). Η ακουστική διαφάνεια των ακουστικών κατά την αμφιωτική αναπαραγωγή του πραγματικού περιβάλλοντος είναι το κυρίως ζητούμενο, δηλαδή η αναπαραγωγή του πραγματικού περιβάλλοντος στον ακροατή χωρίς ακουστικές μεταβολές, δηλαδή σαν να μη φοράει τα ακουστικά [102]. Για να επιτευχθεί αυτό ωστόσο θα πρέπει να ξεπερασθούν τα παρακάτω πρακτικά ζητήματα λόγω της εφαρμογής του ακουστικού στο αυτί.

Αρχικά το πρώτο ζήτημα αποτελεί η μεταβαλλόμενη ακουστική αυτιού λόγω έμφραξης του ακουστικού πόρου από το ακουστικό τύπου In-Ear (Occlusion Effect). Ο ακουστικός πόρος λειτουργεί ως κυλινδρικό αντηχείο ημι-ανοικτό (από την μια πλευρά ανοικτό και κλειστό από την πλευρά που βρίσκεται ο τυμπανικός υμένας). Τα χαρακτηριστικά του είναι η λειτουργία ως αντηχείο σε μήκος κύματος  $\lambda/4$  με την πρώτη αντήχηση να συμβαίνει μεταξύ των 2 kHz - 4 kHz.

Η τοποθέτηση του ακουστικού στον ακουστικό πόρο μετατρέπει το ημι-ανοικτό αντηχείο σε κλειστό. Κατά συνέπεια αλλάζει η ακουστική η οποία διέπεται από τις ιδιότητες ενός κλειστού ηχητικού σωλήνα ο οποίος μετασχηματίζεται σε αντηχείο με πρώτο σημείο αντήχησης στο  $\lambda/2$  και μεταξύ των συχνοτήτων 5 kHz-10 kHz το οποίο εξαρτάται από το μήκος του ακουστικού πόρου και την εφαρμογή του εκάστοτε ακουστικού [45]. Κατά συνέπεια, για να υπάρξει η φυσική αναπαραγωγή του πραγματικού περιβάλλοντος χρειάζεται να εφαρμοστεί τεχνική ακύρωσης της αντήχησης λόγω της εφαρμογής του ακουστικού καθώς και να επανα-

δημιουργηθεί η ακουστική του ανοικτού ακουστικού πόρου. Προηγούμενες εργασίες έχουν παρουσιάσει μεθοδολογία επίλυσης του παραπάνω προβλήματος με τη δημιουργία ακολουθίας φίλτρων [128].

Το επόμενο ζήτημα προκύπτει από τη μη ακριβή εφαρμογή του ακουστικού στον ακουστικό πόρο και την ύπαρξη ηχητικής διαρροής (Leakage Effect). Αυτό προκαλεί τη μερική είσοδο του πραγματικού ηχητικού περιβάλλοντος στον ακουστικό πόρο υπό τη μορφή διαρροής. Ο υπολογισμός της διαρροής, βοηθάει στη δημιουργία της ηχητικής διαφάνειας του ζεύγους ακουστικών. Ωστόσο η ακρίβεια του υπολογισμού είναι περιπτωσιολογική, καθώς η φυσιολογία του ακουστικού οργάνου μεταξύ διαφορετικών ατόμων δεν είναι ίδια. Στο παρελθόν έχει δημιουργηθεί μοντέλο πρόβλεψης της διαρροής του ακουστικού, το οποίο έχει υποδείξει πως, σε γενικό επίπεδο, αυτή κυμαίνεται σε ένα εύρος μεταξύ των 6 kHz - 7 kHz [126].

Επίσης, θα πρέπει να προστεθεί στα παραπάνω, η επίδραση του μεγέθους του θαλάμου του ακουστικού πόρου. Το μικρό μέγεθος του θαλάμου σε συνάρτηση με την επίδραση της υψηλής ακουστικής πίεσης που δημιουργείται από την έμφραξη του θαλάμου, δημιουργεί συσσώρευση ενέργειας με δυσκολία στην εκτόνωσή της. Το φαινόμενο αυτό ορίζεται ως «Αρχή Πίεσης Ακουστικού Θαλάμου» (Pressure Chamber Principle) και επηρεάζει τις χαμηλές και μεσαίες συχνότητες [15]. Η μη τέλεια εφαρμογή του ακουστικού στον ακουστικό πόρο, βοηθάει στην εξασθένιση της επίδρασης της υψηλής συγκέντρωσης ακουστικής πίεσης στον ακουστικό θάλαμο [103].

Τέλος, το ζήτημα της απομόνωσης του ακροατή από το πραγματικό περιβάλλον με τη χρήση ακουστικών, αποτελεί επιπλέον παράγοντα στην αμφιωτική αναπαραγωγή του πραγματικού περιβάλλοντος. Το εξασθενημένο ακουστικό σήμα που προέρχεται από το πραγματικό περιβάλλον και διαπερνάει τα ακουστικά, αποτελεί τον τελευταίο παράγοντα στην αμφιωτική αναπαραγωγή του πραγματικού περιβάλλοντος. Προηγούμενες εργασίες παρουσιάζουν καμπύλες απομόνωσης του ακροατή σε συνάρτηση της συχνότητας σε διαφορετικούς τύπους ακουστικών ανάλογα με το πόσο σφιχτά τοποθετούνται στον ακουστικό πόρο [103]. Σε ακουστικά τύπου In-Ear και στην περίπτωση χαλαρής τοποθέτησης, υπάρχει αύξηση των χαμηλών συχνοτήτων, επιδεινώνοντας την παθητική απομόνωση [103].

Για την επίλυση των παραπάνω ζητημάτων εφαρμόζονται τεχνικές ισοστάθμισης του ηχητικού σήματος εισόδου. Μειονέκτημα των τεχνικών ισοστάθμισης, αποτελεί η καθυστέρηση που μπορεί να εισάγει η εκάστοτε τεχνική στην αναπαραγωγή του σήματος, καθώς ο ήχος του πραγματικού περιβάλλοντος που εισάγεται λόγω διαρροών δεν συνάδει χρονικά με τον αναπαραγόμενο ήχο του ψευδοακουστικού περιβάλλοντος (Comb Filter Effect) [105]. Η χρήση αναλογικής ισοστάθμισης φαίνεται να βοηθάει στην αποτροπή φαινομένων εισαγωγής καθυστέρησης στο σήμα του ψευδοακουστικού περιβάλλοντος [61]. Η χρήση ψηφιακών φίλτρων εισάγει καθυστέρηση στο αναπαραγόμενο σήμα του ψευδοακουστικού περιβάλλοντος, ωστόσο υπάρχει ένα οξύμωρο χαρακτηριστικό. Η καθυστέρηση του σήματος λόγω της έμφραξης του ακουστικού πόρου από το ακουστικό, φαίνεται να αντισταθμίζεται με την καθυστέρηση που μπορεί να εισάγει η τεχνική ισοστάθμισης του σήματος [104]. Παράλληλα, υπάρχουν πλέον υλοποιήσεις με ψηφιακή ισοστάθμιση σε πραγματικό χρόνο η οποία εισάγει καθυστέρηση στο αναπαραγόμενο σήμα λιγότερο από 1.4 ms [102].

Ακουστικά κλειστού τύπου Closed-Back: Τα ακουστικά τύπου Closed-Back, εφαρμόζουν στο περύγιο του εξωτερικού αυτιού. Παρέχουν σχετικά καλή ηχητική απομόνωση στον ακροατή από το πραγματικό ηχητικό περιβάλλον. Σε σύγκριση με τα ακουστικά τύπου In-Ear, δεν έχουν ζητήματα εφαρμογής καθώς δεν εφαρμόζουν στον ακουστικό πόρο και καλύπτουν το περύγιο.

Η έμφραξη του ακουστικού πόρου φαίνεται να επηρεάζεται από την «ακουστική αντίσταση» των ακουστικών. Ως ακουστική αντίσταση εννοείται η δυσκολία που εισάγεται από τα ακουστικά κατά τη μετάδοση του ήχου από το πραγματικό περιβάλλον στο εσωτερικό αυτί παρεμβάλλοντας τον εξοπλισμό των ακουστικών. Τα ακουστικά τύπου Closed-Back έχουν μεγαλύτερη ακουστική αντίσταση σε σχέση με τα ακουστικά τύπου Open-Back. Η αύξηση της αντίστασης των ακουστικών συνάδει με την αύξηση της στάθμης ηχητικής πίεσης στις χαμηλές συχνότητες, όταν υπάρχει για παράδειγμα ομιλία του ακροατή. Η γνώση της ακουστικής αντίστασης του ακουστικού, φαίνεται να μπορεί να οδηγήσει στην πρόβλεψη της επίδρασης της έμφραξης του ακουστικού πόρου και δυνατότητα ισοστάθμισης [116]. Τα τελευταία χρόνια έχουν παρουσιαστεί πειραματικές μεθοδολογίες χρήσης ακουστικών τύπου Closed-Back, με χρήση μηχανισμού ακύρωσης θορύβου, για καλύτερη απομόνωση του ακροατή από το εξωτερικό περιβάλλον και τον υπολογισμό της ιδανικής ισοστάθμισης με την χρήση εσωτερικού μικροφώνου στο κάθε ακουστικό [107, 44]. Η παραπάνω μεθοδολογία αυξάνει την πολυπλοκότητα του εξοπλισμού, μειώνοντας της ευχρηστία ενός τέτοιου συστήματος.

Επιπροσθέτως, πρέπει να σημειωθεί η δημιουργία προσαρμοστικών αλγόριθμων, οι οποίοι υπολογίζουν την απομόνωση που παρέχουν τα ακουστικά και ρυθμίζουν την ισοστάθμιση της τεχνικής HT, λαμβάνοντας υπόψη την τυχόν καθυστέρηση και βοηθούν στην βελτίωση της ακουστικής εμπειρίας [73] ως προς την αναπαραγωγή του ψευδοακουστικού περιβάλλοντος.

Ακουστικά τύπου Open-Back: Δεν παρέχουν τόσο καλή ηχητική απομόνωση όσο τα προηγούμενα ακουστικά. Παρ' όλα αυτά, λόγω του ανοικτού τύπου του ακουστικού, επιτρέπουν να περάσει το πραγματικό ηχητικό περιβάλλον με φυσικότερο τρόπο απ' ότι τα προηγούμενα ακουστικά. Όμοια με τα ακουστικά τύπου Closed-Back, η χρήση εσωτερικών μικροφώνων απαιτείται για την ιδανική ισοστάθμιση του εικονικού περιβάλλοντος σε σχέση με το πραγματικό [106].

Ακουστικά τύπου Bone-Conduction: Τα ακουστικά τύπου Bone-Conduction συναντώνται σε εφαρμογές υποβοηθούμενης ακρόασης (Assisted Listening) μέσα σε περιβάλλοντα επαυξημένης πραγματικότητας. Η τοποθέτησή τους γίνεται εκτός του ακουστικού πόρου και του περυγίου, ακουμπώντας σε οστά του κεφαλιού, όπως φαίνεται στο Σχήμα 2.1δ'. Η χρήση τέτοιου τύπου ακουστικών παρέχει υψηλή πιστότητα του πραγματικού ηχητικού περιβάλλοντος καθώς ο ακουστικός πόρος είναι ανοικτός και δεν απαιτείται ο μετασχηματισμός σε ψευδοακουστικό περιβάλλον. Ωστόσο, πρακτικά υπάρχει ο μη έλεγχος του πραγματικού περιβάλλοντος όπως σε περιπτώσεις φασαρίας, δημιουργώντας δυσκολία στην κατανόηση του εικονικού περιβάλλοντος. Επίσης σε συγκριτικά τεστ με τυπικά ακουστικά, διαπιστώνεται η λειτουργία τους σε ένα περιορισμένο εύρος συχνοτήτων από 0.25 kHz - 4 kHz σε ένα δυναμικό εύρος από 30 dB - 80 dB [120]. Ο περιορισμός λειτουργίας αυτού του είδους των ακουστικών λόγω της λειτουργίας μέσω των δονήσεων σε οστά έναντι του αέρα [139] δημιουργούν περιορισμούς



(α') In-Ear



(β') Closed-Back



(γ') Open-back



(δ') Bone-Conduction

Σχήμα 2.1: Ακουστικά υποστηριζόμενα από τεχνολογία Hear-Through α)Τύπου In-Ear, β)Τύπου Closed-Back, γ)Τύπου Open-Back, δ)Τύπου Bone-Conduction

στην τυχόν προσαρμογή του εικονικού περιβάλλοντος στο πραγματικό περιβάλλον αν λάβουμε ως παράδειγμα την συχνотική απόκριση σε ακουστικά τύπου In-Ear με προσαρμοστική ισοστάθμιση.

Πρέπει να τονιστεί η σημασία της καταγραφής της επίδρασης του αυτιού στα σήματα που εισέρχονται στον ακουστικό πόρο και η διατήρηση ή επανασύσταση της ακουστικής του έξω αυτιού. Στα ακουστικά κλειστού τύπου έχουν πραγματοποιηθεί προσπάθειες για προσαρμοστική παραμετρική ισοστάθμιση σε τεχνική HT σε επαυξημένα ηχητικά περιβάλλοντα με χρήση προϋπολογισμένων παραμετρικών HT φίλτρων σε πραγματικό χρόνο με σκοπό την εισαγωγή των φασματικών χαρακτηριστικών που προκαλούνται από το έξω αυτί [106, 43]. Στα ακουστικά τύπου In-Ear που γίνεται χρήση αμφιωτικής τεχνολογίας εγγραφής για τη δημιουργία

του ψευδοακουστικού περιβάλλοντος, δεν υπάρχει η παραπάνω ανάγκη, καθώς η ακουστική του έξω αυτιού διατηρείται αθέρα.

Η παραπάνω ανάλυση αποτελεί μια επισκόπηση των παραγόντων που επιδρούν στην αμφιωτική αναπαραγωγή του πραγματικού περιβάλλοντος και της τεχνολογίας που έχει αναπτυχθεί για περιβάλλοντα τύπου ARA. Η αμφιωτική αναπαραγωγή περιλαμβάνει το κομμάτι του πραγματικού ηχητικού περιβάλλοντος, το οποίο έχει την ίδια βαρύτητα, όσο και η χωρική τοποθέτηση του εικονικού περιβάλλοντος, η οποία παρουσιάζεται στη συνέχεια.

### 2.2.2 Χωρική τοποθέτηση εικονικού περιβάλλοντος

Το επαυξημένο περιβάλλον περιλαμβάνει επιπλέον από το πραγματικό περιβάλλον, και το εικονικό. Στην Ενότητα 2.2.1 παρουσιάστηκαν οι εναλλακτικές αναπαραγωγής του πραγματικού περιβάλλοντος μέσα από τεχνολογίες που χρησιμοποιούνται σε επαυξημένα περιβάλλοντα ήχου. Σε αυτή την ενότητα της διατριβής, παρουσιάζεται η εικονική συνιστώσα του ARA περιβάλλοντος μέσα από μια επισκόπηση των τρόπων χωρικής παρουσίας της στο τρισδιάστατο εικονικό ακουστικό χώρο.

Στα σενάρια επαυξημένης πραγματικότητας, η εικονική πληροφορία διαδραματίζει σημαντικό ρόλο καθώς μαζί με το πραγματικό ηχητικό περιβάλλον, αποτελούν το επαυξημένο ηχητικό περιβάλλον μεταφέροντας την απαραίτητη πληροφορία σύμφωνα με το εκάστοτε σενάριο. Η αναπαραγωγή και σε αυτή την περίπτωση γίνεται με τη χρήση ειδικού ακουστικού εξοπλισμού ο οποίος πρέπει να λαμβάνει υπόψη την αναπαραγωγή του πραγματικού περιβάλλοντος. Ζητούμενο στην εικονική ηχητική συνιστώσα, είναι η χωρική τοποθέτησή της στον τρισδιάστατο ηχητικό κόσμο. Για αυτή την λειτουργία, είναι συνήθης η χρήση ακουστικών, καθώς συγκεντρώνουν τα εγγενή προτερήματα της αμφιωτικής τεχνολογίας [86]. Αναλυτικότερα, η ελεύθερη μετακίνηση του ακροατή στο χώρο γίνεται εφικτή με την απαλοιφή παρεμβολών τύπου συνακρόασης Cross-talk [36] από το ανοικτό ηχητικό πεδίο, οι οποίες σε αντίθετη περίπτωση, μεταβιβάζουν ακουστική ενέργεια μεταξύ των κοντινών ακουστικών σημάτων. Η τεχνική μετάδοσης αμφιωτικών σημάτων με τη χρήση ηχείων έχει οριστεί ως Transaural [37]. Σε αυτή την τεχνική, γίνεται χρήση ηχείων με δύο τρόπους, είτε με χρήση ηχείων σε μακρινή απόσταση από τον ακροατή όπου απαιτείται η χρήση τεχνικής ακύρωσης των παρεμβολών συνακρόασης, είτε με τη χρήση ηχείων σε πολύ κοντινή απόσταση από το αυτί, τοποθετημένα στον ώμο του ακροατή, όπου δεν απαιτείται σύστημα απαλοιφής των παρεμβολών συνακρόασης [10]. Στην περίπτωση των ακουστικών, η ακουστική ενέργεια του κάθε σήματος οδηγείται απευθείας στον ακουστικό πόρο χωρίς να υπάρχει παρεμβολή από το ανοικτό ηχητικό πεδίο στο γειτονικό ακουστικό κανάλι. Η απαλοιφή τέτοιου είδους παρεμβολής, δημιουργεί προοπτικές ελεύθερης μετακίνησης στον χώρο χωρίς τον υποχρεωτικό καθορισμό ενός βέλτιστου σημείου (Sweet Spot) που απαιτούν πολυκαναλικά συστήματα όπως τα τύπου Ambisonic [77] ή τα τύπου Wavefield Synthesis - WFS [25], και η χρήση τεχνικών χωρικής τοποθέτησης αμφιωτικής τεχνολογίας, παρέχουν την βέλτιστη ηχητική τοποθέτηση των ηχητικών πηγών στον εικονικό χώρο. Επιπροσθέτως, πρέπει να αναφερθεί η φορητότητα και η ευχρηστία του εξοπλισμού των ακουστικών σε σύγκριση με τα παραπάνω συστήματα, καθώς είναι μικρά σε

μέγεθος και εύκολα στην εφαρμογή τους. Ως μειονέκτημα σε σύγκριση με τα παραπάνω συστήματα, είναι η παρεμβατικότητα (Intrusiveness) τους στο υποκείμενο, καθώς χρειάζεται να είναι φορετά σε αντίθεση με πολυκαναλικές εκπομπές ή WFS συστήματα όπου ο ακροατής δεν χρειάζεται να φέρει τίποτα.

Η περιγραφή του τρόπου που αντιλαμβάνεται ο ακροατής τη θέση των ηχητικών πηγών έγκειται στην επιστήμη της ψυχοακουστικής. Σχετικά με το ακουστικό αίσθημα που παράγεται, τα ηχητικά σήματα στο ακουστικό κανάλι είναι τα πιο σημαντικά σήματα στο υποκείμενο για την χωρική ακουστική (Spatial Hearing) και οι μικρές αλλαγές σε αυτά τα σήματα μπορεί να οδηγήσουν σε σημαντικές μεταβολές σε αυτή [10], λαμβάνοντας υπόψη την αντίληψη και την φυσιολογία του ανθρώπινου αυτιού. Παρακάτω παρουσιάζονται οι δύο κύριες προσεγγίσεις της αμφιωτικής ακοής.

Η «Duplex Theory» [76] εμφανίστηκε στις αρχές του προηγούμενου αιώνα και υποστηρίζει πως στον προσδιορισμό της κατεύθυνσης μιας ηχητικής πηγής, σημαντικό ρόλο καταλαμβάνουν οι χρονικές διαφορές (Interaural Time Differences - ITDs) και οι διαφορές στην ηχητική στάθμη (Interaural Level Differences - ILDs) που προκύπτουν από τη στιγμή της μετάδοσης του ηχητικού κύματος μέχρι τη στιγμή που θα γίνει αντιληπτό από το κάθε αυτί. Οι ελάχιστες χρονικές διαφορές που είναι αντιληπτές είναι μεγέθους 10  $\mu\text{sec}$  [53] και οι ελάχιστες διαφορές σε στάθμη που γίνονται αντιληπτές είναι από 0.5 dB [84]. Η μοντελοποίηση της παραπάνω θεωρίας, έχει τον περιορισμό της λειτουργίας στο οριζόντιο επίπεδο χωρίς να μπορεί να γίνει διαχωρισμός σε μπροστινό και πίσω μέρος [10]. Αυτό το φαινόμενο έχει ονομαστεί ως Κώνος Αβεβαιότητας (Cone of Confusion) [91, 39] στο οποίο οι χρονικές διαφοροποιήσεις είναι ίδιες για περισσότερα ηχητικά σημεία μέσα στην περιοχή του Κώνου. Πρέπει να προστεθεί, πως οι χρονικές διαφοροποιήσεις βοηθάνε στην αντίληψη χαμηλών συχνοτήτων ( $f < 1 \text{ kHz}$ ) ενώ διαφοροποιήσεις στις στάθμες βοηθάνε στην χωρική αντίληψη υψηλών συχνοτήτων ( $f > 1.5 \text{ kHz}$ ) καθώς οι χαμηλές συχνότητες αποδυναμώνονται από το φαινόμενο σκίασης του ηχητικού κύματος από το κεφάλι [10].

Η δεύτερη προσέγγιση ξεπερνά το πρόβλημα του Κώνου Αβεβαιότητας και κάνει χρήση των ευρέως γνωστών συναρτήσεων τύπου HRTFs. Οι συναρτήσεις HRTFs μοντελοποιούν μαθηματικά εκτός από τις διαφοροποιήσεις στο χρόνο και στις ηχητικές στάθμες, τον μηχανισμό λειτουργίας του εξωτερικού αυτιού. Είναι γνωστό πως το ανθρώπινο σώμα λειτουργεί ως φίλτρο [10]. Ο ήχος πριν κατευθυνθεί στον ακουστικό πόρο, προσπίπτει στο ανθρώπινο σώμα το οποίο λειτουργεί ως γραμμικό φίλτρο. Η καταγραφή των ακουστικών μεταβολών μπορεί να γίνει με την τοποθέτηση μικροφώνων στην είσοδο του ακουστικού πόρου. Η καταγραφή των στιγμιαίων διεγέρσεων από ηχητικές πηγές σε προκαθορισμένες θέσεις στο τρισδιάστατο επίπεδο, συνθέτουν τις χροστικές διεγέρσεις τύπου HRIR - Head Related Impulse Response. Οι συχνοτικές αποκρίσεις των χροστικών διεγέρσεων αποτελούν οι συναρτήσεις HRTFs [35]. Η ενσωμάτωση των χροστικών διεγέρσεων HRIRs στο πεδίο του χρόνου ή των χροστικών αποκρίσεων HRTFs στο πεδίο της συχνότητας, σε ανηχοϊκά μονοφωνικά σήματα, έχουν σαν αποτέλεσμα την δημιουργία ηχητικού ερεθίσματος το οποίο περιέχει και την χωρική πληροφορία του ηχητικού αντικειμένου [10].

Η τεχνολογία ακουστικών που χρησιμοποιείται στο εκάστοτε ηχητικό σενάριο επαιξη-

μένης πραγματικότητας έχει άμεση σχέση με τη μέθοδο χωρικής τοποθέτησης των εικονικών ηχητικών πηγών. Η χρήση των HRTFs απαιτεί τη χρήση ακουστικών τύπου In-Ear, καθώς η μέτρησή τους έχει γίνει στην είσοδο του ακουστικού πόρου καταγράφοντας όλα τα φασματικά χαρακτηριστικά του σήματος λόγω της επίδρασης του έξω αυτιού. Συνεπώς, η χρήση ακουστικών με εφαρμογή στον ακουστικό πόρο βοηθά στην βέλτιστη αντίληψη των χωρικών χαρακτηριστικών της ηχητικής πηγής. Η χρήση ακουστικών για παράδειγμα τύπου Closed-back χρειάζεται επιπλέον φίλτρα αντιστάθμισης [106] της επίδρασης του εξωτερικού αυτιού για να υπάρχει η βέλτιστη αντίληψη της χωρικότητας των εικονικών πηγών ήχου. Τα ακουστικά τύπου Bone-Conduction σύμφωνα με έρευνες, δεν έχουν την ποιότητα στην χωρική αντίληψη που έχουν τα ακουστικά με τη χρήση HRTFs [71] λόγω της περιορισμένης συχνотικής απόκρισης, αλλά μπορούν να τις υποστηρίξουν σε έναν ικανοποιητικό βαθμό [75, 70, 5, 6].

Συμπερασματικά, το είδος της εφαρμογής μπορεί να κατευθύνει την επιλογή της τεχνολογίας ακουστικών, όπως για παράδειγμα σε εφαρμογές πλοήγησης επαυξημένης ηχητικής πραγματικότητας, η χρήση ακουστικών τύπου Bone-Conduction για χωρική τοποθέτηση ηχητικών πηγών φαίνεται να έχουν καλή απόδοση [138]. Παρόλα αυτά τα φίλτρα αντιστάθμισης εκτός από τις συναρτήσεις HRTFs έχουν σημαντικό ρόλο καθώς, όπως φαίνεται σε μια μελέτη σχετικά με κατανόηση λόγου σε περιβάλλοντα τύπου ARA, η χρήση ακουστικών χωρίς φίλτρα αντιστάθμισης δεν έχουν τόσο καλή απόκριση όσο τα ακουστικά τύπου Bone-Conduction, ενώ σε αντίθεση, με τη χρήση των φίλτρων, η απόκριση είναι καλύτερη από τα ακουστικά τύπου Bone-Conduction [65].

Η επιλογή της τεχνολογίας αναπαραγωγής της εικονικής ηχητικής συνιστώσας, είναι ένα πολύπλοκο ζήτημα. Η πολυπλοκότητα έγκειται στους ακόλουθους παράγοντες: το είδος της ηχητικής πληροφορίας που απαιτείται να αναπαραχθεί, την ποιότητα του χωρικού αποτελέσματος που στοχεύεται να δοθεί, την ευχρηστία του εξοπλισμού, το είδος του πραγματικού περιβάλλοντος και γενικά το σενάριο της εφαρμογής. Για παράδειγμα στην εφαρμογή Sonic-Badminton [63], όπου υπάρχει η ηχητική επαύξηση του αθλητικού παιχνιδιού αντιπέρισης, η αναπαραγωγή της εικονικής πληροφορίας που συνοδεύει το πέταγμα της ειδικής μπάλας, γίνεται με χρήση απλής στερεοφωνίας ελαχιστοποιώντας την πολυπλοκότητα του συστήματος χωρικής τοποθέτησης της εικονικής πηγής παρέχοντας την απαιτούμενη πληροφορία για τους παίκτες και σε περιπτώσεις ατόμων με δυσκολία στην όραση. Η παραπάνω μεθοδολογία δίνει τη δυνατότητα χρήσης ακουστικών ή ηχείων. Σε ένα διαφορετικό παράδειγμα τηλεπαρουσίας τύπου ARA [47], για τη μετάδοση της εικονικής πληροφορίας που αντικατοπτρίζει το περιβάλλον του συνομιλητή και τη μοντελοποίηση της ακουστικής του χώρου, γίνεται χρήση συναρτήσεων τύπου HRTF οι οποίες παρέχουν την πιστότερη αναπαραγωγή του ακουστικού χώρου. Η χρήση της παραπάνω μεθοδολογίας, περιορίζει το χρήστη στη χρήση ακουστικών, οπότε στην περίπτωση της τηλεπαρουσίας τύπου ARA, δεν είναι εφικτή η χρήση ηχείων λόγω του φαινομένου της συνακρόασης [10]. Εκτός από τις τεχνολογίες ηχητικής αναπαραστάσης πραγματικού και εικονικού περιβάλλοντος, σε ένα σενάριο τύπου ARA γίνεται χρήση τεχνολογιών διάδρασης για την επικοινωνία του ακροατή με το σύστημα. Στην επόμενη ενότητα παρουσιάζονται οι τεχνολογίες διάδρασης σε περιβάλλοντα τύπου ARA.

### 2.2.3 Τεχνολογίες διάδρασης σε ARA περιβάλλοντα

Η τεχνολογική εξέλιξη έχει διαμορφώσει ένα νέο τοπίο στον τομέα της επαυξημένης πραγματικότητας. Η παρουσίαση όλο και μικρότερων φορητών συσκευών, όπως είναι τα έξυπνα κινητά τηλέφωνα και ο εμπλουτισμός τους με συστήματα αισθητήρων είναι πλέον κάτι συνηθισμένο. Επιπλέον, είναι προσιτή η κατοχή τέτοιων συσκευών από μεγάλο ποσοστό πληθυσμού. Όλοι οι παραπάνω παράγοντες συνθέτουν συνθήκες στις οποίες υπάρχει η δυνατότητα για νέα προσέγγιση των σεναρίων επαυξημένης πραγματικότητας ήχου μέσα από νέους τρόπους διάδρασης.

Στην επαυξημένη πραγματικότητα ήχου, ο ακροατής εισάγεται στο επαυξημένο περιβάλλον με το οποίο αλληλεπιδρά ενεργητικά και παθητικά [145]. Ως ενεργητικά θεωρούμε την διάδραση με το επαυξημένο περιβάλλον η οποία είναι επιλογή του ακροατή, όπως είναι για παράδειγμα το πάτημα ενός κουμπιού με το οποίο εκτελείται μια ενέργεια μέσα στο περιβάλλον. Η παθητική αλληλεπίδραση του ακροατή με το επαυξημένο περιβάλλον περιέχεται για παράδειγμα στον καθορισμό ακουστικών παραμέτρων οι οποίες προέρχονται από μεταβολή της θέσης του ακροατή. Το αποτέλεσμα γίνεται αντιληπτό από τον ακροατή και στις δύο περιπτώσεις και συμβάλει στην αμφίδρομη επικοινωνία με το επαυξημένο περιβάλλον [99].

Για την υλοποίηση της αλληλεπίδρασης μεταξύ ακροατή και επαυξημένου περιβάλλοντος ήχου, χρησιμοποιούνται τεχνολογίες διάδρασης. Οι διεργασίες που συναντώνται σε περιβάλλοντα επαυξημένης πραγματικότητας ήχου που απαιτούν τέτοιες τεχνολογίες συνοψίζονται παρακάτω.

- Αναγνώριση θέσης ακροατή σε περιβάλλοντα τύπου ARA.

Η αναγνώριση της θέσης του ακροατή στο επαυξημένο περιβάλλον είναι απαραίτητη για την αλληλεπίδραση με ηχητικές παραμέτρους οι οποίες καθορίζουν την ακουστική συμπεριφορά των εικονικών ηχητικών αντικειμένων. Για να επιτευχθεί αυτό, θα πρέπει να γίνει ο καθορισμός της περιοχής στην οποία εκτυλίσσεται το εκάστοτε ηχητικό σενάριο επαυξημένης πραγματικότητας και η συσχέτισή του με τον πραγματικό κόσμο. Το σύστημα δισδιάστατης ή τρισδιάστατης ηχητικής αναπαραγωγής του εικονικού περιβάλλοντος, λαμβάνει τιμές από αισθητήρες που υπάρχουν στο πραγματικό περιβάλλον του ακροατή. Παραδείγματα τέτοιων αισθητήρων είναι τα συστήματα καταγραφής γεωγραφικών συντεταγμένων τύπου GPS [110], συστήματα τύπου Beacon [74] για κλειστούς χώρους, σύστημα υπέρυθρων σημάτων [97] ή συστήματα κάμερας [92].

- Αναγνώριση κατεύθυνσης κεφαλιού.

Η θέση του κεφαλιού και κατά συνέπεια η θέση των αυτιών. Η συγκεκριμένη παράμετρος επηρεάζει τη διάδραση του ακροατή με το σύστημα ηχητικής αναπαραγωγής του εικονικού ηχητικού περιβάλλοντος. Η κατεύθυνση του κεφαλιού μεταβάλλει τις ακουστικές συναρτήσεις τύπου ITD και ILD όπως σημειώθηκε και στην Ενότητα 2.2.2. Η εισαγωγή των κινήσεων του κεφαλιού επιτρέπει τη βέλτιστη αντίληψη του τρισδιάστατου ηχητικού επιπέδου. Δηλαδή σε αντίθετη περίπτωση, το κεφάλι θα πρέπει να μη μετακινείται και να ακολουθεί την κατεύθυνση του σώματος. Για την καταγραφή της κίνησης

του κεφαλιού, χρησιμοποιούνται αισθητήρες γυροσκοπίου, επιταχυντή ή πυξίδα. Σε φορητές συσκευές, όπως για παράδειγμα σε έξυπνα τηλέφωνα, υπάρχουν ενσωματωμένοι. Επίσης, σε εμπορικές συσκευές ακουστικών [16], ενσωματώνονται τεχνολογίες τύπου Head Tracking για εντοπισμό και καταγραφή της κίνησης του κεφαλιού. Το παραπάνω καταδεικνύει τη δυνατότητα εξέλιξης της επαυξημένης πραγματικότητας ήχου για χρήση στο ευρύ κοινό.

- Αναγνώριση κινήσεων ακροατή.

Η διάδραση του ακροατή με το σύστημα τύπου ARA καθορίζεται από το εκάστοτε ηχητικό σενάριο. Η εισαγωγή του ακροατή μέσα στο επαυξημένο περιβάλλον γίνεται μέσω της χωρικής μετατόπισής του, όπως περιγράφηκε και παραπάνω, σηματοδοτώντας έναν παθητικό τρόπο διάδρασης. Εκτός αυτού, σε ένα ηχητικό σενάριο επαυξημένης πραγματικότητας, χρειάζεται και η ενεργητική διάδραση με ηχητικά αντικείμενα του εικονικού περιβάλλοντος. Σε αυτό συχνά χρησιμοποιούνται ενεργοποιητές, όπως είναι κουμπιά συσκευών [92] με καθορισμένη λειτουργία καθώς και αισθητήρες καταγραφής κίνησης από συσκευές που χρησιμοποιούνται ήδη μέσα σε τέτοια σενάρια. Ένα χαρακτηριστικό παράδειγμα είναι οι τρόποι διάδρασης με την χρήση νευμάτων από τον χρήστη για την αντιμετώπιση εχθρών σε ηχητικά παιχνίδια τύπου ARA [111] με τη χρήση φορητών συσκευών τύπου Tablet.

Από τα παραπάνω, καθίσταται σαφές ότι, οι νέες έξυπνες συσκευές παρέχουν ένα πλούσιο όγκο πληροφορίας που προέρχεται από τους ενσωματωμένους αισθητήρες που περιέχουν, δημιουργώντας όλες τις προοπτικές για την υλοποίηση μιας βέλτιστης εμπειρικής εμβυθιστικής μέσα στα ηχητικά περιβάλλοντα επαυξημένης πραγματικότητας. Επιπλέον, η ευχρηστία και το μικρό μέγεθος τους επιτρέπουν την φορητότητα τέτοιων συστημάτων και την χρήση από το κοινό σε πραγματικά περιβάλλοντα [46, 2].

## 2.3 Εφαρμογές ARA τεχνολογίας

Η ARA τεχνολογία χρησιμοποιείται από ένα ευρύ φάσμα εφαρμογών. Η εξέλιξη των φορητών συσκευών όπως έχει ήδη προαναφερθεί, επιτρέπει την υλοποίηση εφαρμογών με την υποστήριξη εξοπλισμού που μπορεί να φορεθεί από τον χρήστη [47, 60, 50]. Στις περισσότερες περιπτώσεις πρόκειται για πειραματικές εφαρμογές, καθώς πρόκειται για ένα διεπιστημονικό πεδίο το οποίο ερευνά και συνδυάζει τομείς όπως είναι η τεχνολογία ήχου, ο ηχητικός σχεδιασμός, η τεχνολογία αισθητήρων με το πεδίο της ψυχοακουστικής. Η συνένωση όλων των παραπάνω για την υλοποίηση ηχητικών σεναρίων επαυξημένης πραγματικότητας αποτελεί μια ενεργή ερευνητική περιοχή με την παραγωγή ποικίλων εφαρμογών τύπου ARA. Παρακάτω παρουσιάζονται παραδείγματα τέτοιων εφαρμογών.

Η τηλεπαρουσία του ακροατή σε πραγματικό χρόνο αποτελεί μία από τις πρώτες προσεγγίσεις σεναρίου επαυξημένης ηχητικής πραγματικότητας όπου εικονική ηχητική παρουσίασή του χρήστη στο πραγματικό περιβάλλον του συνομιλητή του γίνεται με όρους τοποθέτησης

στον τρισδιάστατο ηχητικό χώρο. Στο παράδειγμα το Cohen [24], γίνεται προσπάθεια «ευθυγράμμισης» των αμφιωτικών σημάτων με τεχνητά χωρικά τοποθετημένα σήματα τα οποία συνθέτονται από τη συνέλιξη μονοωτικών σημάτων με ζευγάρια κατευθυντικών συναρτήσεων. Ο παραπάνω μηχανισμός χρησιμοποιείται ως εργαλείο για την υλοποίηση τηλεπαρουσίας με την οποία αναθέεται σε ένα ρομπότ να ενεργεί για λογαριασμό του ανθρώπου. Ο έλεγχος γίνεται απομακρυσμένος από τον άνθρωπο ο οποίος ως πιλότος φοράει αισθητήρες ενεργοποίησής του. Η τηλεπαρουσία σύμφωνα με το συγγραφέα, δίνει τη δυνατότητα στον άνθρωπο πιλότο να μπει σε επικίνδυνα περιβάλλοντα όπως για παράδειγμα πυρκαγιές, ή πυρηνικούς σταθμούς.

Στο δεύτερο παράδειγμα τηλεπαρουσίας [47], ο ακροατής βρίσκεται σε επαυξημένο περιβάλλον και ο άλλος χρήστης βιώνει την εμπειρία της τηλεπαρουσίας με τη χρήση αμφιωτικής εγγραφής σε πραγματικό χρόνο. Στην αμφιωτική τηλεφωνία ενεργοποιείται η τηλεπαρουσία με την μετάδοση του τρισδιάστατου ηχοτοπίου στο δέκτη. Με τη χρήση τεχνικών HRTFs και μοντελοποίηση της ακουστικής του χώρου, δημιουργείται το τρισδιάστατο εικονικό περιβάλλον, στο οποίο τοποθετείται ο ομιλητής, ουσιαστικά, «έξω» από το κεφάλι του ακροατή, μέσα στον εικονικό ηχητικό χώρο. Η παραπάνω εφαρμογή υποστηρίζει τη χρήση περισσότερων από δύο συνομιλητές.

Η φυσικότητα και η ευκολότερη κατανόηση από τον χρήστη συνεδριάσεων με πολλούς συνομιλητές αποτελεί το ζητούμενο της εφαρμογής ηχητικής επαύξεσης συνεδριάσεων του Virolainen [135]. Αυτό υποστηρίζει πως επιτυγχάνεται με τη χωρική τοποθέτηση των συνομιλητών στον τρισδιάστατο ηχητικό χώρο.

Στο πεδίο της επικοινωνίας, υπάρχει η χρήση ηχητικής επαύξεσης σε εφαρμογή συνεδρίασης με σκοπό την χωρική τοποθέτηση συνομιλητών στον τρισδιάστατο ηχητικό χώρο [135]. Σκοπός είναι να βοηθήσει τον χρήστη να ξεχωρίσει τους συνομιλητές κάνοντας τη συνεδρίαση πιο φυσική και ευκολότερη στην κατανόηση από το χρήστη. Γενικά, η βελτίωση της κατανόησης του πραγματικού περιβάλλοντος μέσα από την ηχητική επαύξηση, αποτελεί στόχο εφαρμογών τύπου ARA.

Ο Barde [7] πρότεινε την βελτίωση της οπτικής αντίληψης με χρήση τεχνολογίας τύπου ARA δημιουργώντας μια ακουστική αναπαράσταση για την αντιμετώπιση του προβλήματος των πολλαπλών ροών πληροφορίας και της παρουσίας τους από τις μικρές οθόνες των φορητών συσκευών. Η εφαρμογή χρήσης ακουστικών φράσεων που αναπαράγονται μέσω ακουστικών τύπου Bone-Conduction, χωρίς να απομονώνεται ο ακροατής από το πραγματικό περιβάλλον, ενισχύουν την οπτική απόδοση. Οι χρόνοι απόκρισης και οι χρόνοι εύρεσης του στόχου μειώθηκαν σημαντικά με τη χρήση αμφιωτικών χωρικών σημάτων σε σύγκριση με τη μονοφωνική αναπαραγωγή. Αυτό έδειξε ότι ένα ηχητικό σήμα με κατευθυντικές πληροφορίες βελτιώνει εργασίες αναζήτησης ή προειδοποίησης του χρήστη κατευθύνοντας την προσοχή του. Επίσης, σε άλλο παράδειγμα εφαρμογής, γίνεται χρήση ακουστικών και ισοστάθμισης σε πραγματικό χρόνο [105] με σκοπό την εξασθένιση του ψευδοακουστικού περιβάλλοντος σε σύγκριση με την ακρόαση χωρίς ακουστικά, και τη βελτίωση της κατανόησης του λόγου και της μουσικής ποιότητας σε πραγματικό χρόνο.

Η επόμενη κατηγορία που παρουσιάζεται, είναι εκείνη των εφαρμογών τύπου ARA που προορίζονται για καθημερινές εργασίες. Το πρώτο παράδειγμα δείχνει πως η τεχνολογία ARA,

εισάγεται σε μια εφαρμογή διαχείρισης ημερολογίου [137]. Σε αυτή την εφαρμογή ημερολογίου, τοποθετούνται τα προγραμματισμένα συμβάντα σε μια χωρική ηχητική αναπαράσταση. Η τοποθέτηση των συμβάντων γίνεται με την αντιστοίχιση της πληροφορίας της ώρας σε θέσεις στον εικονικό χώρο γύρω από το κεφάλι του ακροατή δηλαδή για παράδειγμα συνάντηση στις 3 το μεσημέρι, ακούγεται από δεξιά ενώ στις 6 το απόγευμα, ακούγεται από τα αριστερά.

Μια άλλη προσέγγιση της παραπάνω εφαρμογής, αποτελεί το έργο «Nomadic Radio project» [113, 114] το οποίο παρουσιάζει μια ιδέα φορητής ηχητικής πλατφόρμας η οποία παρέχει πρόσβαση σε όλες τις υπηρεσίες πληροφοριών και μηνυμάτων του υπολογιστή. Τα μηνύματα παρουσιάζονται με τη μορφή συνθετικού λόγου και τοποθετούνται ηχητικά σε ένα εικονικό τρισδιάστατο ηχητικό χώρο γύρω από το κεφάλι του χρήστη με βάση το χρόνο άφιξης και ανάλογα με την χρησιμότητα του μηνύματος τοποθετείται και σε άλλη απόσταση στον τρισδιάστατο εικονικό χώρο.

Στο πεδίο του εργασιακού περιβάλλοντος, έχει δοκιμαστεί η υπέρθεση εικονικών πηγών στο πραγματικό περιβάλλον, με ήχους για παράδειγμα ειδοποιήσεων ή οδηγιών. Η εφαρμογή Audio-Aura [97] αποτελεί ένα κατανεμημένο σύστημα που χρησιμοποιεί ηχητικό σήμα για τη μετάδοση εξατομικευμένης πληροφορίας σε ασύρματα ακουστικά. Στόχος είναι να παρέχει στο χρήστη ακουστικές φράσεις για τις ενέργειες των υπόλοιπων ατόμων στο χώρο εργασίας, όπως είναι οι πληροφορίες για το πριν πόση ώρα ένα άτομο έχει φύγει από το δωμάτιο, ή που βρίσκεται εκείνη τη στιγμή. Το σύστημα αναγνωρίζει την κίνηση και ανανεώνει τη βάση δεδομένων. Ορισμένες αλλαγές στη βάση δεδομένων έχουν αποτέλεσμα την αποστολή των ακουστικών σημάτων.

Το εγγενές χαρακτηριστικό της χωρικής τοποθέτησης του ακροατή μέσα στο επαυξημένο περιβάλλον όπως φαίνεται και από τα προηγούμενα παραδείγματα, δημιουργεί τις προϋποθέσεις για εμπλουτισμό πραγματικών περιβαλλόντων με εικονική πληροφορία, η οποία βελτιώνει την εμπειρία του θεατή. Μια ακόμα κατηγορία εφαρμογών τύπου ARA, αποτελούν οι αυτόματοι οδηγοί επαυξημένης πραγματικότητας ήχου.

Το έργο «Listen» [147], αποτελεί μια προσπάθεια για χρήση της εγγενούς ενσωμάτωσης της καθημερινής οπτικής και ακουστικής αντίληψης δημιουργώντας ένα προσαρμοσμένο στα μέτρα του ακροατή και εμπυθιστικό, ηχητικά επαυξημένο περιβάλλον κατά την επίσκεψη καλλιτεχνικής έκθεσης. Η πρόκληση του συγκεκριμένου έργου αποτελεί η προσαρμοστικότητα του περιεχομένου στα χαρακτηριστικά του κάθε επισκέπτη (ενδιαφέροντα, προτιμήσεις κίνησης), παρέχοντας ένα έξυπνο ηχητικό περιβάλλον.

Ο αυτόματος οδηγός μουσείων του Bederson [9, 26], είχε σκοπό να αντικαταστήσει τις κασέτες των ηχητικών οδηγιών εκθέσεων. Η πρωτότυπη εφαρμογή είχε τη δυνατότητα τυχαίας πρόσβασης ψηφιακών ηχητικών πηγών με τη χρήση ενός μικροεπεξεργαστή για τον έλεγχο των ηχητικών πηγών, ενός δέκτη υπέρυθρων για τον εντοπισμό και ενημέρωση του μικροεπεξεργαστή σχετικά με τη θέση του θεατή και ενός πομπού στην οροφή πάνω από κάθε αντικείμενο ώστε να ενημερώνει το σύστημα να αναπαράγει τις προηχογραφημένες περιγραφές του κάθε εκθέματος καθώς και να τις σταματάει μόλις ο θεατής απομακρυνθεί από αυτό.

Η εφαρμογή «SARIM (Sound Augmented Reality Interface)», έχει αντικείμενο την επαύξηση της μουσειακής επίσκεψης [59]. Σκοπός είναι η παροχή ενός ηχητικού οδηγού που

θα εμβυθίσει τον επισκέπτη σε μια ηχητική σκηνή αποτελούμενη από ήχους περιβάλλοντος και σχόλια σχετικά με τα εκθέματα, ελαχιστοποιώντας την προσπάθειά που χρειάζεται για να ανακαλύψει τα αντικείμενα και να αλληλεπιδράσει με το ηχητικό περιβάλλον. Η εφαρμογή αξιολογήθηκε σε πραγματικό περιβάλλον μουσείου και επιβεβαίωσε τον θετικό αντίκτυπο της ευχρηστίας, του εκπαιδευτικού και παιγνιακού χαρακτήρα της χρήσης της επαυξημένης πραγματικότητας ήχου σε σύγκριση με τους παραδοσιακούς ηχητικούς οδηγούς.

Παραδείγματα εφαρμογής τεχνολογίας επαυξημένης πραγματικότητας με σκοπό την βελτίωση της εμπειρίας πλοήγησης σε μουσειακό χώρο και ενσωμάτωση του ίδιου του επισκέπτη μέσα στο καλλιτεχνικό έργο, αποτελεί το έργο «Mapping the Symphony» και το έργο ακροάσεις του «National Science and Media Museum Gallery». Και στις δύο περιπτώσεις, ο επισκέπτης εξερευνά το χώρο και γίνεται ένα με την τέχνη, τα τεχνουργήματα, αρχαιακό ηχητικό υλικό και τις σχετιζόμενες ιστορίες. Το σύστημα χρησιμοποιεί την κάμερα του κινητού για τον εντοπισμό και αναγνώριση αντικειμένων και ακολούθως την ενεργοποίηση της τρισδιάστατης ηχητικής αναπαραγωγής.

Η εφαρμογή «Corona» [49] αποτελεί ένα ακόμα παράδειγμα ηχητικού οδηγού επαυξημένης πραγματικότητας ήχου εφαρμοσμένο σε έναν ιστορικό χώρο της πόλης Άαχεν της Γερμανίας μέσα στον οποίο γινόταν η τελετή στέψης. Στο εικονικό περιβάλλον του επαυξημένου χώρου, διαδραματιζόταν μια τελετή στέψης του 16ου αιώνα. Εικονικά πρόσωπα συζητούσαν για αυτή την τελετή και ο επισκέπτης με τη χρήση ειδικού εξοπλισμού μπορούσε να περιηγηθεί στο χώρο και ανάλογα με τη θέση του να ακούει τους αντίστοιχους διαλόγους.

Μια διαφορετική προσέγγιση με γνώμονα τον τουρισμό, αποτελεί η εφαρμογή με το όνομα «AudioNear» [14], η οποία σχεδιάστηκε με σκοπό να βοηθήσει την τουριστική εξερεύνηση του ανοικτού αστικού περιβάλλοντος με χρήση τεχνολογίας επαυξημένης πραγματικότητας ήχου για την παροχή ηχητικής πληροφορίας βασισμένη σε λόγο σχετικά με αξιοθέατα πάντα σε σχέση με τη θέση του επισκέπτη στο περιβάλλον. Τα αποτελέσματα της αξιολόγησης έδειξε την θετική δυναμική της επαυξημένης πραγματικότητας ήχου για την παροχή ηχητικής ενημερωτικής πληροφορίας και συναρπαστικής εμπειρίας.

Η επόμενη κατηγορία που παρουσιάζει αυξημένη δραστηριότητα με αρκετά παραδείγματα εφαρμογών, είναι η πλοήγηση στον δρόμο. Η χρήση συστημάτων τύπου ARA και η ικανότητα αντίληψης του εικονικού και του πραγματικού περιβάλλοντος από πεζούς αποτελεί άλλο ένα ζητούμενο. Παρακάτω θα παρουσιαστούν παραδείγματα εφαρμογών τέτοιου τύπου.

Αρχικά, παρουσιάζεται η μελέτη [132] όπου διερευνά τις επιδόσεις, την ασφάλεια και την εμπειρία των χρηστών ενός τρισδιάστατου ηχητικού συστήματος πλοήγησης πεζών. Η συσκευή δοκιμής αποτελείται από κινητή συσκευή, όπου υλοποιείται το λογισμικό πλοήγησης και η τρισδιάστατη επεξεργασία ήχου, καθώς και ένα πρότυπο ζεύγος ακουστικών επαυξημένης πραγματικότητας, με το οποίο ο εικονικός ήχος μπορεί να επαυξήσει το πραγματικό ηχητικό περιβάλλον γύρω από τον χρήστη. Σύμφωνα με τα αποτελέσματα της έρευνας, φαίνεται να βοηθάει τους πεζούς η χρήση τεχνολογίας ARA οδηγώντας τους στον προσορισμό με ευχέρεια και με ασφάλεια. Ταυτόχρονα, είναι σε θέση να εστιάσουν την προσοχή τους και σε άλλα ακουστικά συμβάντα, όπως η ακρόαση μουσικής από τη συσκευή ή η ομιλία ενός άλλου ατόμου.

Παράδειγμα εφαρμογής πλοήγησης για ποδηλάτες και για πεζούς, κάνει χρήση μουσικής και τεχνολογιών ARA για την καθοδήγησή τους διαπιστώνοντας πως η χωρική τοποθέτηση της μουσικής πληροφορίας μπορεί να λειτουργήσει ως ηχητική καθοδήγηση σε εφαρμογή πλοήγησης [3]. Η ενσωμάτωση της μουσικής ακρόασης με μια υπηρεσία πλοήγησης δίνει τη δυνατότητα στο χρήστη να ακολουθήσει την αντιληπτή κατεύθυνση της μουσικής για να φτάσει στον προορισμό του. Το σύστημα δοκιμάστηκε από πεζούς σε ένα κέντρο πόλης και από ποδηλάτες σε προάστιο. Τα αποτελέσματα έδειξαν ότι η χωρική τοποθέτηση της μουσικής στον τρισδιάστατο χώρο, μπορεί να καθοδηγήσει πεζούς και ποδηλάτες σε ένα προορισμό χωρίς να έχει προηγηθεί εκπαίδευση προσφέροντας μια ευχάριστη εμπειρία.

Η εφαρμογή «NAVINKO» [78] αποτελεί ένα κοινωνικό δίκτυο και ένα σύστημα πλοήγησης με χρήση επαυξημένης πραγματικότητας ήχου για ποδηλάτες στο Τόκιο. Χρησιμοποιείται από κινητές συσκευές και οι χρήστες μπορούν να μοιράζονται σημεία ενδιαφέροντος. Η ηχητική διεπαφή του συστήματος, χρησιμοποιεί τεχνολογία επαυξημένης πραγματικότητας ήχου για την χωρική πλοήγηση του χρήστη, τον εντοπισμό των σημείων ενδιαφέροντος καθώς και άλλων χρηστών.

Οι εφαρμογές σοβαρού σκοπού τύπου ARA, αποτελούν μια κατηγορία με αρκετά παραδείγματα τα οποία κυρίως στοχεύουν σε εκπαιδευτικούς σκοπούς και σε δυσκολίες που αντιμετωπίζουν τα άτομα με δυσκολία όρασης. Παρακάτω δίνεται η παρουσίαση τέτοιου είδους εφαρμογών.

Ένα παράδειγμα εφαρμογής τεχνολογίας ARA με χρήση χωρικής ακουστικής απεικόνισης (Spatial Auditory Display) για αξιολόγηση από πλήρωμα αεροπλάνου σε εξομίωση νυχτερινής αποστολής πτήσης υλοποιήθηκε από τον Shilling [118]. Η αξιολόγηση περιελάμβανε τη σύγκριση μιας τυπικής οπτικής απεικόνισης, με μια χωρικά επαυξημένη ακουστική απεικόνιση. Τα αποτελέσματα έδειξαν πως ένα βελτιωμένο ακουστικό σύστημα, μπορεί να επιφέρει βελτίωση στην πλοήγηση.

Μια άλλη κατηγορία εφαρμογών σοβαρού σκοπού, αποτελούν εκείνες που προορίζονται για τη βελτίωση πλοήγησης ατόμων με πρόβλημα όρασης. Η εφαρμογή «NAVIG» [62], αποτελεί ένα σύστημα βοήθειας επαυξημένης πραγματικότητας για άτομα με δυσκολία όρασης. Στόχος είναι η ενίσχυση της αυτονομίας τους και της κίνησής τους μέσα σε ένα πλαίσιο αίσθησης του άμεσου περιβάλλοντος και της καθοδήγησής τους ως πεζοί. Γίνεται χρήση δορυφορικών συντεταγμένων, εικόνες και αισθητήρων για τη μέγιστη ακρίβεια εύρεσης θέσης. Γίνεται επιπλέον αναγνώριση εικόνες και χωρική τοποθέτηση ήχου που βοηθούν την επίτευξη μεγάλων ή μικρών διαδρομών. Επιπλέον ειδικό χειριστήριο, έχει αναπτυχθεί για τη βέλτιστη απόδοση του χρήστη.

Στην ίδια κατηγορία εφαρμογών σοβαρού σκοπού, αποτελεί το παρακάτω παράδειγμα εφαρμογής [13] το οποίο χρησιμοποιεί σημεία ενδιαφέροντος τύπου POI (Points of Interest) από παγκόσμια βάση δεδομένων. Η ηχητική πληροφορία παρουσιάζεται υπό τη μορφή λόγου. Ο εξοπλισμός που επιλέχθηκε, ήταν ένα κινητό τηλέφωνο από το οποίο έγινε χρήση των διαθέσιμων αισθητήρων για την εύρεση της κατεύθυνσης και της θέσης του ακροατή στον πραγματικό περιβάλλον καθώς και η χωρική αναπαραγωγή της ηχητικής πληροφορίας.

Ο Belkhamza, υλοποίησε ένα σύστημα πλοήγησης για άτομα με προβλήματα όρασης στο

οποίο, το οπτικό περιβάλλον του ατόμου, λαμβάνεται από κάμερα και μετατρέπεται σε ηχητική σκηνή που παρέχει πληροφορίες για εμπόδια ή πιθανό κίνδυνο [146]. Ο εξοπλισμός που χρησιμοποιείται είναι μια κάμερα, μια φορητή υπολογιστική μονάδα και ένα ζευγάρι ακουστικά.

Το επόμενο πεδίο εφαρμογών σοβαρού σκοπού αποτελεί η εκπαίδευση ατόμων με δυσκολία στην όραση. Ο Inman παρουσίασε μέθοδο [54] για διδασκαλία ατόμων με δυσκολία στην όραση με την οποία, περιβάλλοντα εξομοίωσης χρησιμοποιούνται για εκπαίδευση ηχητικής ταυτοποίησης, χωρική τοποθέτηση και εντοπισμό ηχητικών πηγών. Για την υλοποίηση της εξομοίωσης, χρησιμοποιήθηκε μια υπολογιστική μονάδα, ένας ζεύγος ακουστικών, ένα σύστημα εντοπισμού κεφαλής, και ένα χειριστήριο. Η κίνηση του κεφαλιού του μαθητή, μετέβαλε δυναμικά την ακουστική των ηχητικών πηγών του εικονικού περιβάλλοντος. Οι εργασίες για την εκμάθηση του συστήματος περιελάμβαναν την ταυτοποίηση και αντίδραση στις ηχητικές πηγές της εξομοίωσης με συγκεκριμένο τρόπο, ολοκλήρωση μιας δράσης όπως η διάσχιση ενός δρόμου στον εικονικό κόσμο και ολοκλήρωση μιας δράσης λαμβάνοντας ηχητική ανάδραση από ένα σύνολο ενεργειών δράσης που έπρεπε να ακολουθηθεί. Το περιβάλλον επίσης, απεικονιζόταν σε μια οθόνη όπου ο εκπαιδευτής ακολουθούσε τις δράσεις του μαθητή.

Στο ίδιο πεδίο μπορεί να ενταχθεί και η εφαρμογή για την παρουσίαση γεωγραφικών πληροφοριών σε άτομα με αδυναμία όρασης με τη χρήση εικονικού χάρτη αφής πάνω σε φωτιζόμενη επιφάνεια [67]. Η καταγραφή των κινήσεων, γίνεται από κάμερα η οποία βρίσκεται σε οροφή. Τα χαρακτηριστικά αφής που καταγράφονται, μεταφέρονται υπό τη μορφή ηχητικών σημάτων και περιγράφουν τα αντίστοιχα αντικείμενα που αγγίχτηκαν.

Ο τομέας της διασκέδασης αποτελεί ένα ενεργό πεδίο εφαρμογών τύπου ARA στον οποίο αρχικά στοχεύει η παρούσα διατριβή. Εφαρμογές που εντάσσονται σε αυτό το πεδίο αποτελούν τα ηχητικά παιχνίδια επαυξημένης πραγματικότητας, τα οποία αποτελούν την εφαρμογή στόχο για την παρούσα διατριβή. Η σύνδεση των εγγενών χαρακτηριστικών της επαυξημένης πραγματικότητας με τα χαρακτηριστικά των ηχητικών παιχνιδιών έχουν δημιουργήσει αυτή την κατηγορία παιχνιδιών. Οι υποκατηγορίες που άτυπα συναντώνται στα ηχητικά παιχνίδια τύπου ARA είναι η πλοήγηση, η εξερεύνηση, η διασκέδαση [111].

Το παιχνίδι «Guided by Voices» [74], είναι ένα παράδειγμα ηχητικού παιχνιδιού το οποίο χρησιμοποιεί ένα σύστημα επαυξημένης πραγματικότητας ήχου για την παροχή ηχητικού περιεχομένου συνυφασμένο με θέση στον πραγματικό χώρο. Οι ήχοι εξαρτώνται από την στιγμιαία θέση του θεατή και το ιστορικό των περιοχών του κάθε παίκτη. Το σύστημα βασίζεται σε φορητές υπολογιστικές μονάδες και σύστημα ραδιοσυχνοτήτων για τον εντοπισμό της θέσης των πηγών, τοποθετημένο στο περιβάλλον. Το σενάριο του παιχνιδιού δομείται μέσα από ένα μεσαιωνικό κόσμο όπου ο παίκτης περιπλανιέται μέσα από ένα ηχοτοπίο από τρολ, δράκους και μάγους. Διάφοροι ήχοι από το σκηνικό παρουσιάζονται καθώς ο παίκτης εξερευνά το περιβάλλον περνώντας από τη φωλιά του δράκου, και τη γέφυρα ενός τρολ, όπου συναντάει φίλους ή εχθρούς και συγκεντρώνει αντικείμενα όπως ένα σπαθί και ένα έλατο. Ο τρόπος παιχνιδιού είναι εφάμιλλος των παραδοσιακών παιχνιδιών ρόλου. Η αφήγηση της ιστορίας αλλάζει ανάλογα με τη διαδρομή του παίκτη. Ο φορετός εξοπλισμός, παρακολουθεί την κατάσταση του παίκτη και ελέγχει αν έχει αποκτήσει τα απαραίτητα αντικείμενα της κάθε τοποθεσίας. Ο εμπλουτισμός του πραγματικού περιβάλλοντος μπορεί να πραγματοποιηθεί ε-

πειδή η κάθε τοποθεσία μπορεί να προσφέρει πολλαπλούς σκοπούς και εναλλαγή ρόλων σε διαφορετικές καταστάσεις.

Το παιχνίδι «Sleuth» [29], βασίζεται σε ένα σύστημα ηχητικής έκθεσης (Auditory Display System) στο οποίο ο παίκτης ως ερευνητής, τοποθετείται στη μέση μια σκηνής εγκλήματος όπου θα πρέπει να χρησιμοποιήσει την πληροφορία που δίνεται από ηχητικές φράσεις και να διερευνήσει την πορεία του εγκλήματος και να δηλώσει τον δολοφόνο, το όπλο και την τοποθεσία. Το παιχνίδι διαδραματίζεται σε πραγματικό χώρο αποτελούμενο από δωμάτια στα οποία έχει εγκατασταθεί σύστημα εντοπισμού θέσης. Ο παίκτης φοράει ειδικό ακουστικό εξοπλισμό με τον οποίο εισέρχεται στο επαυξημένο ηχητικό χώρο του παιχνιδιού. Η τυπική αλληλεπίδραση του παίκτη με το παιχνίδι ξεκινάει με το χρήστη να ακούει μια αφήγηση, που περιγράφει το σενάριο, καταλήγοντας στη δολοφονία και καθορίζοντας το ρόλο του παίκτη ως ντετέκτιβ. Τα ονόματα των καλεσμένων, ο κατάλογος των πιθανών όπλων, και τα ονόματα των δωματίων στην έπαυλη, παρέχονται σε χαρτί για να ενεργούν ως το σημειωματάριο του ντετέκτιβ. Παρέχεται επίσης ένα σχέδιο δαπέδου για τα δωμάτια του σπιτιού, αλλά χωρίς λεπτομέρειες για το περιεχόμενο ή τα ονόματα των δωματίων. Ο ντετέκτιβ στη συνέχεια πλοηγείται από δωμάτιο σε δωμάτιο, έχοντας την επιλογή να πατήσει μάλιστα ακούσει κάτι που τραβάει την προσοχή του, και να κινηθεί προς τους ήχους του ενδιαφέροντος του. Με αυτόν τον τρόπο, ο ντετέκτιβ προσπαθεί να εξερευνήσει το σπίτι και να συγκεντρώσει αρκετά στοιχεία για να φτιάξει μια κατηγορία. Διάφοροι ήχοι και συζητήσεις του περιβάλλοντος μεταξύ των άλλων επισκεπτών στο σπίτι χρησιμεύουν ως απόδειξη στην υπόθεση, καθώς και για την παροχή στοιχείων πλοήγησης στον ντετέκτιβ. Για παράδειγμα, αν ο χαρακτήρας του καθηγητή Plum είναι γνωστό ότι έχει κοιμηθεί ήσυχα τη νύχτα, τότε είναι απίθανο να είναι ο δολοφόνος. Σε μια άλλη περίπτωση, ο ήχος των σφαιριδίων του μπιλιάρδου συγκρούεται μεταξύ τους και ο ληστής των παικτών δείχνει ότι ο ντετέκτιβ μπήκε στο δωμάτιο παιχνιδιών. Μόλις ο ντετέκτιβ αποφασίσει ότι έχει αποκτηθεί αρκετές πληροφορίες για να απομονώσει τον δράστη, ο ντετέκτιβ δηλώνει ένα δολοφόνο, ένα όπλο και μια τοποθεσία. Αν η εικασία είναι σωστή, ο ντετέκτιβ κερδίζει το παιχνίδι. Σε περίπτωση λάθους, ο ντετέκτιβ θα πρέπει να επανεξετάσει τα δωμάτια του σπιτιού για να συγκεντρώσει περισσότερες αποδείξεις.

Το ηχητικό παιχνίδι επαυξημένης πραγματικότητας ήχου «SoundPacman» [20], στο οποίο δημιουργείται ένας χάρτης τύπου Pacman [19] ο οποίος τοποθετείται πάνω σε πραγματικό χάρτη. Η πληροφορία του παιχνιδιού είναι μόνο ηχητική. Το παιχνίδι, βασίζεται στην τοποθεσία, το οποίο απαιτεί σύνδεση στο διαδίκτυο και ενεργοποιημένο δέκτη «GPS» στην κινητή συσκευή του παίκτη. Ο χώρος του παιχνιδιού καθορίζεται στην αρχικοποίηση του παιχνιδιού, λαμβάνοντας υπόψη τα πραγματικά τμήματα του δρόμου γύρω από τη θέση του χρήστη ως πιθανούς διαδρόμους για τον παίκτη. Ο χρήστης συλλέγει τα εικονικά μπισκότα που βρίσκονται στους δρόμους. Η κίνηση των εικονικών αντικειμένων διεκπεραιώνεται από τη μηχανή χωρικής τοποθέτησης στον τρισδιάστατο ηχητικό χώρο. Τα κινούμενα εικονικά αντικείμενα ως φαντάσματα κυνηγούν να πιάσουν τον παίκτη. Βασικό χαρακτηριστικό του παιχνιδιού είναι η ότι οι απαραίτητες πληροφορίες για το παιχνίδι παρέχονται στον παίκτη μέσω τρισδιάστατου ήχου. Οι παίκτες συνειδητοποιούν τη θέση των φαντασμάτων και των μπισκότων μέσω της αντίληψης των συγκεκριμένων ήχων που παράγονται από αυτά λαμβάνοντας τα χωρικά χαρα-

κτηριστικά από την εικονική τοποθέτησή τους στο πραγματικό περιβάλλον. Επιπλέον, στην υλοποίηση του παιχνιδιού, έγινε αξιολόγηση με χρήση αισθητήρων εγκεφαλογραφήματος, με θετικά αποτελέσματα ως προς τη συνεισφορά της ηχητικής επαύξησης στο επίπεδο εμπύθισης του παίκτη.

Ένα παράδειγμα βασισμένο σε επαυξημένη πραγματικότητα ήχου και σε πρωτότυπα παιχνίδια σοβαρού σκοπού αποτελεί μια έκδοση παιχνιδιού χαμένου θησαυρού με σκοπό την βελτίωση της τουριστικής εμπειρίας του επισκέπτη [31]. Σκοπός του παιχνιδιού είναι η δημιουργία διαδραστικής και εμπυθιστικής εμπειρίας κατά την οποία ο επισκέπτης περιηγείται σε διαφορετικά σημεία ενδιαφέροντος. Η επαυξημένη πραγματικότητα ήχου σε συνεργασία με το σενάριο και την ηχοποίηση δημιουργούσαν κίνητρο στον επισκέπτη να εξερευνήσει την περιοχή. Ο παίκτης φοράει ένα ζεύγος ακουστικών με τα οποία ακούει το ηχητικό υλικό, αλληλεπιδρά με το ηχητικό περιβάλλον και απαντάει ερωτήσεις για να ολοκληρώσει το παιχνίδι. Για την καθοδήγηση του παίκτη χρησιμοποιείται μουσική ή κάποιο ηχοτοπίο σχετικό με την αναζητούμενη τοποθεσία δράσης. Για την ηχητική επαύξηση, χρησιμοποιείται μέθοδος τρισδιάστατης χωρικής αναπαραγωγής των ήχων παράλληλα με τη χρήση αμφιωτικής τεχνολογίας.

Παράδειγμα χρήσης της επαυξημένης πραγματικότητας για παιχνίδι αθλητικού χαρακτήρα σοβαρού σκοπού αποτελεί η μετατροπή ενός παιχνιδιού μπάντμιντον σε παιχνίδι επαυξημένης πραγματικότητας ήχου κατάλληλο και για άτομα με δυσκολία στην όραση [63]. Η υλοποίηση του σεναρίου, περιλαμβάνει την υλοποίηση ειδικής μπάλας με ενσωμάτωση αισθητήρα κίνησης εννέα αξόνων καθώς και ειδικής ρακέτας με αντίστοιχο αισθητήρα θέσης ασύρματο. Ο ήχος μεταδίδεται με χρήση στερεοφωνίας από ακουστικά ή ηχεία. Σε όλη τη διάρκεια του παιχνιδιού υπάρχει ηχητική ανάδραση από την ειδική μπάλα και τη ρακέτα δίνοντας τη δυνατότητα σε άτομα με αδυναμία όρασης να πάρουν μέρος στο παιχνίδι. Η αξιολόγηση του παιχνιδιού έγινε από άτομα με πρόβλημα στην όραση καθώς και από άτομα με κανονική όραση. Τα αποτελέσματα δείχνουν πως το παιχνίδι ήταν διασκεδαστικό για όλους. Σε αυτό βοήθησε η χρήση πραγματικής ρακέτας και η χρήση απλής στερεοφωνίας για την οδήγηση του ήχου της μπάλας.

Η εφαρμογή «Music Blocks» [85] αποτελεί ένα ηχητικό και μουσικό παιχνίδι επαυξημένης πραγματικότητας με εκπαιδευτική προέκταση. Σε αυτή την εφαρμογή χρησιμοποιούνται ειδικοί κύβοι με ενσωματωμένους αισθητήρες σχεδιασμένοι για να επαυξήσουν ηχητικά, παιχνίδια αξιολόγησης γνωστικών και κινητικών δεξιοτήτων.

Προηγούμενες δουλειές του συγγραφέα, αποτελούν τα ακόλουθα παιχνίδια. Το παιχνίδι «Eidola» [92], αντιπροσωπεύει ένα πρωτότυπο ηχητικό παιχνίδι το οποίο λαμβάνει τα πλεονεκτήματα των χαρακτηριστικών της επαυξημένης πραγματικότητας για την υλοποίηση πολύπλοκων σεναρίων ηχητικών παιχνιδιών. Το πρωτότυπο ηχητικό παιχνίδι υλοποιήθηκε ως μια οπτικοακουστική διαδραστική εγκατάσταση επιτρέποντας την περαιτέρω σύνδεση του φυσικού χώρου ως δευτερεύουσα συνιστώσα για την δημιουργία του πραγματικού περιβάλλοντος του παιχνιδιού. Μια ακολουθία από τεστ έδειξε ότι το προτεινόμενο πρωτότυπο μπορεί να υποστηρίξει πολύπλοκα σενάρια παιχνιδιών καθώς είναι διαθέσιμοι οι απαραίτητοι τρόποι αλληλεπίδρασης.

Το πρωτότυπο παιγνιακό σενάριο που υλοποιήθηκε, βασίστηκε στην παρουσία ενός αριθ-

μού από άορατα όντα τα οποία ζουν μέσα στο δωμάτιο του παιχνιδιού. Το όνομα των όντων είναι «Είδωλα». Ο παίκτης του ηχητικού παιχνιδιού επαυξημένης πραγματικότητας έχει την αποστολή να εισέλθει στο δωμάτιο και να νικήσει όλα τα όντα τα οποία κινούνται μέσα στο δωμάτιο ακολουθώντας μη γραμμικές τροχιές στον τρισδιάστατο χώρο. Για να επιτευχθεί αυτό, ο παίκτης έχει να τραβήξει μια εικονική σκανδάλη (να πατήσει ένα ειδικό κουμπί) όταν η θέση του συμπίπτει (μέσα σε ένα σταθερό όριο) με τη στιγμιαία θέση του “Ειδώλου” στην τρισδιάστατη τροχιά που ακολουθεί.

Το ηχητικό παιχνίδι «Eidola Multiplayer» [94], είναι ένα πρωτότυπο παιχνίδι το οποίο υλοποιεί την έννοια της ελεγχόμενης μεικτής πραγματικότητας με σκοπό την επαύξηση του ηχητικού περιβάλλοντος του κάθε παίκτη. Το πρωτότυπο υλοποιείται ως ένα σύμπλεγμα οπτικοακουστικών εγκαταστάσεων πολλαπλών παικτών στο οποίο υπάρχει αλληλοσύνδεση μεταξύ τους με σκοπό την επικοινωνία της κατάστασης επιλεγμένων παραμέτρων σε πραγματικό χρόνο. Το πρωτότυπο εμφανίζει σημαντική συνάφεια με τον γνωστό τομέα των «on-line» παιχνιδιών πολλαπλών παικτών, αλλά καινοτομεί στο γεγονός ότι η αλληλεπίδραση των παικτών πραγματοποιείται σε περιβάλλοντα ήχου επαυξημένης πραγματικότητας.

Το σενάριο σχεδιάστηκε ως προέκταση του ηχητικού παιχνιδιού επαυξημένης πραγματικότητας «Είδωλα» [92]. Το σενάριο του παιχνιδιού εμπλουτίστηκε, εισάγοντας δύο παίκτες που κυνηγούσαν τα πλάσματα “Είδωλα” σε δύο απολύτως όμοια δωμάτια. Η εξολόθρευση των «Ειδώλων» πραγματοποιείται όταν η θέση του παίκτη συμπίπτει με την τρέχουσα θέση του «Ειδώλου». Ωστόσο, ένα κινούμενο ον δεν ελέγχεται από τη μηχανή του παιχνιδιού. Αντ’ αυτού, αντιστοιχεί στον απομακρυσμένο χρήστη, ορίζοντας έτσι την έννοια της αλληλεπίδρασης στη μικτή πραγματικότητα ήχου. Επιπλέον, ορισμένα αντικείμενα (όπως ραδιόφωνα, τηλεοράσεις κλπ.) βρίσκονται σε συγκεκριμένες θέσεις και στα δύο δωμάτια. Αυτά τα αντικείμενα παράγουν φυσικούς ήχους που αντικατοπτρίζουν το πραγματικό περιβάλλον ήχου. Επιπλέον, οι ήχοι περιβάλλοντος αναπαράγονται μέσω της εικονικής συνιστώσας του ήχου, χρησιμοποιώντας προ-εγγεγραμμένους ήχους και προσαρμοστικό έλεγχο του κέρδους αναπαραγωγής με βάση τη στιγμιαία θέση του παίκτη.

Πρόσφατο παράδειγμα ηχητικών παιχνιδιών επαυξημένης πραγματικότητας αποτελεί και η τριλογία παιχνιδιών «Audio Legends» [111]. Ο σχεδιασμός τους ερευνά την ηχητική διάδραση, όχι μόνο ως στοιχείο της πλοήγησης, αλλά και ως τρόπος αλληλεπίδρασης του ακροατή μέσα από νεύματα και χειρονομίες για την επιτυχή ολοκλήρωση του παιχνιδιού. Το σενάριο των παιχνιδιών σχεδιάστηκε με γνώμονα τρεις ιστορίες του νησιού της Κέρκυρας. Ειδικότερα η σχεδίαση του σεναρίου βασίστηκε στους τοπικούς μύθους σχετικά με την προστασία του νησιού από τον Άγιο Σπυρίδωνα σε τρεις διαφορετικές περιπτώσεις.

- Τον 16ο αιώνα καθοδήγησε ιταλικά πλοία που κουβαλούσαν σιτάρι μέσα από καταιγίδα στην Κέρκυρα όπου είχε πληγεί από λιμό.
- Τον 17ο αιώνα κατάφερε και έδιωξε την πανούκλα έξω από το νησί με ένα σταυρό η οποία περιγράφεται ως ένα ον όπου είναι μισή γυναίκα και μισή το τέρας της ασθένειας.
- Τον 18ο αιώνα υπερασπίστηκε το νησί ενάντια στη ναυτική πολιορκία των Τούρκων καταστρέφοντας το μεγαλύτερο μέρος του στόλου με καταιγίδα.

Λαμβάνοντας υπόψιν τις παραπάνω ιστορίες, σχεδιάστηκε η τριλογία παιχνιδιών «Audio Legends».

- «Ο Λιμός»: Σε αυτό το παιχνίδι ο παίκτης έπρεπε αρχικά να εντοπίσει έναν Ιταλό ναύτη και κατόπιν να τον οδηγήσει στο σημείο τερματισμού.
- «Η Πανούκλα»: Οι παίκτες έπρεπε να εντοπίσουν ένα τέρας και στη συνέχεια σε μία μονομαχία μαζί του να το νικήσουν
- «Η Πολιορκία»: Οι παίκτες έπρεπε να βρουν το σημείο που γινόταν ο βομβαρδισμός και να προστατευθούν από τους κανονιοβολισμούς.

Η μηχανική των παραπάνω παιχνιδιών αποτελείται από δύο φάσεις διάδρασης: την εξερεύνηση και την δράση. Στην πρώτη φάση ο παίκτης πρέπει να εντοπίσει τη θέση του εικονικού ήχου και να κατευθυνθεί κοντά σε αυτόν. Στη δεύτερη φάση, όταν ο παίκτης εισέρχεται σε ακτίνα πέντε μέτρων από το σημείο δράσης, χρειάζεται να επικεντρωθεί στη χρήση χειρονομιών και την ανάδραση που λαμβάνει από τα ηχητικά ερεθίσματα του εικονικού ακουστικού χώρου. Σε καμία φάση δεν υπάρχει οπτική επαφή με την διεπαφή του παιχνιδιού αλλά μόνο η χρήση του εξοπλισμού ως μέσο ηχητικής αναπαραγωγής και μέσου διεπαφής για την αναγνώριση χειρονομιών για το εκάστοτε σενάριο [111].

Μια πρωτότυπη χρήση τεχνολογίας ARA αποτελεί το παράδειγμα συναυλίας ηλεκτροακουστικής μουσικής [95]. Σε αυτό το παράδειγμα, υλοποιείται ένα δίκτυο επαυξημένης πραγματικότητας ήχου το οποίο επαυξάνει το ακουστικό περιβάλλον μιας συναυλίας ηλεκτροακουστικής μουσικής με το ηχητικό περιβάλλον μιας πόλης. Ένα σημαντικό σημείο στην παραπάνω επαύξηση είναι ο τρόπος που τα σημεία της πόλης ως «Ηχητικές Μικρογραφίες» αναπαράγονται μέσα στη συναυλία. Η προσέγγιση της χωρικής τοποθέτησης ηχητικά μέσα στη συναυλία, βασίζεται στις γεωγραφικές συντεταγμένες των ηχητικών σημείων σε σχέση με τον τόπο της συναυλίας. Η αναπαραγωγή του επαυξημένου ήχου γίνεται από ένα πολυκαναλικό σύστημα το οποίο συνδυάζει τεχνικές τρισδιάστατης αναπαραγωγής του ήχου στον χώρο με τη χρήση αμφιωτικής τεχνολογίας, Ambisonic τεχνολογίας και πολυκαναλικής ηχητικής διάχυσης. Η αξιολόγηση του παραπάνω συστήματος έγινε από το κοινό και από τους μουσικούς. Τα αποτελέσματα έδειξαν την θετική αντιμετώπιση του κοινού σε ένα τέτοιο είδος συναυλίας και ειδικότερα από τους μουσικούς δίνοντάς τους νέες οδούς για δημιουργικότητα μέσα από την μουσική σύνθεση και τον πειραματισμό συμφωνώντας πως θα ήθελαν να ξαναχρησιμοποιήσουν τεχνολογία τύπου ARA.

Παρόμοιο παράδειγμα χρήσης τεχνολογίας επαυξημένης πραγματικότητας ήχου με το παραπάνω, αποτελεί το έργο SonicMaps [101]. Πρόκειται για μια εφαρμογή ηχητικής γεωτοποθέτησης. Με τη χρήση των αισθητήρων GPS του κινητού, συνθέτεται το εικονικό περιβάλλον το οποίο τοποθετείται χωρικά πάνω στο πραγματικό. Η επιλογή των ήχων που θα τοποθετηθούν βρίσκονται σε μια κεντρική βάση δεδομένων, όπου ανάλογα με κριτήρια όπως η ώρα ή η θερμοκρασία, ενεργοποιούνται. Η θέση του χρήστη και τα δεδομένα των αισθητήρων στέλνονται σε έναν διακομιστή στον οποίο υπάρχει πρόσβαση από έναν συναυλιακό χώρο στον οποίο μεταφράζονται μουσικά δημιουργώντας μια σύνδεση μεταξύ απομακρυσμένων χώρων.

Πρέπει να προστεθεί στις παραπάνω καλλιτεχνικές προσεγγίσεις η εφαρμογή «Sound-Delta», η οποία αποτελεί ένα έργο για την παραγωγή καλλιτεχνικού τρισδιάστατου ηχοτοπίου για δημόσιες εκθέσεις. Η αρχιτεκτονική του συστήματος διάχυσης βασίζεται σε μια υλοποίηση τεχνικής «Ambisonic» για αμφιωτική αναπαραγωγή. Στο σενάριο, ο χώρος διαιρείται σε ακουστικά κελιά μέσα στα οποία τοποθετούνται οι ηχητικές πηγές. Ο ακροατής περιηγείται στον χώρο και ακούει τον εικονικό ηχητικό κόσμο με ειδικό ακουστικό εξοπλισμό.

Η εισαγωγή εικονικών κόσμων μέσα στο πραγματικό περιβάλλον τυπικών εφαρμογών όπως είναι η τηλεπαρουσίαση, η πλοήγηση, τα ηχητικά παιχνίδια ή συναυλία, αναδεικνύει τις νέες δυνατότητες που απορρέουν από τέτοιου είδους εφαρμογές. Η εφαρμοσμένη χρήση εννοιών όπως η ακουστική αντίληψη, η χωρική αντίληψη του ήχου, η σχεδίαση νέων μορφών ηχητικής διάδρασης όπως για παράδειγμα η χρήση χειρονομιών, αποτελούν στοιχεία τα οποία καταδεικνύουν την ανάγκη για εξέλιξη της ARA τεχνολογίας μέσα από γενικευμένα μοντέλα ηχητικής διάδρασης και ακουστικής αντίληψης.

## 2.4 Σύνοψη Κεφαλαίου

Σε αυτό το κεφάλαιο παρουσιάστηκε η στάθμη τεχνικής της ARA τεχνολογίας με την επισκόπησή των τεχνολογιών που χρησιμοποιούνται στα επαυξημένα περιβάλλοντα ήχου καθώς και παραδείγματα εφαρμογών. Η εξέλιξη των συσκευών κινείται παράλληλα με την εξέλιξη της επαυξημένης τεχνολογίας ήχου, προσφέροντας νέες δυνατότητες για ανάπτυξη πολύπλοκων ηχητικών σεναρίων. Η τεχνολογία ARA μπορεί να διαιρεθεί σε τρεις πυλώνες εξέλιξης. Ο πρώτος περιέχει τις τεχνολογίες αναπαραγωγής του πραγματικού περιβάλλοντος στα περιβάλλοντα τύπου ARA. Η εξέλιξη σε αυτές τις τεχνολογίες στηρίζεται στην ανάπτυξη των τεχνικών HT οι οποίες λαμβάνουν υπόψη τη φυσιολογία του αυτιού καθώς και την τεχνολογία του ακουστικού εξοπλισμού που χρησιμοποιείται. Ο δεύτερος πυλώνας περιέχει τη χωρική τοποθέτηση του εικονικού περιβάλλοντος. Οι τεχνολογίες που χρησιμοποιούνται στηρίζονται στην αμφιωτική τεχνολογία και στην χωρική τοποθέτηση ηχητικών πηγών στον τρισδιάστατο εικονικό ηχητικό χώρο. Ο τρίτος πυλώνας αποτελείται από την τεχνολογία σχεδίασης της ηχητικής διάδρασης. Σε αυτό εντάσσονται οι τεχνολογίες διάδρασης που υλοποιούνται μέσα σε περιβάλλοντα τύπου ARA. Η ενσωμάτωση αισθητήρων σε φορητές συσκευές επιτρέπουν την εξερεύνηση νέων τρόπων διάδρασης μέσα στα επαυξημένα περιβάλλοντα ήχου. Παραδείγματα εφαρμογών τύπου ARA δείχνουν την διαρκή εξέλιξη του τομέα της επαυξημένης πραγματικότητας ήχου με την παρουσίαση των νέων τεχνολογικών δυνατοτήτων μέσα από παραδείγματα νέου τύπου ηχητικών παιχνιδιών, μουσειακών περιηγήσεων, διαδραστικών πλοηγήσεων, και καλλιτεχνικών δραστηριοτήτων.



## Κεφάλαιο 3

# Αναλυτική μελέτη ARA μείξης

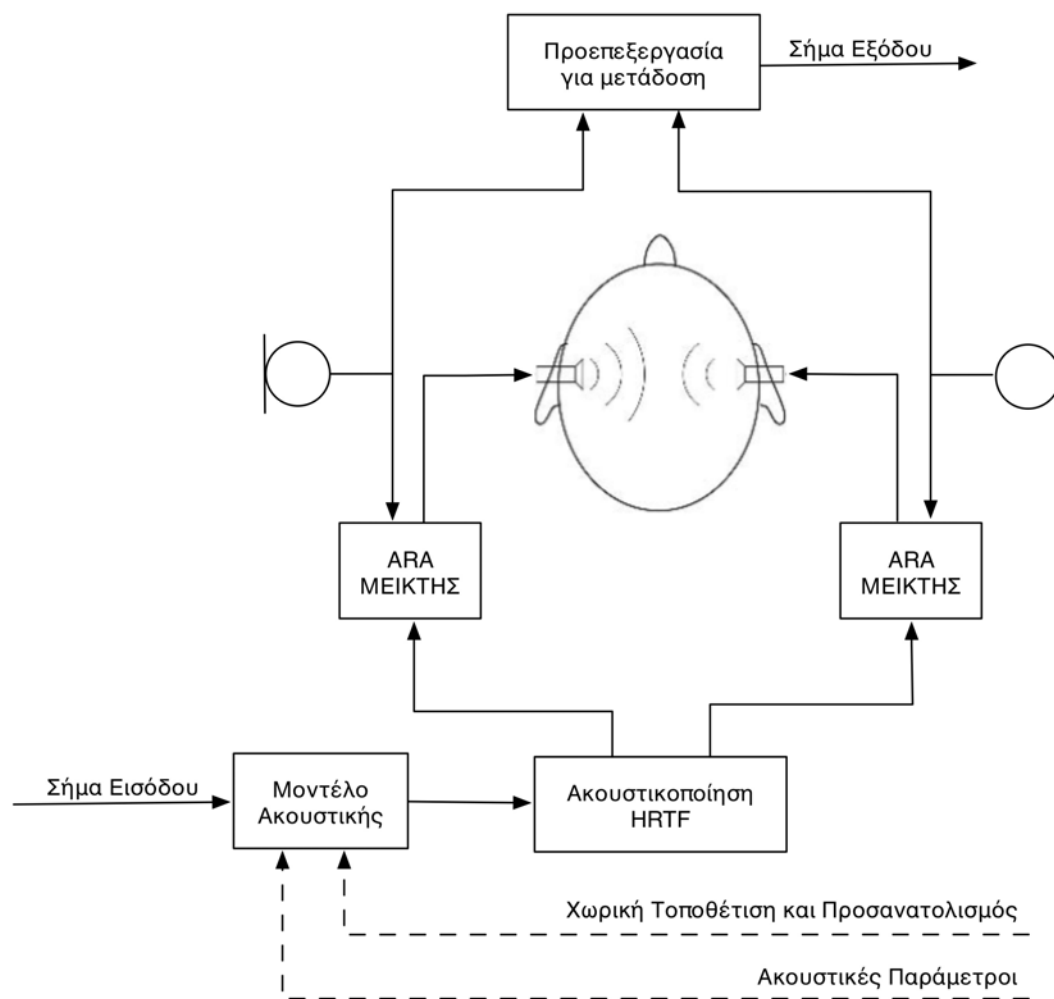
### 3.1 Εισαγωγή

Η μείξη του ψευδοακουστικού περιβάλλοντος, το οποίο αποτελεί την ηχητική αναπαράσταση του πραγματικού περιβάλλοντος, με το εικονικό περιβάλλον συνθέτουν τον πυρήνα του ARA συστήματος και ορίζεται ως ARA μείξη. Το τελικό αποτέλεσμα της ARA μείξης αποτελεί το επαυξημένο περιβάλλον ήχου το οποίο μεταφέρεται στον ακροατή μέσω του ειδικού εξοπλισμού αναπαραγωγής. Η τελική ARA μείξη, δομείται από το συνδυασμό των τεχνολογιών αναπαραγωγής του πραγματικού περιβάλλοντος, της χωρικής τοποθέτησης του εικονικού περιβάλλοντος και των τεχνολογιών διάδρασης. Κατά συνέπεια, λόγω της πολυπλοκότητας και της πολύπλευρης επίδρασης των παραπάνω παραγόντων στην ARA μείξη, είναι επιτακτική ανάγκη να υπάρχει ένας έλεγχος του τελικού αποτελέσματος της μείξης, λαμβάνοντας υπόψιν τον ανθρώπινο παράγοντα και ειδικότερα, την ανθρώπινη ακουστική αντίληψη.

Η γενίκευση του μοντέλου της ARA μείξης, χρειάζεται μια στοχευμένη αντιμετώπιση, λαμβάνοντας υπόψιν τη μέχρι στιγμής στάθμη τεχνικής των τεχνολογιών που χρησιμοποιούνται μέσα σε αυτήν και των υφιστάμενων προβλημάτων που επηρεάζουν την ποιότητα της μείξης. Παραδείγματα τέτοιων προβλημάτων αποτελούν παραδοσιακά ηχητικά φαινόμενα όπως είναι εκείνο του ηχητικού μασκαρίσματος [87] και της συναχρόασης ηχητικών [10] σημάτων. Στόχος αυτού του κεφαλαίου είναι η κριτική αποτύπωση του παραδοσιακού μοντέλου ARA μείξης μέσα από την λειτουργία του και τους περιορισμούς που τίθενται.

### 3.2 Αρχιτεκτονική παραδοσιακού μοντέλου ARA μείξης

Το παραδοσιακό μοντέλο ARA μείξης, όπως παρουσιάζεται στο Σχήμα 3.1, υλοποιεί τη σύνδεση του εικονικού περιβάλλοντος με το ψευδοακουστικό. Πυρήνας της μείξης αποτελεί ο ARA μείκτης. Οι λειτουργίες που επιτελεί είναι η οδήγηση των σημάτων που προέρχονται από τα αμφιωτικά μικρόφωνα στα ακουστικά και η προσθήκη των εικονικών ηχητικών αντικειμένων στην αναπαραγωγή του ψευδοακουστικού περιβάλλοντος [46]. Επιπροσθέτως, δίνεται και η δυνατότητα αναμετάδοσης του πραγματικού περιβάλλοντος σε απομακρυσμένο ακροατή για περιπτώσεις εφαρμογών επικοινωνίας [24, 47].



Σχήμα 3.1: Γενικευμένο Μοντέλο ARA Μείξης (Τροποποιημένο σχήμα από το άρθρο του Aki Härmä σελ.625) [46]

Οι αρχικές προσεγγίσεις του επαυξημένου μείκτη ήχου [129, 46] ήταν η απλή μείξη του ψευδοακουστικού περιβάλλοντος με τις εικονικές ηχητικές πηγές, χωρίς την χρήση ισοσταθμιστικών ρυθμίσεων για το επαυξημένο περιβάλλον. Η απαίτηση για τη βέλτιστη ηχητική αναπαραγωγή του φυσικού περιβάλλοντος και η διαφάνεια του ακουστικού εξοπλισμού, δημιούργησε την ανάγκη για χρήση ισοστάθμισης, με σκοπό την εξάλειψη της επίδρασης του ζεύγους ακουστικών στη μεταβολή της ακουστικής του αυτιού. Ο πραγματικός κόσμος πρέπει να γίνεται αντιληπτός φυσικά, σαν να μην παρεμβάλλεται το ακουστικό ανάμεσα στο πραγματικό και εικονικό περιβάλλον. Οι πρώτες προσεγγίσεις για ισοστάθμιση, έκαναν χρήση αναλογικού εξοπλισμού, λόγω της χαμηλής καθυστέρησης που εισήγαγε μια τέτοιου είδους επεξεργασία στο σύστημα για την αποφυγή του φαινομένου του Comb filter Effect [128]. Η εξέλιξη των ψηφιακών επεξεργασιών ήχου ως προς την χαμηλή καθυστέρηση κατά την επεξεργασία σημάτων, επέτρεψε την εισαγωγή ψηφιακού ισοσταθμιστή στον ARA μείκτη [102].

Η χρήση ψηφιακής ισοστάθμισης επιτρέπει τη χρήση σε συστήματα αμφιωτικής τηλεφωνίας και τηλεσυνεδρίασης [102].

Για την αναπαραγωγή του εικονικού περιβάλλοντος, χρησιμοποιούνται οι συναρτήσεις τύπου HRTF και, μέσα από την διαδικασία της ακουστικοποίησης (Auralization), η εικονική πηγή ήχου λαμβάνει τα χωρικά χαρακτηριστικά που την τοποθετούν στον τρισδιάστατο εικονικό ηχητικό χώρο. Στόχος για το εικονικό περιβάλλον είναι η αναπαραγωγή του, διασφαλίζοντας την υψηλή στάθμη αυθεντικότητας και αντιληπτής ποιότητας [57].

### 3.3 Περιορισμοί παραδοσιακού μοντέλου ARA μείξης

Η λειτουργία του ARA μείκτη μπορεί να προσεγγιστεί με δύο τρόπους. Η πρώτη προσέγγιση στηρίζεται στο ότι κατά την επαυξημένη μείξη δύναται να μην είναι εφικτός ο διαχωρισμός του πραγματικού περιβάλλοντος από το εικονικό ηχητικό περιβάλλον [57]. Για να επιτευχθεί αυτό, χρειάζεται ένα ακουστικό σύστημα που να συνδυάζει εξαιρετικής πιστότητας αμφιωτική αναπαραγωγή, σύστημα εντοπισμού θέσης και κεφαλιού σε συνδυασμό με την υλοποίηση συστήματος πλοήγησης και υπολογισμού της απόστασης των εικονικών ηχητικών πηγών. Η δεύτερη προσέγγιση περιγράφει τα σενάρια στα οποία ο εικονικός ήχος αναπαράγεται με τέτοιο τρόπο, ώστε ο θεατής να μπορεί να τον διαχωρίσει από τον πραγματικό κόσμο [128, 60]. Η πρώτη προσέγγιση η οποία στηρίζεται στην προσαρμογή του εικονικού περιβάλλοντος στο πραγματικό, θεωρείται πως προσδίδει ρεαλιστικότητα στο επαυξημένο ηχητικό περιβάλλον [106]. Παρόλα αυτά δεν μπορεί να αποκλειστεί καμία από τις παραπάνω προσεγγίσεις καθώς υπάρχει και ο παράγοντας του ηχητικού σεναρίου στόχου ο οποίος καθορίζει τη βαρύτητα της σημαντικότητας των δύο περιβαλλόντων.

Παραδοσιακά, το μοντέλο ARA μείξης, ορίζει σταθερά κέρδη στα συνιστάμενα μέρη του (βλπ Εξ. 3.1) [61]).

$$ARA_{mix}(n) = R_S(n) + G * V_S(n) \quad (3.1)$$

όπου  $R_S(n)$  είναι το σήμα του πραγματικού περιβάλλοντος,  $V_S(n)$  το σήμα του εικονικού περιβάλλοντος και  $G$  ο σταθερός συντελεστής κέρδους για το εικονικό περιβάλλον.

Προφανώς, στην παραγόμενη μείξη υπάρχει ο αστάθμητος παράγοντας της δυναμικής μεταβολής του πραγματικού περιβάλλοντος. Σε ένα περιβάλλον για παράδειγμα εσωτερικού χώρου, μπορεί να υπάρξει έλεγχος της σχετικής ηχητικής συνιστώσας  $R_S(n)$ . Σε αντίθεση, για ένα περιβάλλον εξωτερικού χώρου, ο έλεγχος του πραγματικού περιβάλλοντος είναι πιο δύσκολος, με αποτέλεσμα να δημιουργείται κίνδυνος επικάλυψης του εικονικού στοιχείου σε τέτοιο βαθμό, που να δημιουργείται πρόβλημα στην κατανόησή του.

Σε ένα ενδεικτικό παράδειγμα εφαρμογής του παραδοσιακού μοντέλου ARA μείξης σε σενάριο ηχητικού παιχνιδιού επαυξημένης πραγματικότητας ήχου ανοικτού πεδίου [111], διαπιστώθηκαν τα ακόλουθα: Η χρήση του παραδοσιακού μοντέλου για την πρώτη φάση του παιχνιδιού, εκείνη της εξερεύνησης και εύρεσης του σημείου δράσης, οι παίκτες δεν αντιμετώπισαν σοβαρό πρόβλημα, παρ' όλους του εξωτερικούς παράγοντες, όπως φυσικός θόρυβος λόγω αέρα ή γάβγισμα σκυλιών. Στην ακόλουθη φάση της δράσης, στην οποία οι παίκτες

βρίσκονται σε στενό εύρος του πεδίου διάδρασης κάνουν χρήση του προκαθορισμένου τρόπου διάδρασης για την ολοκλήρωση του σεναρίου, παρατηρείται ότι οι εξωτερικοί παράγοντες που περιλαμβάνονται στο πραγματικό περιβάλλον, επηρεάζουν την ακουστική αντίληψη του παίκτη και κατά συνέπεια και την απόδοσή του στο παιχνίδι [93]. Στην εργασία του Rämö [47], παρουσιάζεται το πρόβλημα όπου το πραγματικό περιβάλλον είναι υψηλής στάθμης με αποτέλεσμα να υπάρχει πιθανότητα πρόκλησης προβλήματος στην ακοή ή ελάττωση της ακουστικής ποιότητας του περιβάλλοντος. Σε αυτό το πρόβλημα η τεχνολογία ARA με την υλοποίηση κατάλληλου ισοσταθμιστή σε πραγματικό χρόνο, φαίνεται να υπάρχει δυνατότητα βελτίωσης της ηχητικής ποιότητας εξασθενίζοντας ή τονίζοντας επιλεγμένες κατάλληλα συχνοτικές ζώνες του επαυξημένου ηχητικού σήματος. Στο παραπάνω παράδειγμα, η επιλογές ισοστάθμισης είναι προκαθορισμένες χωρίς προσαρμογή στο εκάστοτε πραγματικό περιβάλλον.

Οι παραπάνω προσεγγίσεις, δημιουργούν την ανάγκη ελέγχου της ARA μείξης, η οποία έχει να ξεπεράσει τα παρακάτω προβλήματα. Η υπόσταση του πραγματικού περιβάλλοντος είναι σύνθετη και δυναμικά μεταβαλλόμενη. Για παράδειγμα, το ηχητικό περιβάλλον ενός δρόμου σε ώρα αιχμής, περιλαμβάνει συνεχείς μεταβολές. Το εικονικό περιβάλλον με βάση την πρώτη προσέγγιση, θα πρέπει να επανακαθορίζει τις ηχητικές του στάθμες έχοντας ως κριτήριο την ακουστική αντίληψη του ακροατή ως προς το εικονικό και πραγματικό περιβάλλον. Μια χειροκίνητη τέτοια διαδικασία, είναι πολύπλοκη και αδύνατη για τον ακροατή, καθώς δεν γνωρίζει το ζητούμενο του ηχητικού σεναρίου, πράγμα το οποίο γνωρίζει μόνο ο σχεδιαστής. Η λύση που προτείνεται με την παρούσα διατριβή είναι η δυναμική ARA μείξη μέσα από ένα προσαρμοστικό μοντέλο το οποίο να καθορίζεται από ένα γενικευμένο μοντέλο ακουστικής αντίληψης [96]. Ο έλεγχος της ARA μείξης με τη δυναμική προσαρμογή των ηχητικών περιβαλλόντων, αποτελεί το συνδυαστικό κρίκο όλων των τεχνολογιών που χρησιμοποιούνται στο πεδίο της επαυξημένης πραγματικότητας ήχου. Η εφαρμογή ενός τέτοιου μοντέλου, αρχιτεκτονικά τοποθετείται πριν την εισαγωγή των χωρικών χαρακτηριστικών στο εικονικό περιβάλλον, μηδενίζοντας την πιθανότητα αλλοίωσης των χωρικών χαρακτηριστικών που δίνονται κατά την διαδικασία χωρικής τοποθέτησης των εικονικών ηχητικών πηγών.

Η παραπάνω προτεινόμενη μέθοδος δημιουργεί κάποιους πρωταρχικούς περιορισμούς στο θέμα του εξοπλισμού. Η χρήση ακουστικών δονήσεων οστού (Bone-Conduction) δυσκολεύει τον ακουστικό έλεγχο του πραγματικού περιβάλλοντος, καθώς σε περιπτώσεις που το εικονικό περιβάλλον μεταφέρει κρίσιμη πληροφορία, είναι αδύνατον να ελεγχθεί το πραγματικό ηχητικό περιβάλλον, λόγω του ανοικτού ακουστικού καναλιού. Οπότε σε περιπτώσεις ενός θορυβώδους πραγματικού περιβάλλοντος, δεν μπορεί να καθορισθεί το κέρδος του.

Οι παραπάνω περιορισμοί δημιουργούν την ανάγκη για εξέλιξη της επαυξημένης μείξης με την εισαγωγή τεχνικών προσαρμοστικότητας για τον έλεγχο της ισορροπίας μεταξύ των συνιστωσών του επαυξημένου περιβάλλοντος και την ενίσχυση της εμπύθισης του ακροατή με άμεση επίδραση στην διάδρασή του με το σύστημα:

$$AdptARA_{mix}(n) = R_S(n) + G * V_S(n), \quad G = \frac{R_L}{V_L} \quad (3.2)$$

Στην Εξίσωση 3.2 παρουσιάζεται η θεωρητική προσέγγιση της προσαρμοστικής μείξης

σύμφωνα με τα εξαγόμενα ακουστικά χαρακτηριστικά του πραγματικού και του εικονικού περιβάλλοντος  $R_L$  και  $V_L$ . Η σχέση των ακουστικών χαρακτηριστικών που εξορύσσονται από τα αντίστοιχα ηχητικά περιβάλλοντα, καθορίζουν τον συντελεστή μείξης του εικονικού περιβάλλοντος ως προς το πραγματικό (συντελεστής  $G$ ). Η παραπάνω ενέργεια αποτελεί την βασική διαφορά μεταξύ του παραδοσιακού μοντέλου μείξης ήχου επαυξημένης πραγματικότητας και του προτεινόμενου προσαρμοστικού.

### 3.4 Σύνοψη Κεφαλαίου

Ο τεχνολογικός - αλγοριθμικός πυρήνας στα επαυξημένα περιβάλλοντα ήχου, είναι η ARA μείξη. Η ανάμειξη των ηχητικών περιβαλλόντων, υλοποιείται μέσα στα πλαίσια που ορίζονται από τον ARA μείκτη. Μέχρι στιγμής, η λειτουργία του παραδοσιακού μοντέλου, συνοψίζεται στην οδήγηση των ηχητικών σημάτων και την τοποθέτηση των εικονικών ηχητικών αντικειμένων στο ψευδοακουστικό περιβάλλον. Η τεχνολογική εξέλιξη και η πολυπλοκότητα των ηχητικών σεναρίων, συμβάλει στην απαίτηση της αναβάθμισης του ρόλου του ARA μείκτη. Η δυναμική προσαρμογή των συνισταμένων περιβαλλόντων της επαυξημένης μείξης ήχου λαμβάνοντας υπόψιν την ακουστική αντίληψη, είναι το νέο ζητούμενο το οποίο στην παρούσα διατριβή εξετάζεται μέσα από το προτεινόμενο πρωτότυπο μοντέλο ARA μείξης με όνομα AR-ION - Augmented Reality of IONian University που παρουσιάζεται στα επόμενα κεφάλαια.



## Κεφάλαιο 4

# Μεθοδολογία Ενσωμάτωσης Προσαρμοστικότητας στην ARA Μείξη

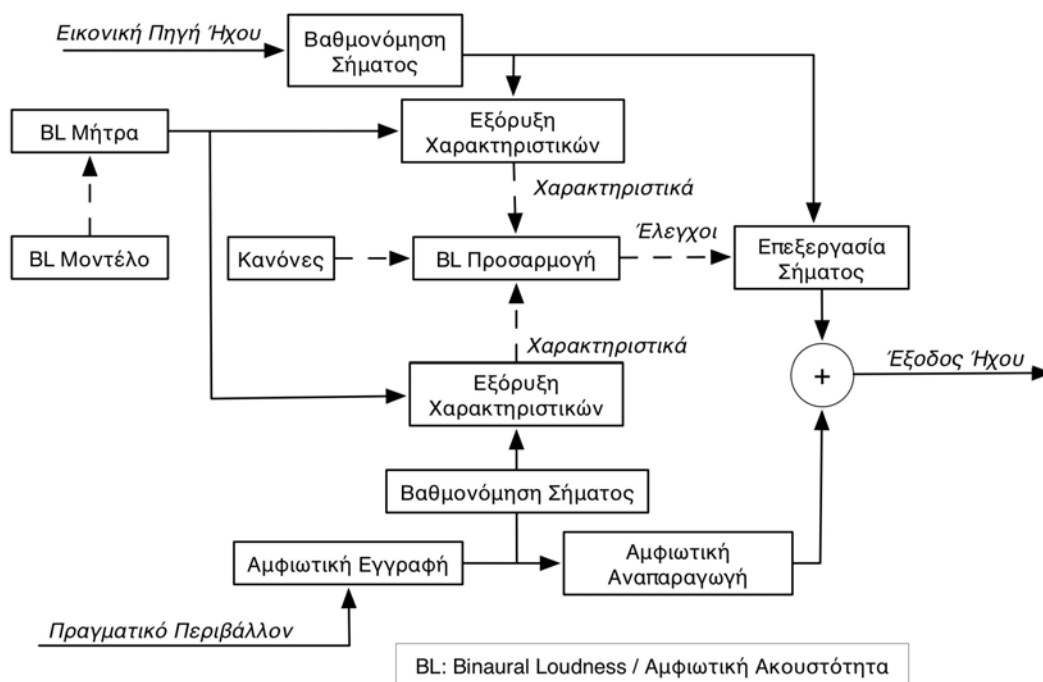
### 4.1 Εισαγωγή

Η ιδιότητα της προσαρμοστικότητας αποτελεί τη βάση για την αναβάθμιση της ακουστικής απόδοσης της ARA μείξης. Στο παρόν κεφάλαιο, παρουσιάζεται η μεθοδολογία με την οποία υλοποιείται η εισαγωγή της προσαρμοστικότητας στο μείκτη επαυξημένης πραγματικότητας ήχου. Η ρύθμιση της ακουστικής ισορροπίας μεταξύ των ηχητικών περιβαλλόντων με όρους ακουστικής αντίληψης συνθέτουν το νέο περιβάλλον επαυξημένης πραγματικότητας ήχου. Η μέθοδος εκτίμησης της αμφιωτικής ακουστότητας, αποτελεί τον κρίσιμο παράγοντα για τον καθορισμό της προσαρμοστικής μείξης. Η ομαλή λειτουργία του προτεινόμενου μοντέλου και η μεθοδολογία ενσωμάτωσής του μέσα στο περιβάλλον ARA δίνεται σε αυτό το κεφάλαιο.

Το παρόν κεφάλαιο ακολουθεί την παρακάτω δομή. Αρχικά εισάγεται η αρχιτεκτονική του γενικευμένου μοντέλου προσαρμοστικής ARA μείξης. Έχοντας σαν βάση το προτεινόμενο μοντέλο, γίνεται μια ανάλυση στα κρίσιμα σημεία του. Κατόπιν παρουσιάζεται η μεθοδολογία που ακολουθείται για την λήψη των σημάτων εισόδου. Στη συνέχεια, γίνεται ανάλυση της μεθόδου εξόρυξης δεδομένων ακουστικής αντίληψης από τα σήματα εισόδου των περιβαλλόντων της ARA μείξης. Κατόπιν παρουσιάζονται οι διεργασίες υπολογισμού της ενέργειας διέγερσης του αισθητήριου της ακοής για τα σήματα εισόδου και η εκτίμηση της αμφιωτικής ακουστότητας, σύμφωνα με την οποία πραγματοποιείται η δυναμικά μεταβαλλόμενη επαυξημένη μείξη. Ακολούθως, δίνεται η μεθοδολογία ανάλυσης και επανασύνθεσης του εικονικού περιβάλλοντος σε πραγματικό χρόνο, η οποία είναι απαραίτητη για την εφαρμογή των συντελεστών προσαρμογής. Τέλος, αναλύεται ο αλγόριθμος προσαρμογής μεταξύ των περιβαλλόντων επαυξημένης πραγματικότητας, καθώς και η προαιρετική σταθερά διαβάθμισης της προσαρμοστικής ARA μείξης.

## 4.2 Γενικευμένο μοντέλο προσαρμοστικής ARA μείξης

Η μέθοδος αναβάθμισης της λειτουργίας της ARA μείξης στηρίζεται στην ενσωμάτωση ενός μοντέλου εξόρυξης αντιληπτικών χαρακτηριστικών ακουστότητας και στην χρήση τους στην διαδικασία της προσαρμογής για τον καθορισμό της σχέσης πραγματικού με εικονικό περιβάλλον. Λόγω της αυξημένης μεταβλητότητας του πραγματικού περιβάλλοντος, η επιλογή της προσαρμοστικότητας αποτελεί την λύση για την άμεση ανάλυση και επεξεργασία του επαυξημένου περιβάλλοντος και τον καθορισμό της ισορροπίας κατά την ARA μείξη. Η εισαγωγή μοντέλου εκτίμησης της αμφιωτικής ακουστότητας (Binaural Loudness - BL) των ηχητικών συνιστωσών του επαυξημένου περιβάλλοντος, στοχεύει στην εξάλειψη του φαινομένου του μασκαρίσματος μεταξύ των ηχητικών περιβαλλόντων δίνοντας έμφαση στην ενίσχυση της κατανόησης των περιβαλλόντων, τη διαφάνεια της μείξης [68] και την βελτίωση της αλληλεπίδρασης του ακροατή με το σύστημα, με βελτιωμένα αποτελέσματα σε συνθήκες λήψης αποφάσεων σε ηχητικά σενάρια ειδικά αυτά όπου το πραγματικό περιβάλλον έχει σημαντικό ρόλο [93].



Σχήμα 4.1: Γενικευμένο Μοντέλο Προσαρμοστικής ARA Μείξης

Στο Σχήμα 4.1 αποτυπώνεται η αρχιτεκτονική του προτεινόμενου γενικευμένου μοντέλου ARA μείξης βασισμένο σε μοντέλο εκτίμησης αμφιωτικής ακουστότητας. Η αρχιτεκτονική του προτεινόμενου μοντέλου αποτελεί ένα πολυεπίπεδο σύστημα διεργασιών, με σκοπό την προσαρμογή του εικονικού περιβάλλοντος στο πραγματικό για την επιτέλεση της τελικής μείξης. Παρακάτω επισημαίνονται τα στάδια διεργασιών που επιτελούνται μέσα από το μοντέλο.

Πρώτο στάδιο: Εισαγωγή ηχητικών περιβαλλόντων στο σύστημα. Για το πραγματικό

ηχητικό περιβάλλον, πραγματοποιείται αμφιωτική εγγραφή με έμφαση στην διαφάνεια του ακουστικού που χρησιμοποιείται. Η διαφάνεια ορίζεται από την ισοστάθμιση της επίδρασης του ακουστικού εξοπλισμού λόγω της έμφραξης του ακουστικού πόρου, που έχει συνέπεια τη μεταβολή της συχνотικής απόκρισης του ακουστικού οργάνου κατά την εισαγωγή του πραγματικού περιβάλλοντος. Για το εικονικό περιβάλλον, γίνεται εισαγωγή του ανηχοϊκού σήματος της εικονικής πηγής ήχου στο σύστημα.

Δεύτερο στάδιο: Βαθμονόμηση σημάτων του εκάστοτε ηχητικού περιβάλλοντος πριν την εισαγωγή τους στο σύστημα εξόρυξης χαρακτηριστικών αμφιωτικής ακουστότητας και τελικής επεξεργασίας.

Τρίτο στάδιο: Εξόρυξη χαρακτηριστικών αμφιωτικής ακουστότητας η οποία περιλαμβάνει τον υπολογισμό της ακουστικής διέγερσης ηχητικών σημάτων εισόδου και την εκτίμηση της αμφιωτικής ακουστότητας από την ακουστική διέγερση του κάθε σήματος σύμφωνα με την προτεινόμενη μέθοδο προϋπολογισμένης μήτρας αμφιωτικής ακουστότητας (BL Μήτρα) σύμφωνα με το προεπιλεγμένο μοντέλο εκτίμησης αμφιωτικής ακουστότητας (BL Μοντέλο). Αυτή η διεργασία υλοποιείται για την εικονική και πραγματική ηχητική συνιστώσα του επαυξημένου περιβάλλοντος.

Τέταρτο στάδιο: Σε αυτό το επίπεδο γίνεται ο υπολογισμός των συντελεστών προσαρμογής του εικονικού περιβάλλοντος από τον αλγόριθμο προσαρμογής σύμφωνα με τις εκτιμήσεις αμφιωτικής ακουστότητας του προηγούμενου επιπέδου και του κανόνες προσαρμογής.

Πέμπτο στάδιο: Ενσωμάτωση των συντελεστών προσαρμογής που εξήχθησαν κατά την διεργασία του τέταρτου σταδίου στο σήμα της εικονικής πηγής ήχου.

Έκτο στάδιο: Τελική μείξη πραγματικού ηχητικού περιβάλλοντος με το προσαρμοσμένο εικονικό περιβάλλον.

Σχετικά με την σειρά εισαγωγής του μοντέλου στην αλυσίδα της ARA μείξης τίθεται ένα εύλογο ζήτημα. Η τοποθέτηση του να γίνει πριν την χωρική τοποθέτηση του εικονικού περιβάλλοντος ή μετά από την εισαγωγή των χωρικών χαρακτηριστικών στην εικονική ηχητική πηγή. Λόγω της ευαισθησίας των χωρικών χαρακτηριστικών από μια αλλοίωση στο συχνотικό φάσμα το οποίο μπορεί να προκαλέσει η επεξεργασία του εικονικού σήματος από το μοντέλο προσαρμοστικής μείξης, η επιλογή συγχλίνει στην τοποθέτηση της διεργασίας προσαρμογής πριν την χωρική επεξεργασία της εικονικής πηγής.

Το βάρος της ισορροπίας μεταξύ του εικονικού και του πραγματικού περιβάλλοντος στην προσαρμοστική μείξη δημιουργεί δύο προσεγγίσεις. Την προσαρμογή του εικονικού περιβάλλοντος στο πραγματικό ή το αντίθετο. Στο πλαίσιο της διατριβής, επιλέχθηκε η προσαρμογή του εικονικού περιβάλλοντος στο πραγματικό για να διασφαλιστεί η φυσική ενσωμάτωση του εικονικού στο πραγματικό, μέσω της οποίας προσδίδεται αίσθηση εμπύθισης σύμφωνα και με προηγούμενες έρευνες [107]. Η διατήρηση της φυσικής αναπαραγωγής του πραγματικού περιβάλλοντος μέσα από ένα διάφανο ισοσταθμισμένο σύστημα ακουστικών [81] αποτελεί παράγοντα που σε μεγάλο βαθμό καθορίζει την επιτυχία της προσαρμογής της ARA μείξης. Τα παραπάνω επίπεδα θα αναλυθούν στα επόμενα κεφάλαια μέσα από την μεθοδολογική προσέγγιση των διεργασιών του μοντέλου προσαρμοστικής ARA μείξης.

### 4.3 Τεχνική Αμφιωτικής Επεξεργασίας Σημάτων Εισόδου ARA μοντέλου

Τα σήματα εισόδου στο μοντέλο προσαρμοστικής ARA μείξης διαιρούνται σε δύο κατηγορίες. Το αμφιωτικό σήμα του πραγματικού περιβάλλοντος το οποίο καταλαμβάνει δύο κανάλια εισόδου και το εικονικό περιβάλλον το οποίο αποτελείται από ανηχοϊκές ηχητικές πηγές οι οποίες καταλαμβάνουν ένα κανάλι η κάθε μία. Η εικονική πηγή παράγεται ως ανηχοϊκό σήμα, το οποίο αρχικά πρέπει να προσαρμοστεί στην ARA μείξη με το πραγματικό περιβάλλον και, ακολούθως, να της δοθούν τα χωρικά χαρακτηριστικά τοποθέτησής της στον τρισδιάστατο ηχητικό χώρο. Παρακάτω θα αναλυθεί η τεχνική αμφιωτικής επεξεργασίας του πραγματικού περιβάλλοντος που χρησιμοποιείται από το προσαρμοστικό μοντέλο ARA μείξης.

Μεταξύ των τεχνικών αμφιωτικής επεξεργασίας πραγματικών περιβαλλόντων που χρησιμοποιούνται σε περιβάλλοντα τύπου ARA [81], επιλέχθηκε η μέθοδος χρήσης μικροφώνου στην είσοδο του ακουστικού πόρου του κάθε αυτιού για την αμφιωτική εγγραφή του πραγματικού περιβάλλοντος. Για την αναπαραγωγή του ληφθέντος ήχου, χρησιμοποιείται ακουστικό τύπου In-Ear το οποίο τοποθετείται στον ακουστικό πόρο του κάθε αυτιού μπλοκάροντας αυτόν. Η χρήση της τεχνικής αυτής έχει σημαντικά προτερήματα και μειονεκτήματα τα οποία όμως σύμφωνα και με την βιβλιογραφία μπορούν να αντιμετωπισθούν [102]. Πιο αναλυτικά τα μειονεκτήματα συνοψίζονται στη μεταβολή της ακουστικής του ακουστικού πόρου λόγω της έμφραξης του από τον εξοπλισμό ο οποίος τοποθετείται στην είσοδο του ακουστικού πόρου. Η επανασύσταση της ακουστικής του ανοικτού ακουστικού πόρου, είναι εφικτή με την χρήση κατάλληλων ισοσταθμιστικών φίλτρων. Σε αυτή την περίπτωση δεν χρειάζεται επανασύσταση της ακουστικής επίδρασης του έξω αυτιού το οποίο περιέχει την επίδραση του ήχου από το σώμα και το κεφάλι του ακροατή καθώς προσπίπτει ο ήχος πάνω σε αυτό και οδηγείται μέσω του εξωτερικού αυτιού στον ακουστικό πόρο. Η διαδικασία επανασύστασης των παραπάνω αναχλάσεων αποτελεί πολυπλοκότερη διαδικασία καθώς θα πρέπει να ληφθούν υπόψιν ανθρωπομετρικοί παράμετροι οι οποίοι θα έπρεπε να υπολογιστούν για κάθε ακροατή ξεχωριστά.

Συνεπώς, βασικός λόγος χρήσης In-Ear ακουστικών με εξωτερικό μικρόφωνο σε περιβάλλοντα ARA, είναι η διατήρηση της χωρικής πληροφορίας του ήχου που λαμβάνεται από το εξωτερικό μικρόφωνο του κάθε ακουστικού, και προέρχεται από την επίδραση του κεφαλιού και του σώματος η οποία συγκεντρώνεται στο εξωτερικό αυτί του ακροατή [18]. Στην περίπτωση ακουστικών non In-Ear όπως είναι τα Closed-Back, θα χρειαζόταν ειδική τεχνική μέτρησης της χροστική απόκρισης του έξω αυτιού με την πιθανή εισαγωγή εσωτερικού μικροφώνου, και κατόπιν δημιουργίας του κατάλληλου φίλτρου επανασύστασης της ακουστικής του έξω αυτιού και επιπλέον και της ακουστικής του ακουστικού πόρου. Στην περίπτωση χρήσης In-Ear ακουστικών, λόγω μόνο της έμφραξης του ακουστικού πόρου, χρειάζονται να γίνει ισοσταθμιστική παρέμβαση μόνο για την επανασύσταση της ακουστικής του ακουστικού πόρου για τη διάφανη αναπαραγωγή του πραγματικού περιβάλλοντος.

Είναι κρίσιμο η τεχνική αμφιωτικής αναπαραγωγής να προσδίδει διαφάνεια στο ακουστικό και να διατηρεί τα χωρικά χαρακτηριστικά του πραγματικού περιβάλλοντος. Παρακάτω παρου-

σιάζονται τα τεχνικά ζητήματα που προκύπτουν από την χρήση της προτεινόμενης τεχνικής για το μοντέλο προσαρμοστικής ARA μείξης.

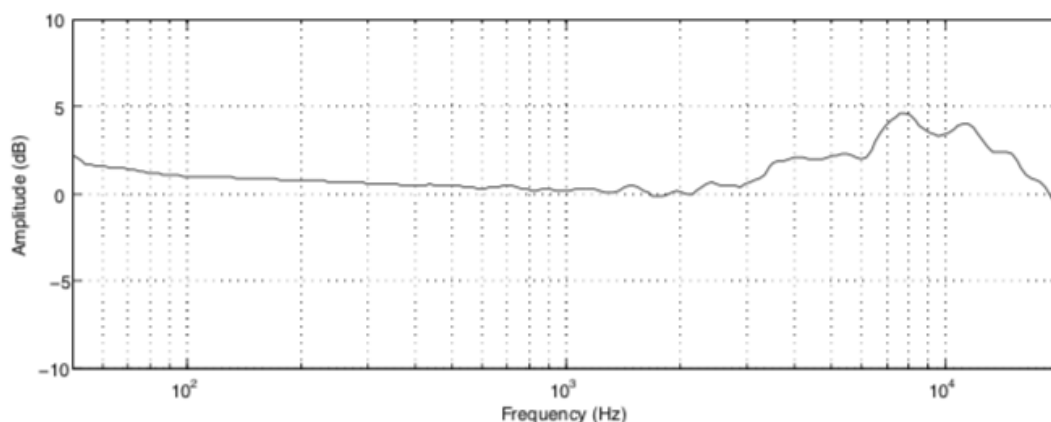
Η θέση τοποθέτησης των ακουστικών για την αμφιωτική καταγραφή του πραγματικού περιβάλλοντος θα πρέπει να είναι όσο το δυνατόν πιο κοντά στην είσοδο του ακουστικού πόρου του κάθε αυτιού. Με αυτό τον τρόπο, είναι δυνατόν να καταγραφεί το σύνολο της ακουστικής επίδρασης του έξω αυτιού, το οποίο προσδίδει τα χωρικά χαρακτηριστικά του πραγματικού ηχητικού περιβάλλοντος. Οι επιδράσεις προέρχονται από ανακλάσεις του σώματος του ακροατή, των ώμων, του κεφαλιού και του έξω αυτιού. Η χρήση ακουστικών εξωτερικού αυτιού σύμφωνα με τα παραπάνω, θα εισήγαγε επιπλέον πολυπλοκότητα στην ανάκτηση της συγκεκριμένης πληροφορίας [106]. Η πολυπλοκότητα έγκειται στην ανάγκη χρήσης δεδομένων ανθρωπομετρίας τα οποία είναι δύσκολο να ληφθούν κατά τη διάρκεια ενός σεναρίου τύπου ARA. Η ευχρηστία του εξοπλισμού, είναι σημαντικός παράγοντας για την επιλογή του. Η χρήση ακουστικών που In-Ear για τη μέγιστη διαφάνεια κατά την αναπαραγωγή του ψευδο-ακουστικού περιβάλλοντος, θα χρειαζόταν επιπλέον εξοπλισμό μέτρησης της επίδρασης του έξω αυτιού το οποίο προσθέτει μια επιπλέον δυσκολία στη χρήση του εξοπλισμού. Η χρήση ακουστικών τεχνολογίας in-Ear αποτελεί αυτή τη στιγμή την ιδανική επιλογή για το μοντέλο προσαρμοστικής ARA μείξης.

Το μπλοκάρισμα του ακουστικού πόρου, δημιουργεί μεταβολές στην ακουστική του. Ο ανοικτός ακουστικός πόρος λειτουργεί ως ανοικτός ακουστικός σωλήνας με πρώτη αντήχηση να συμβαίνει στο  $\lambda/4$  του μήκους κύματος μεταξύ των 2 kHz - 4 kHz. Η χρήση In-Ear ακουστικού, μεταβάλλει την ακουστική του αυτιού σε κλειστό ακουστικό σωλήνα με πρώτη αντήχηση σε  $\lambda/2$  του μήκους κύματος μεταξύ των 5 kHz - 10 kHz ακυρώνοντας την αρχική ακουστική αντήχηση του ανοικτού ακουστικού πόρου. Η αντιμετώπιση του παραπάνω φαινομένου είναι καθοριστική για την φυσική αμφιωτική αναπαραγωγή του πραγματικού περιβάλλοντος. Συνεπώς, θα πρέπει να ακυρωθεί η πρώτη αντήχηση σε  $\lambda/2$  του μήκους κύματος με την εφαρμογή ενός φίλτρου κορυφής με κεντρική συχνότητα μεταξύ των 700Hz-3200Hz. Επιπροσθέτως, θα πρέπει να επαναδημιουργηθεί η αρχική αντήχηση του ανοικτού ακουστικού πόρου με την εισαγωγή ενός φίλτρου εγχοπής με κεντρική συχνότητα μεταξύ των 1800Hz-8500Hz [102].

Το επόμενο πρόβλημα που προκαλείται από την εφαρμογή των ακουστικών τύπου In-Ear είναι η συγκέντρωση ακουστικής ενέργειας στον κλεισμένο ακουστικό πόρο. Αυτή η συγκέντρωση παρατηρείται κυρίως σε μεσαίες και χαμηλές συχνότητες. Για να αποτραπεί αυτή η συγκέντρωση, μπορεί να χρησιμοποιηθεί ένα υψηλοπερατό (High-Pass) φίλτρο με συχνότητα αποκοπής από τα 6Hz-720Hz [102].

Επίσης πρέπει να προστεθεί, πως η απόκριση των ακουστικών τύπου In-Ear, παρουσιάζει έλλειμμα στην απόδοση των χαμηλών συχνοτήτων [98]. Στο Σχήμα 4.2, παρουσιάζεται η απόκριση της συχνότητας των ακουστικών τύπου In-Ear. Τα ακουστικά του συγκεκριμένου παραδείγματος, χρησιμοποιήθηκαν στο στάδιο της υλοποίησης της παρούσας διατριβής. Σύμφωνα με το Σχήμα 4.2, παρουσιάζεται ελάττωση στην απόκριση στις χαμηλές συχνότητες. Για την αντιστάθμιση της παραπάνω ιδιαιτερότητας, είναι σύνηθες να χρησιμοποιούνται τεχνικές επεξεργασίας με εισαγωγή αρμονικών («έξυπνα ακουστικά») [144], όπου στην παρούσα περίπτωση, θα ήταν αιτία ακύρωσης του High-Pass φίλτρου. Λόγω του φαινομένου συγκέντρω-

σης ενέργειας του κλειστού ακουστικού πόρου, δεν απαιτείται χρήση τεχνικών διόρθωσης της συχνοτικής απόκρισης των χαμηλών συχνοτήτων, αντιθέτως, χρειάζεται ελάττωση της υφιστάμενης ενέργειας των χαμηλών συχνοτήτων με την εισαγωγή του High-Pass φίλτρου.



Σχήμα 4.2: Συχνοτική απόκριση ακουστικών τύπου In-Ear Philips SHN 2500 (Σχήμα από την εργασία του Jussi Rämö, σελ.43 [58])

Η εφαρμογή του ακουστικού στον ακουστικό πόρο, δεν είναι συνήθως τέλεια με αποτέλεσμα να υπάρχει διαρροή του εξωτερικού περιβάλλοντος προς τον ακουστικό πόρο και το αντίθετο. Συνήθως αυτή η διαρροή συμβαίνει σε συχνότητες μεσαίες και χαμηλές, κάτω από το 1 kHz [102]. Παράλληλα, υπάρχει και η επίδραση της απομόνωσης του ακροατή από το εξωτερικό περιβάλλον, λόγω της εφαρμογής του ακουστικού. Κατά τη διάρκεια αυτού του φαινομένου, εισάγονται προφανώς φασικές διαφορές οι οποίες σε σφιχτή τοποθέτηση των ακουστικών, εισάγουν καθυστέρηση περίπου 0.21 ms, ενώ, σε πιο χαλαρή τοποθέτηση, η καθυστέρηση μειώνεται σε 0.06ms [102]. Αυτή η καθυστέρηση επιτρέπει τη μη αντιληπτή ακουστική καθυστέρηση που εισάγει η χρήση ψηφιακών φίλτρων χωρίς να επηρεάζεται η απόδοση του ARA μείκτη.

Σύμφωνα με τα παραπάνω, η τεχνική αμφιωτικής αναπαραγωγής του πραγματικού περιβάλλοντος που χρησιμοποιείται στον προσαρμοστικό ARA μείκτη, βασίζεται στη χρήση ακουστικών In-Ear με την χρήση ακολουθίας φίλτρων για την εξάλειψη των προαναφερόμενων επιδράσεων της χρήσης αυτού του τύπου ακουστικών. Προφανώς, ο κάθε άνθρωπος έχει διαφορές στη διάπλαση του οργάνου του αυτιού, ωστόσο, οι παραπάνω επιδράσεις, συναντώνται στις περισσότερες περιπτώσεις.

#### 4.4 Εισαγωγή δεδομένων ακουστικής αντίληψης στην ARA μείξη

Η ARA μείξη, εκτός από την οδήγηση των ηχητικών περιβαλλόντων στον ακροατή και την ισοστάθμιση της αμφιωτικής αναπαραγωγής του πραγματικού περιβάλλοντος, εισάγει τα εικονικά ηχητικά αντικείμενα στην επαυξημένη μείξη. Ο καθορισμός του τρόπου προσαρμο-

γής των εικονικών πηγών στο επαυξημένο περιβάλλον, δεν είναι απόλυτα καθορισμένος. Η παρούσα διατριβή, υποστηρίζει, πως πρέπει να διέπεται από κανόνες που ορίζει η ακουστική αντίληψη των επαυξημένων περιβαλλόντων ήχου.

Η μέθοδος εκτίμησης της αμφιωτικής ακουστότητας αποτελεί αυτόνομο τμήμα του μοντέλου. Κεντρικό ρόλο αποτελεί η προϋπολογισμένη μήτρα τιμών αμφιωτικής ακουστότητας η οποία έχει δομηθεί με συγκεκριμένη μεθοδολογία η οποία θα περιγραφεί στην επόμενη ενότητα. Η εξόρυξη του χαρακτηριστικού της αμφιωτικής ακουστότητας σε συνδυασμό με ένα σύμπλεγμα κανόνων, διέπουν την μέθοδο για την υλοποίηση της προσαρμοστικής μείζης για περιβάλλοντα τύπου ARA.

Σε αυτό το κεφάλαιο θα παρουσιαστεί μια επισκόπηση των μοντέλων πρόβλεψης της αμφιωτικής ακουστότητας σε πραγματικό χρόνο και ακολούθως, η προτεινόμενη μέθοδος εκτίμησης της αμφιωτικής ακουστότητας που υλοποιείται. Επιπλέον, περιγράφεται η μέθοδος υλοποίησης των κανόνων που εισάγονται στον αλγόριθμο προσαρμογής του μοντέλου της ARA μείζης σύμφωνα με τους οποίους γίνονται οι απαραίτητες ρυθμίσεις στο πεδίο της συχνότητας για την τελική ενσωμάτωση της εικονικής συνιστώσας στην επαυξημένη μείζη ήχου.

#### 4.4.1 Υφιστάμενα μοντέλα πρόβλεψης Αμφιωτικής Ακουστότητας

Τα μοντέλα πρόβλεψης αμφιωτικής ακουστότητας έχουν στόχο την ποσοτικοποίηση της ακουστικής αντίληψης του ηχητικού στο πλαίσιο της αμφιωτικής ακοής. Παραδείγματα τέτοιων μοντέλων αποτελεί το μοντέλο του Zwicker - Fastl [148, 150] και του Glasberg - Moore [89]. Τα μοντέλα αυτά προβλέπουν την ανθρώπινη αντίληψη μιας ακουστικής διέγερσης που θα προκληθεί στο αισθητήριο όργανο της ακοής και την ποσοτικοποιούν, χρησιμοποιώντας κλίμακες ακουστότητας τύπου Phon ή Sone. Παρακάτω θα παρουσιαστούν τα κυρίαρχα μοντέλα πρόβλεψης της αμφιωτικής ακουστότητας μέσα από τον τρόπο λειτουργίας τους.

Οι κλίμακες ακουστότητας αναπτύχθηκαν αρχικά από τον S.S. Stevens [122, 123] και τον Barkhausen [32]. Η μεθοδολογία που ακολουθήθηκε για την δημιουργία των κλιμάκων ήταν δύο. Η πρώτη μέθοδος, ονομαζόταν υπολογισμός πλάτους, στην οποία το υποκείμενο έπρεπε να αντιστοιχήσει ήχους με διαφορετικές ηχητικές στάθμες σε αριθμούς σύμφωνα με την ληφθείσα ακουστότητα. Η δεύτερη μέθοδος, ονομάστηκε παραγωγή ηχητικού πλάτους, στην οποία το υποκείμενο ρυθμίζει τη στάθμη του ήχου σε απόλυτες ή σχετικές τιμές, σε σύγκριση με τον αρχικό ήχο, για παράδειγμα διπλασιασμός της ακουστότητας ή την μισή ακουστότητα.

Ο Stevens, πρότεινε ότι η αντιληπτή ακουστότητα  $L$ , ήταν συνάρτηση της δύναμης της φυσικής έντασης  $I$ :

$$L = kI^{0.6} \quad (4.1)$$

Όπου  $k$  είναι μια σταθερά που εξαρτάται από το υποκείμενο και την μονάδα μέτρησης που χρησιμοποιείται.

Η Εξίσωση 4.1, μπορεί να εκφραστεί σε ακουστική πίεση, όπου η ακουστότητα ενός ήχου είναι ανάλογη της RMS πίεσης σε δύναμη 0.6.

Η μονάδα μέτρησης ακουστότητας Sone, προτάθηκε από τον Stevens. Ως ένα Sone ορίζεται η ακουστότητα των 1000 Hz στην στάθμη των 40 dB SPL σε αμφιωτική αναπαραγωγή από εμπρόσθια κατεύθυνση σε ελεύθερο πεδίο. Μια αύξηση κατά 10 dB SPL υπολογίζεται σε αύξηση 2 Sones. Αυτή η σχέση όμως δεν ισχύει για στάθμες κάτω από 40 dB SPL, όπου η ακουστότητα με την ακουστική στάθμη αλλάζουν γρήγορα.

Η δεύτερη μέθοδος υπολογισμού της ακουστότητας, εισήχθη από τον Barkhausen ο οποίος δημιούργησε τη μονάδα Phon για την χρήση της από την κλίμακα κρίσιμων ζωνών συχνότητας που δημιούργησε, με το όνομα Bark [119].

| Αριθμός<br>Μπάντας | Κεντρικές<br>Συχνότητες<br>Hz | Κατώτατο<br>Όριο<br>Hz | Ανώτατο<br>Όριο<br>Hz | Εύρος |
|--------------------|-------------------------------|------------------------|-----------------------|-------|
| 1                  | 50                            | 0                      | 100                   | 100   |
| 2                  | 150                           | 100                    | 200                   | 100   |
| 3                  | 250                           | 200                    | 300                   | 100   |
| 4                  | 350                           | 300                    | 400                   | 100   |
| 5                  | 450                           | 400                    | 510                   | 110   |
| 6                  | 570                           | 510                    | 630                   | 120   |
| 7                  | 700                           | 630                    | 770                   | 140   |
| 8                  | 840                           | 770                    | 920                   | 150   |
| 9                  | 1000                          | 920                    | 1080                  | 160   |
| 10                 | 1170                          | 1080                   | 1270                  | 190   |
| 11                 | 1370                          | 1270                   | 1480                  | 210   |
| 12                 | 1600                          | 1480                   | 1720                  | 240   |
| 13                 | 1850                          | 1720                   | 2000                  | 280   |
| 14                 | 2150                          | 2000                   | 2320                  | 320   |
| 15                 | 2500                          | 2320                   | 2700                  | 380   |
| 16                 | 2900                          | 2700                   | 3150                  | 450   |
| 17                 | 3400                          | 3150                   | 3700                  | 550   |
| 18                 | 4000                          | 3700                   | 4400                  | 700   |
| 19                 | 4800                          | 4400                   | 5300                  | 900   |
| 20                 | 5800                          | 5300                   | 6400                  | 1100  |
| 21                 | 7000                          | 6400                   | 7700                  | 1300  |
| 22                 | 8500                          | 7700                   | 9500                  | 1800  |
| 23                 | 10500                         | 9500                   | 12000                 | 2500  |
| 24                 | 13500                         | 12000                  | 15500                 | 3500  |

Πίνακας 4.1: Κρίσιμες Ζώνες συχνότητας της Κλίμακας Bark

Η μονάδα μέτρησης Phon ουσιαστικά, στηρίζεται στη σύγκριση της ακουστότητας ενός καθαρού τόνου συχνότητας 1000Hz με ένα ηχητικό σήμα. Οι Fletcher και Munson [33]

πρωτοεισάγουν την παραπάνω μέθοδο προσαρμογής της ακουστότητας στην οποία βασίστηκε η κλίμακα Phon.

Οι παραπάνω μέθοδοι, χρησιμοποιούν ζώνες συχνοτήτων, οι οποίες αποτελούν αποτέλεσμα της μοντελοποίησης της λειτουργίας της βασικής μεμβράνης του αυτιού, όπου, σύμφωνα με μελέτες, καθορίζει το συχνοτικό εύρος του ακουστικού ερεθίσματος. Λαμβάνοντας υπόψη την παραπάνω λειτουργία, θεωρείται πως η βασική μεμβράνη λειτουργεί ως μια σειρά από επικαλυπτόμενα ζωνοπερατά φίλτρα τα οποία είναι γνωστά ως ακουστικά φίλτρα. Η βασική μεμβράνη ως ένα σύστημα παθητικής ισοστάθμισης, περιέχει μια τράπεζα φίλτρων. Για την μοντελοποίηση και τον υπολογισμό των ακουστικών φίλτρων υπάρχουν δύο βασικές μεθοδολογίες καθορισμού των ακουστικών ζωνών συχνοτήτων. Η πρώτη είναι η κλίμακα κρίσιμων ζωνών συχνότητας (Critical Band Rate - CB) [151] με μονάδα μέτρησης το Bark. Σε αυτή την κλίμακα έχουν καθορισθεί 24 κρίσιμες ζώνες συχνοτήτων (βλπ Πίν. 4.1) που αντιστοιχούν σε 24 Barks.

Στην κλίμακα CB όπως φαίνεται στον Πίνακα 4.1, αντιστοιχίζεται σε κάθε κρίσιμη συχνότητα το μέγεθος της ζώνης της σε μονάδες Hz. Στο Σχήμα 4.3, φαίνεται πως η κλίμακα CB, είναι πολύ κοντά στην κλίμακα 1/3 οκτάβας για συχνότητες από 500 Hz και πάνω, ενώ στις πιο χαμηλές συχνότητες το εύρος ζώνης παραμένει αμετάβλητο.

Στην Κλίμακα Ισοδύναμου Ορθογώνιου Εύρους Ζώνης (Equivalent Rectangular Bandwidth - ERB [41]), τα ακουστικά φίλτρα προέρχονται από τη χρήση της μεθόδου του Patterson [100], όπου το εύρος ζώνης του φίλτρου σε σχέση με την κεντρική συχνότητα είναι μεταξύ του 11% με 17%. Η μέση τιμή του ERB κάθε ακουστικού φίλτρου μετρήθηκε χρησιμοποιώντας μέτριο επίπεδο ηχητικών σταθμών από νέους ανθρώπους με κανονική ακοή και δηλώθηκε ως  $ERB_N$ . Η περιγραφή της τιμής της  $ERB_N$  σε συνάρτηση της κεντρικής συχνότητας δίνεται από την Εξίσωση 4.2 [41].

$$ERB_N = 24.7(4.37F + 1) \quad (4.2)$$

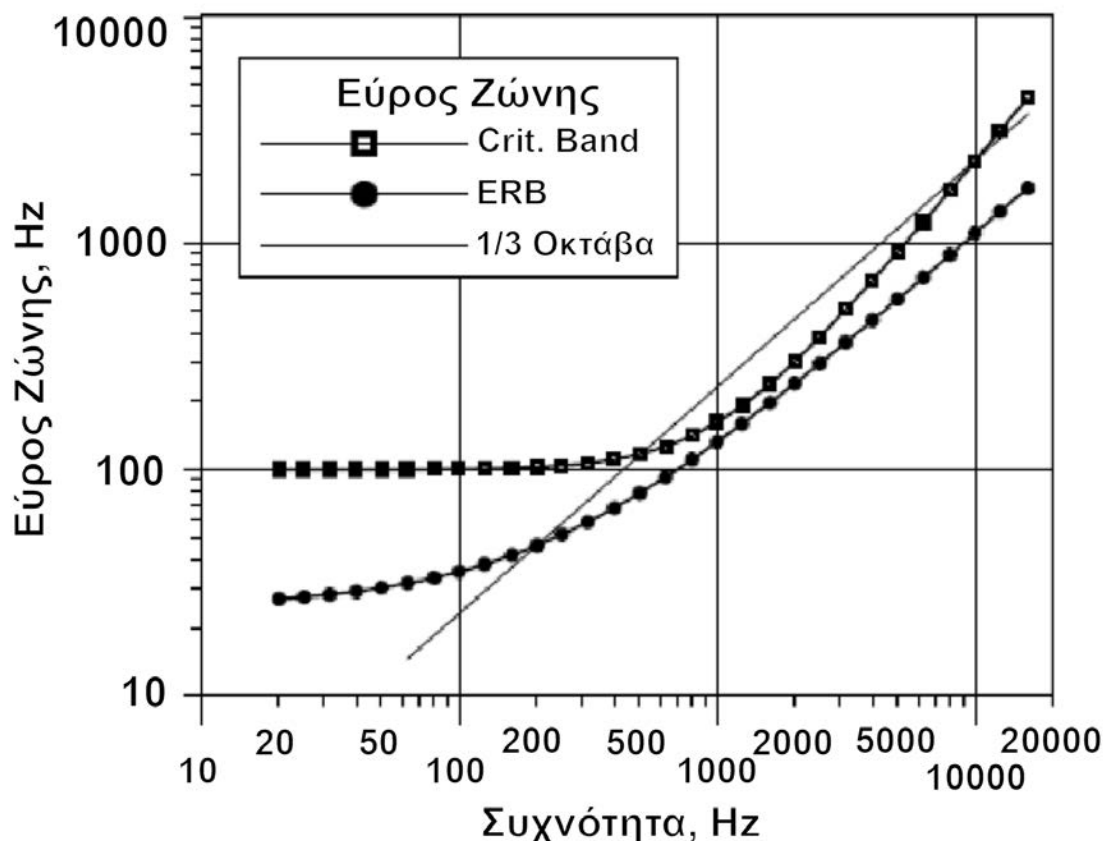
Όπου F είναι η κεντρική συχνότητα και μετριέται σε kHz και η  $ERB_N$  τιμή σε Hz.

Για την αποτύπωση των διαφορετικών  $ERB_N$  τιμών μέσα στην κλίμακα ERB, υφίσταται η Εξίσωση 4.3, η οποία δίνει μια τιμή στην κάθε διαφορετική ζώνη  $ERB_N$  και έχει ως μονάδα μέτρησης το Cam [1]. Πιο αναλυτικά, η μονάδα μέτρησης Cam, εκφράζει το βήμα μεταξύ των ζωνών της  $ERB_N$  κλίμακας.

$$ERB_N \text{ number} = 21.4 \log_{10}(4.37F + 1) \quad (4.3)$$

Σε σύγκριση με την κλίμακα CB, η ERB σε συχνότητες κάτω από 500Hz, έχει μικρότερο εύρος ζώνης. Την CB κλίμακα τη συναντάμε στο μοντέλο του Zwicker ενώ την κλίμακα ERB στο μοντέλο του Glasberg-Moore.

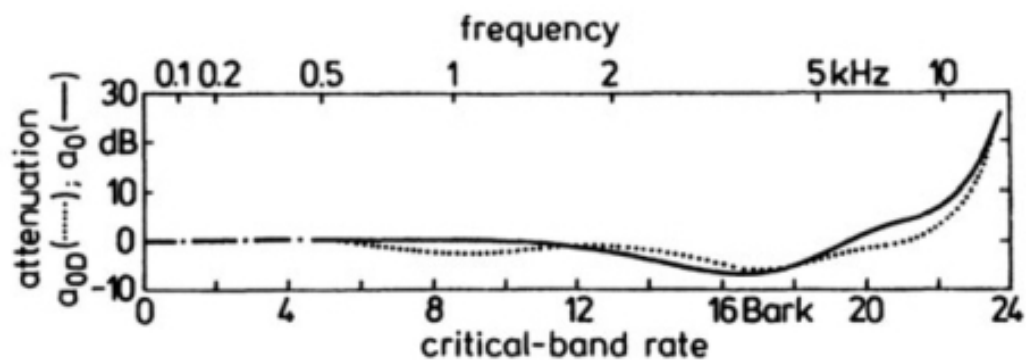
Οι παραπάνω κλίμακες, χρησιμοποιούνται για τον υπολογισμό του διαγράμματος διέγερσης της βασικής μεμβράνης. Το διάγραμμα διέγερσης υπολογίζεται με βάση τη στάθμη ηχητικής πίεσης του ήχου που εισέρχεται από το εξωτερικό αυτί μέσω του ακουστικού πόρου και φτάνει στον κοχλία. Συνεπώς για να υπολογιστεί το διάγραμμα διέγερσης, θα πρέπει να έχει



Σχήμα 4.3: Διάγραμμα κλίμακας CB, ERB και 1/3 Οκτάβας (Τροποποιημένη εικόνα από τη δημοσίευση του Jacobsen F., Poulsen T., Rindel J. H., Gade A. C., Ohlrich, M. σελ.72 [56])

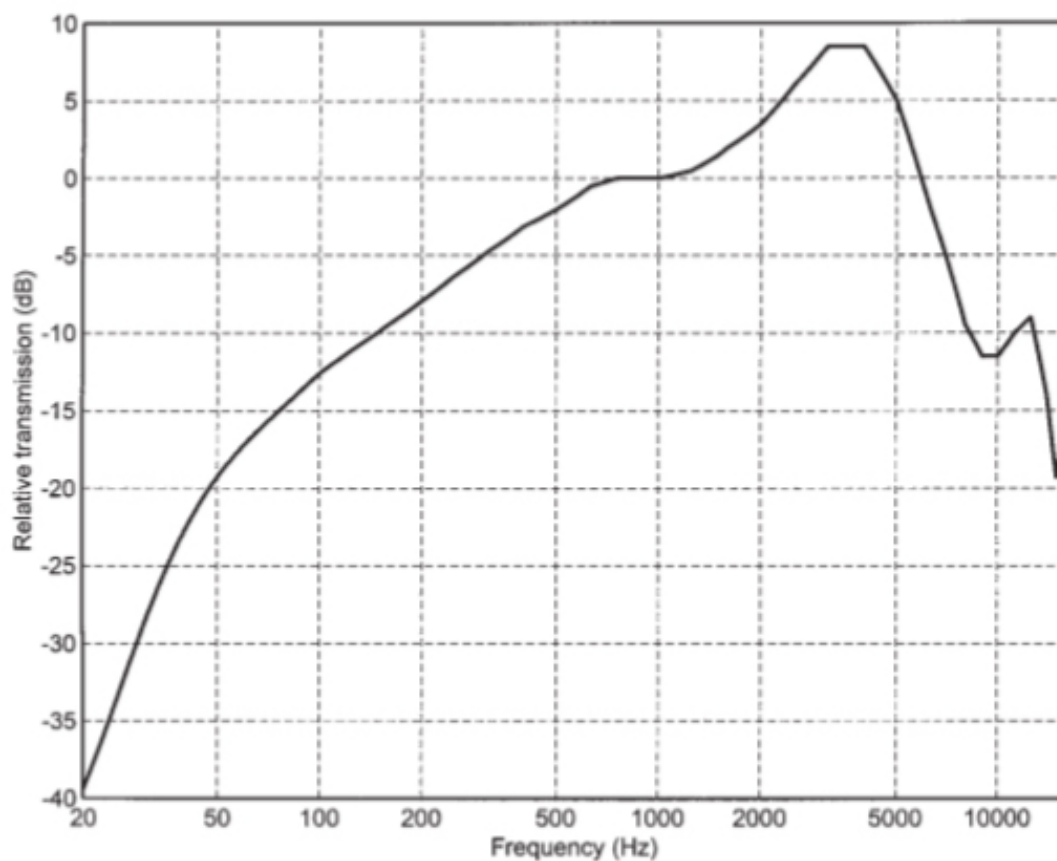
υλοποιηθεί κατάλληλο φιλτράρισμα. Στην περίπτωση του μοντέλου του Zwicker το Σχήμα 4.4 παρουσιάζει την εξασθένιση του ήχου στο πεδίο της συχνότητας σε ανοικτό και κλειστό πεδίο.

Αντίστοιχα, στο μοντέλο του Glasberg - Moore χρησιμοποιείται μια διορθωτική καμπύλη ίσης στάθμης όπου βασίζεται σε δύο υποθέσεις (βλπ Σχ. 4.5). Η πρώτη είναι ότι, σε συχνότητες πάνω από 1000 Hz, το εσωτερικό αυτί έχει την ίδια ευαισθησία σε όλες τις συχνότητες. Το εξωτερικό αυτί, επηρεάζει τη διακύμανση του απόλυτου κατωφλιού ακοής σε σχέση με τη συχνότητα. Η καμπύλη μετάδοσης είναι όμοια με την καμπύλη του απόλυτου κατωφλιού ακοής, αλλά αντεστραμμένη και είναι όμοια με εκείνη του διαγράμματος του Zwicker (βλπ Σχ. 4.4). Η δεύτερη υπόθεση που ολοκληρώνει την καμπύλη για τις συχνότητες κάτω από 1000 Hz, υποστηρίζει ότι η καμπύλη μετάδοσης έχει το ίδιο σχήμα με εκείνη της καμπύλης ίσης ακουστότητας στα 100 Phons αλλά αντεστραμμένη (βλπ Σχ. 4.5). Συνοπτικά, στο μοντέλο του Zwicker, για συχνότητες κάτω από 2 kHz, η μετάδοση από το έξω και μέσω αυτί είναι ομοιόμορφη για όλες τις συχνότητες, ενώ στο μοντέλο του Glasberg-Moore, για συχνότητες κάτω από 1 kHz η συνάρτηση μετάδοσης αντικατοπτρίζει το σχήμα της καμπύλης



Σχήμα 4.4: Διάγραμμα εξασθένισης για τον παράγοντα μετάδοσης του ήχου σε ελεύθερο πεδίο ( $a_0$ ) και κλειστό πεδίο ( $a_{0D}$ ). (Εικόνα από το βιβλίο του Zwicker Eberhard, Fastl Hugo σελ.226 [149])

του απόλυτου κατωφλιού.

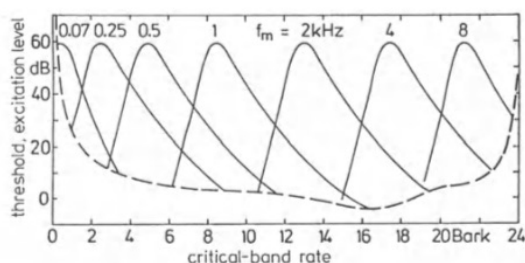


Σχήμα 4.5: Διάγραμμα εξασθένισης για τον παράγοντα μετάδοσης του ήχου. (Μη τροποποιημένη εικόνα από το άρθρο του Brian R. Glasberg, Brian C. J. Moore στο περιοδικό της JAES σελ.333 [42])

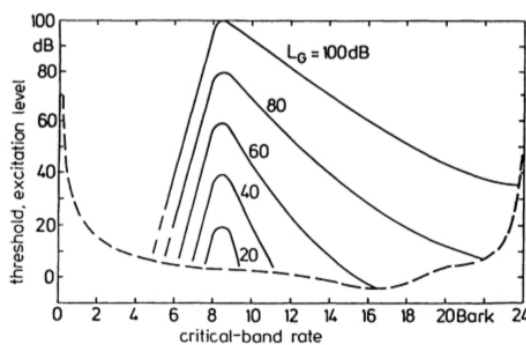
Ο υπολογισμός του διαγράμματος διέγερσης και στις δύο περιπτώσεις (Moore και Zwicker),

βασίζεται στο γνωστό φαινόμενο του ηχητικού μασκαρίσματος (Masking Effect) [34]. Στην περίπτωση όπου συμβαίνει το φαινόμενο του μασκαρίσματος, τα νεύρα της βασικής μεμβράνης διεγείρονται. Η διέγερση είναι μεγαλύτερη όπου υπάρχει εντονότερο το φαινόμενο του ηχητικού μασκαρίσματος. Το μασκάρισμα μπορεί να είναι για το εκάστοτε σήμα είτε ολικό, είτε μερικό [34].

Ο Zwicker, ορίζει τις κύριες διεγέρσεις ως τη μέγιστη τιμή της στάθμης της κρίσιμης ζώνης. Στην περίπτωση των κύριων διεγέρσεων, (δηλαδή στις τιμές της κρίσιμης ζώνης που αντιστοιχούν στην κεντρική συχνότητα της κρίσιμης ζώνης θορύβου μεγάλου εύρους, ή στην συχνότητα τόνων) η στάθμη της κρίσιμης ζώνης και η στάθμη διέγερσης είναι όμοιες. Ενώ στην περίπτωση λευκού θορύβου, λόγω του ότι δημιουργούνται μόνο κύριες διεγέρσεις, η στάθμη διέγερσης και η στάθμη της κρίσιμης ζώνης, είναι δυσδιάκριτη [149]. Οι καμπύλες μασκαρίσματος, δείχνουν την επιλεκτικότητα του αυτιού και εξαρτώνται από την κεντρική συχνότητα και την ηχητική στάθμη (βλπ Σχ. 4.6).



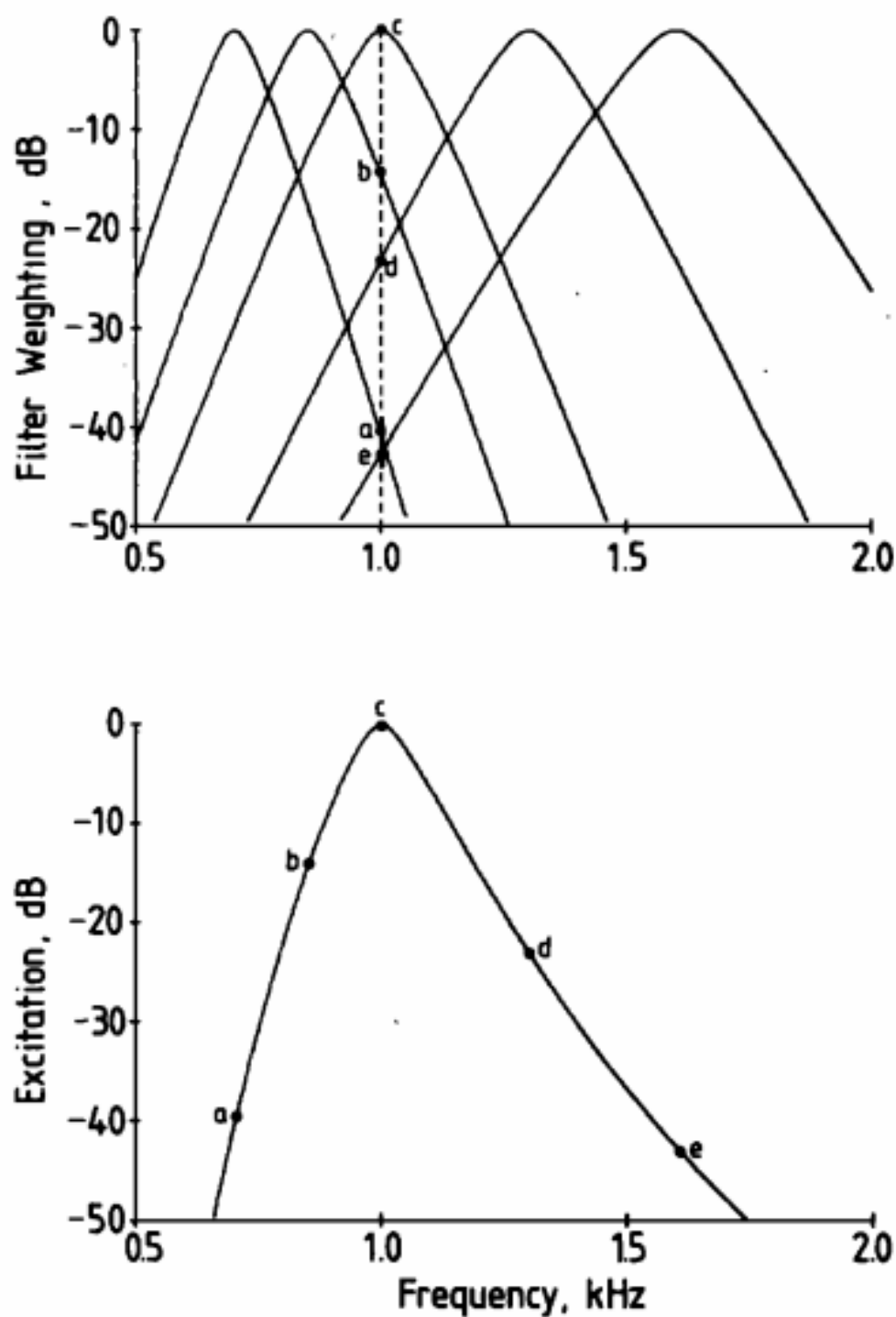
(α')



(β')

Σχήμα 4.6: Υπολογισμός διέγερσης, Μοντέλο Zwicker. (α) Διάγραμμα στάθμης διέγερσης ως προς τιμή κρίσιμης ζώνης για θορύβους στενής ζώνης με προκαθορισμένη κεντρική συχνότητα και στάθμη ακουστικής πίεσης 60 dB SPL (Μη τροποποιημένη εικόνα από το βιβλίο του Hugo Fastl, Eberhard Zwicker σελ.168 [32]). (β) Διάγραμμα στάθμης διέγερσης ως προς τιμή κρίσιμης ζώνης για θορύβους ευρείας ζώνης με κεντρική συχνότητα 1 kHz και στάθμες ακουστικής πίεσης κρίσιμων ζωνών όπως καταγράφονται. (Μη τροποποιημένη εικόνα από το βιβλίο του Hugo Fastl, Eberhard Zwicker σελ.169 [32]). Η διακεκομμένη γραμμή απεικονίζει και στα δύο διαγράμματα το κατώφλι ησυχίας.

Στο μοντέλο του Glasberg-Moore, το διάγραμμα διέγερσης, υπολογίζεται από την έξοδο των ακουστικών φίλτρων των οποίων οι κεντρικές συχνότητες συνθέτουν τον ήχο [88].



Σχήμα 4.7: Διάγραμμα υπολογισμού καμπύλης διέγερσης ημιτόνου 1 KHz από το μοντέλο του Moore - Glasberg. (Μη τροποποιημένη εικόνα από τη δημοσίευση του Moore, B. C. J., Glasberg, B. R. (1983). σελ.752 [88])

Για τον υπολογισμό της ειδικής ακουστότητας, η οποία αποτελεί την ακουστότητα για μια δεδομένη χρονική στιγμή, και της συνολικής ακουστότητας, χρησιμοποιείται ο νόμος του Stevens, ο οποίος συσχετίζει την ένταση του ερεθίσματος στο μέγεθος της προκαλούμενης αίσθησης και δίνεται από την Εξίσωση 4.4 [121].

$$N' = c.E^a \quad (4.4)$$

Όπου  $N'$  η ειδική ακουστότητα όπου είναι σχετική της έντασης της διέγερσης  $E$  με τον λόγο δύναμης. Η σταθερά  $C$  μεταβάλλει το μέγεθος της μονάδας μέτρησης και  $a$  είναι η εκθετική σταθερά όπου εξαρτάται από τον τύπο της ενέργειας όπου στην συγκεκριμένη περίπτωση είναι η ακουστότητα [149].

Στο μοντέλο του Zwicker, η ειδική ακουστότητα υπολογίζεται από την διέγερση της κάθε κρίσιμης ζώνης με την χρήση της Εξίσωσης 4.5, η οποία βασίζεται στη διέγερση στο κατώφλι ησυχίας, στη διέγερση στη συγκεκριμένη συχνότητα και στη διέγερση με ένταση αναφοράς  $I_0 = 10^{-12} \text{W/m}^2$ , ήτοι:

$$N' = c. \left( \frac{E_{TQ}}{E_0} \right)^a \left[ \left( 0.5 + 0.5 \frac{E}{E_{TQ}} \right)^a - 1 \right] \quad (4.5)$$

Όπου  $N'$  είναι η ειδική ακουστότητα (*sones/Bark*),  $E_{RQ}$  είναι η διέγερση στο κατώφλι ησυχίας,  $E$  είναι η διέγερση συγκεκριμένης συχνότητας,  $E_0$  είναι η διέγερση για την ένταση αναφοράς  $I_0$  και σταθερές  $c$ ,  $a$ . Η σταθερά  $a$  ορίζεται από το μοντέλο του Zwicker στα 0.23 για την ακουστότητα και η  $c$  αποτελεί τη σταθερά καθορισμού του μεγέθους της κλίμακας. Η συνολική ακουστότητα, υπολογίζεται από το άθροισμα της ειδικής ακουστότητας όλων των κρίσιμων ζωνών.

Το μοντέλο του Glasberg-Moore, είναι παρόμοιο με το μοντέλο του Zwicker. Η ειδική ακουστότητα υπολογίζεται από την καμπύλη διέγερσης του μοντέλου Glasberg-Moore (βλπ Σχ. 4.7). Σε αυτή την περίπτωση ο θόρυβος στο φόντο του αυτιού θεωρείται μη ακουστός. Αυτή η υπόθεση μοντελοποιείται με την παρακάτω αφαίρεση:

$$\begin{aligned} N' &= N'_{St} - N'_{IntNoise} = a.(E_{St})^b - a.(E_{TQ})^b \\ N' &= a. \left[ (E_{St})^b - (E_{TQ})^b \right] \end{aligned} \quad (4.6)$$

Όπου για την ίδια ERB ζώνη, υπολογίζεται ως  $N'$  η ειδική ακουστότητα της,  $N'_{St}$  η ειδική ακουστότητα του ερεθίσματος και  $N'_{IntNoise}$  η ειδική ακουστότητα του κατωφλιού ησυχίας. Το άθροισμα από τις ειδικές ακουστότητες κάθε ERB ζώνης, έχει ως αποτέλεσμα την συνολική ακουστότητα.

Τα μοντέλα ακουστότητας, υπολογίζουν την ακουστότητα για δύο γενικές κατηγορίες ήχων. Η πρώτη κατηγορία είναι οι σταθεροί ήχοι και η δεύτερη οι μεταβλητοί. Η διαφορά στον υπολογισμό της ακουστότητας μεταξύ των δύο κατηγοριών, είναι η χρονική και φασική επικάλυψη, η οποία μεταφέρει τον υπολογισμό της ακουστότητας στο πεδίο του χρόνου.

Η υλοποίηση του μοντέλου του Zwicker έχει γίνει με αναλογικό τρόπο και του Glasberg-Moore ψηφιακά [42]. Το μοντέλο του Glasberg-Moore, υλοποιείται σύμφωνα με τις παρακάτω

φάσεις [42]. Αρχικά χρησιμοποιείται ένα φίλτρο πεπερασμένης κρουστικής απόκρισης (FIR) αποτελούμενο από 4097 συντελεστές για την μοντελοποίηση του φιλτραρίσματος εξωτερικού και μέσου αυτιού. Για τον υπολογισμό του τρεχούμενου φάσματος και της καμπύλης διέγερσης, χρησιμοποιούνται 6 παράλληλες αναλύσεις τύπου FFT με παραθυροποίηση τύπου Hanning και διάρκειες  $2ms$ ,  $4ms$ ,  $8ms$ ,  $16ms$ ,  $32ms$ ,  $64ms$  με τα κέντρα τους χρονικά ευθυγραμμισμένα. Η κάθε ανάλυση FFT, περιλαμβάνει τα συχνοτικά εύρη:  $20Hz - 80Hz$ ,  $80Hz - 500Hz$ ,  $500Hz - 1250Hz$ ,  $1250Hz - 2540Hz$ ,  $2540Hz - 4050Hz$ ,  $4050Hz - 15000Hz$  με τις αντίστοιχες παραπάνω διάρκειες. Ακολούθως, υπολογίζεται η στιγμιαία ακουστότητα, η μακροπρόθεσμη και η βραχυπρόθεσμη ακουστότητα [90].

Η υλοποίηση τέτοιων μοντέλων σε πραγματικό χρόνο, αποτελεί μεγάλη πρόκληση, λόγω της υπολογιστικής πολυπλοκότητας και της επεξεργαστικής καθυστέρησης που εισάγουν στα ψηφιακά ηχητικά συστήματα [131]. Παραδείγματα υλοποίησης συστημάτων εκτίμησης αμφιωτικής ακουστότητας έχουν υπάρξει στο παρελθόν, χρησιμοποιώντας αναλογικά συστήματα για εξάλειψη τυχόν καθυστερήσεων [148], αλλά και χρησιμοποιώντας τεχνικές ψηφιακής επεξεργασίας [130, 140, 141].

Στην επαυξημένη πραγματικότητα ήχου, λόγω της χρήσης αμφιωτικής τεχνολογίας και ειδικότερα ακουστικών, ο υπολογισμός της ακουστότητας θα πρέπει να λάβει υπόψιν τον αμφιωτικό χαρακτήρα του ηχητικού ερεθίσματος. Τα μοντέλα αμφιωτικής ακουστότητας λαμβάνουν υπόψιν την άθροιση του ηχητικού ερεθίσματος, λόγω της αμφιωτικής ακοής, και μέσω αυτής της ιδιότητάς τους, μπορούν να συνεισφέρουν στην υλοποίηση του μηχανισμού προσαρμοστικής ARA μείξης.

Στη συνέχεια θα παρουσιαστεί η μεθοδολογία ανάπτυξης του μοντέλου υπολογισμού αμφιωτικής ακουστότητας που χρησιμοποιήθηκε μέσα στο σύστημα προσαρμοστικής μείξης επαυξημένης πραγματικότητας ήχου. Στόχος ήταν η υλοποίηση ενός συστήματος με λειτουργία σε πραγματικό χρόνο, ώστε να είναι εφικτή η αξιολόγηση της επίδρασης της προσαρμοστικής ARA μείξης μέσα σε επαυξημένα περιβάλλοντα ήχου.

#### 4.4.2 Προτεινόμενη μέθοδος δημιουργίας προϋπολογισμένης μήτρας με βάση την Αμφιωτική Ακουστότητα

Για την εμβύθιση του ακροατή στα επαυξημένα περιβάλλοντα ήχου, υπάρχει η άποψη της ακουστικής ισότητας μεταξύ του πραγματικού και του εικονικού ηχητικού περιβάλλοντος [106]. Η παραπάνω ισότητα για να είναι εφικτή, απαιτεί θεωρητικά τον υπολογισμό της συνολικής ακουστότητας του κάθε περιβάλλοντος και την προσαρμογή του εικονικού στο πραγματικό.

Η λειτουργία του προσαρμοστικού μοντέλου ARA μείξης σε πραγματικό χρόνο, δημιουργεί την ανάγκη υλοποίησης μεθόδου άμεσης σύγκρισης του πραγματικού με το εικονικό ηχητικό περιβάλλον με όρους ακουστικής αντίληψης και πρόβλεψης της απαιτούμενης ηχητικής ισοστάθμισης που χρειάζεται να εφαρμοστεί στο τελικό επαυξημένο εικονικό σήμα. Τα μοντέλα αμφιωτικής ακουστότητας, επιτρέπουν την εκτίμηση της ακουστότητας του κάθε περιβάλλοντος ωστόσο για να γίνει αυτό σε πραγματικό χρόνο μέσα σε ένα επαυξημένο περιβάλλον ήχου χρειάζεται να ληφθούν υπόψιν συγκεκριμένοι παράμετροι ενός ARA συστήματος και οι

περιορισμοί των υφιστάμενων μοντέλων υπολογισμού αμφιωτικής ακουστότητας.

Αρχικά, ένα ARA σύστημα, αποτελείται από πολλαπλές ηχητικές ροές. Συνεπώς, χρειάζεται παράλληλη επεξεργασία σημάτων και όσο το δυνατόν μικρότερη επεξεργαστική καθυστέρηση. Η εισαγωγή και η υλοποίηση σε πραγματικό χρόνο ενός από τα υφιστάμενα μοντέλα αμφιωτικής ακουστότητας, φαίνεται αδύνατη υπό αυτούς τους όρους σύμφωνα και με την Ενότητα 4.4.1.

Για το λόγο αυτό, επιλέχθηκε ως μεθοδολογία η υλοποίηση ενός νέου μοντέλου εκτίμησης αμφιωτικής ακουστότητας, το οποίο κάνει χρήση μιας μήτρας προϋπολογισμένων τιμών αμφιωτικής ακουστότητας, όπου για τον υπολογισμό τους χρησιμοποιείται ένα υφιστάμενο μοντέλο. Στην παρούσα διατριβή επιλέχθηκε η χρήση του μοντέλου του Glasberg-Moore. Ο συγκεκριμένος τρόπος είναι παραμετροποιήσιμος ως προς την επιλογή του μοντέλου, οπότε θα μπορούσε να επιλεγεί και το μοντέλο του Zwicker. Λαμβάνοντας υπόψη την μήτρα πραγματοποιείται ο μετασχηματισμός την ενέργεια διέγερσης (με μονάδα μέτρησης  $dB SPL$ ) σε μονάδες ακουστότητας *Phons*, μέσω του υπολογισμού της αμφιωτικής ακουστότητας από την καμπύλη διέγερσης του κάθε περιβάλλοντος της ARA μείζης.

Για τον καθορισμό των κρίσιμων ζωνών επιλέχθηκε η κλίμακα οκτάβας. Σύμφωνα με την κλίμακα οκτάβας, επιλέχθηκαν οι κεντρικές συχνότητες των κρίσιμων ζωνών (βλπ Πίν. 4.2). Η επιλογή της οκταβικής κλίμακας, διευκολύνει την μαθηματική μοντελοποίηση και βρίσκεται πολύ κοντά στα εύρη ζώνης που παρουσιάζουν οι κλίμακες CB και ERB (βλπ Σχ. 4.3) για συχνότητες πάνω από  $250Hz$ .

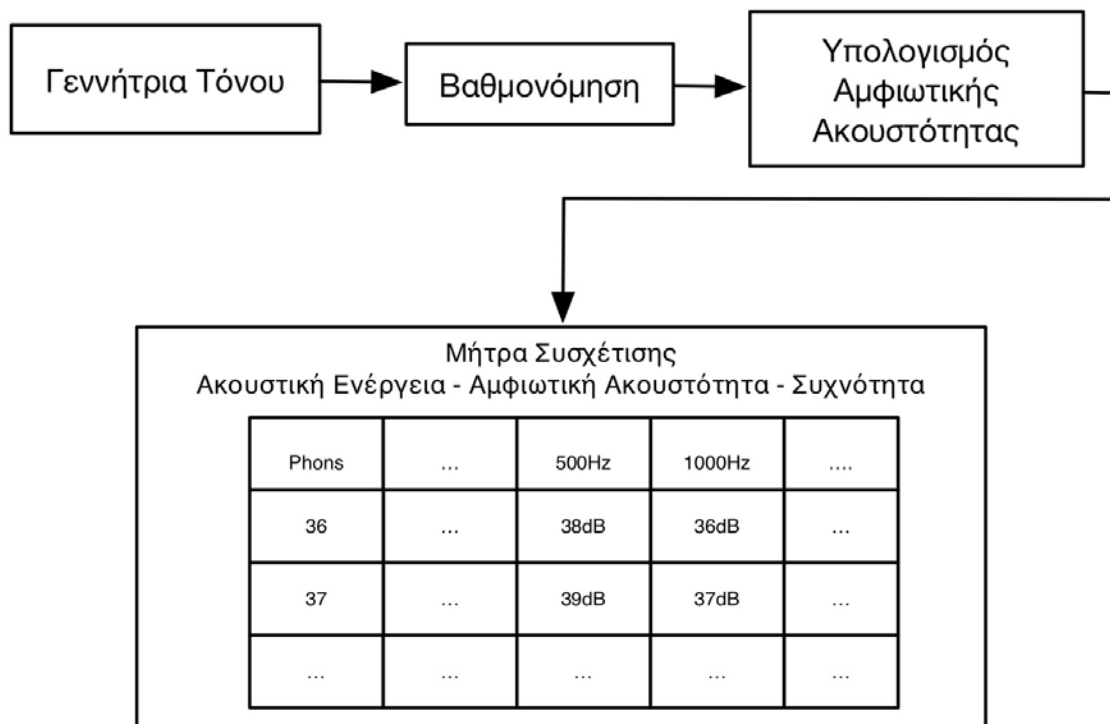
| Κεντρική Συχνότητα | Ζώνη Συχνοτήτων     |
|--------------------|---------------------|
| $250Hz$            | $0Hz - 335Hz$       |
| $500Hz$            | $335Hz - 710Hz$     |
| $1000Hz$           | $710Hz - 1420Hz$    |
| $2000Hz$           | $1420Hz - 2840Hz$   |
| $4000Hz$           | $2840Hz - 5680Hz$   |
| $8000Hz$           | $5680Hz - 11360Hz$  |
| $16000Hz$          | $11360Hz - 22720Hz$ |

Πίνακας 4.2: Κρίσιμες Ζώνες Συχνοτήτων σε διαστήματα Οκτάβας

Για την δημιουργία της προϋπολογισμένης μήτρας, επιλέχθηκε ο υπολογισμός της αμφιωτικής ακουστότητας καθαρών τόνων οι οποίοι αντιπροσωπεύουν την κεντρική συχνότητα των επιλεγμένων κρίσιμων ζωνών. Στο Σχήμα 4.8 φαίνεται η μεθοδολογία που σχεδιάστηκε για τον υπολογισμό της μήτρας.

Σύμφωνα με το Σχήμα 4.8, δημιουργείται ο κατάλληλος τόνος σύμφωνα με τις κεντρικές συχνότητες των επιλεγμένων κρίσιμων ζωνών (βλπ Πίν. 4.2). Ο συγκεκριμένος τόνος βαθμονομείται σε κλίμακα ακουστικής πίεσης ( $dB SPL$ ) και ακολούθως υπολογίζεται η αμφιωτική ακουστότητά του με χρήση υφιστάμενου μοντέλου όπως παρουσιάστηκε στην Ενότητα 4.4.1.

Στόχος είναι να υπολογιστούν και να αποθηκευτούν οι ακουστότητες του κάθε τόνου για μεταβολές της ακουστικής πίεσης ίσης με  $1 dB SPL$ . Η παραγόμενη μήτρα απεικονίζεται στο



Σχήμα 4.8: Μεθοδολογία Υπολογισμού Πίνακα Μετατροπής Ακουστικής Ενέργειας ( $dB SPL$ ), Αμφιωτικής Ακουστότητας ( $Phons$ ), Συχνότητας ( $Hz$ )

Παράρτημα Α'.

Ο πίνακας μετατροπής (βλπ Παράρτημα Α'), αποτελεί τα δεδομένα εισόδου για τη γρήγορη εκτίμηση της αμφιωτικής ακουστότητας, σύμφωνα με τις τιμές διέγερσης του εικονικού και ψευδοακουστικού περιβάλλοντος για την οποία η μέθοδος υπολογισμού περιγράφεται στην επόμενη ενότητα.

## 4.5 Μεθοδολογία εκτίμησης Αμφιωτικής Ακουστότητας των σημάτων εισόδου

Σε αυτή την ενότητα θα παρουσιαστεί η μεθοδολογία που χρησιμοποιήθηκε για την εκτίμηση της αμφιωτικής ακουστότητας των εισερχόμενων ηχητικών ροών στο σύστημα προσαρμοστικής ARA μείξης. Τα στάδια που θα περιγραφούν είναι η βαθμονόμηση του σήματος, η τεχνική υπολογισμού της καμπύλης ακουστικής διέγερσης των σημάτων και η τεχνική εκτίμησης της αμφιωτικής ακουστότητας των σημάτων.

### 4.5.1 Βαθμονόμηση Σημάτων Εισόδου

Τα σήματα εισόδου αποτελούνται από διακριτά δείγματα στο χρόνο με κβαντισμένες τιμές. Οι τιμές κβάντισης, μεταφράζονται από το εκάστοτε πρόγραμμα επεξεργασίας με διαφορετικό τρόπο. Για να γίνει η μετατροπή των κβαντισμένων τιμών σε μια κοινά αντιληπτή μονάδα

μέτρησης, χρειάζεται να είναι γνωστός ο τρόπος μετατροπής και οι τιμές που αποδίδει η εκάστοτε διεργασία στις κβαντισμένες τιμές. Η διαδικασία που περιγράφει την παραπάνω μετατροπή καλείται βαθμονόμηση. Στην περίπτωση του προτεινόμενου μοντέλου, όλα τα σήματα εισόδου βαθμονομούνται. Στην περίπτωση του προτεινόμενου μοντέλου, το κβαντισμένο σήμα θα πρέπει να βαθμονομηθεί σε μονάδα ακουστικής πίεσης ( $dB SPL$ ) και επιπλέον να εισαχθεί και ένας συντελεστής βαθμονόμησης λαμβάνοντας υπόψιν την επιθυμητή τιμή του σήματος κατά την αναπαραγωγή του. Παρακάτω παρατίθενται τα στάδια υπολογισμού του συντελεστή βαθμονόμησης του σήματος.

Η επιθυμητή τιμή ακουστικής πίεσης για να υπολογιστεί θα πρέπει να καθοριστεί η στάθμη αναφοράς για μετάδοση του ήχου στον αέρα η οποία ορίζεται στα  $20 \times 10^{-6} pascal$  (βλπ Εξ. 4.7).

$$P_{ref} = 20 \times 10^{-6} Pa \quad (4.7\alpha')$$

$$x_{refscaled} = \frac{x}{P_{ref}} \quad (4.7\beta')$$

Όπου  $x$  είναι το εισερχόμενο σήμα και  $x_{refscaled}$  το σήμα σε κλίμακα  $pascal$  με μονάδα μέτρησης τα  $dB SPL$ . Κατόπιν υπολογίζεται η  $RMS$  τιμή του σήματος (βλπ Εξ. 4.8).

$$RMS(x_{refscaled}) = \sqrt{\frac{\sum_1^N x_{refscaled}^2}{N}} \quad (4.8)$$

Η έκφραση της  $RMS$  τιμή του σήματος σε στάθμη ακουστικής πίεσης δίνεται από την Εξίσωση 4.9.

$$SPL_{RMS(x_{refscaled})} = 20 \log_{10}(RMS(x_{refscaled})) \quad (4.9)$$

Όπου για τον υπολογισμό του συντελεστή, θεωρούμε την στάθμη ακουστικής πίεσης που έχει το σήμα στον πραγματικό κόσμο την στιγμή της μέτρησης (βλπ Εξ. 4.10). Για παράδειγμα σε περιπτώσεις ομιλίας η στάθμη μέτρησης είναι  $70 dB SPL$ .

$$c = \frac{10^{((SPL_{Measure} - SPL_{RMS(x_{refscaled})})/20)}}{rms(x)} \quad (4.10)$$

Όπου  $SPL_{Measure}$  η στάθμη μέτρησης στον πραγματικό κόσμο και  $c$  ο συντελεστής βαθμονόμησης (βλπ Εξ. 4.11) που πρέπει να εφαρμοστεί ώστε το σήμα να βαθμονομηθεί στο επίπεδο της μέτρησης. Πρέπει να σημειωθεί πως η στάθμη μέτρησης δεν είναι ίδια για όλες τις ηχητικές ροές. Δηλαδή η αμφιωτική αναπαραγωγή του ψευδοακουστικού περιβάλλοντος διαιρείται σε δύο ηχητικά κανάλια όπου το καθένα βαθμονομείται, λαμβάνοντας υπόψιν την διαιρεμένη κατανομή ενέργειας σε δύο κανάλια, ενώ το εικονικό περιβάλλον ως μια εισερχόμενη ροή, βαθμονομείται ως αμφιωτική αναπαραγωγή.

Στο τελικό στάδιο, πρέπει να γίνει η εφαρμογή του συντελεστή βαθμονόμησης στο σήμα για να παραχθεί η βαθμονομημένη στάθμη ακουστικής πίεσης.

$$x_{calibrated} = c \cdot x_{refscaled} \quad (4.11)$$

Η μέθοδος της βαθμονόμησης, χρησιμοποιείται εκτός από τα σήματα εισόδου, και στην δημιουργία του πίνακα μετατροπής Αμφιωτικής Ακουστότητας (βλπ Εν. 4.4.2). Σε αυτή την περίπτωση όλα τα σήματα που αναλύονται, έχουν πρωτύτερα βαθμονομηθεί σε προκαθορισμένες ηχητικές στάθμες.

#### 4.5.2 Τεχνική υπολογισμού ενέργειας διέγερσης ηχητικών ροών

Σε αυτή την ενότητα θα περιγραφεί η μέθοδος που ακολουθήθηκε για τον υπολογισμό της διέγερσης των ακουστικών ροών του επαυξημένου περιβάλλοντος. Η μέθοδος που ακολουθήθηκε, τόσο στο σήμα του εικονικού περιβάλλοντος, όσο και στο ψευδοακουστικό, είναι κοινή. Σε αυτό το στάδιο, η επεξεργασία του σήματος μεταβαίνει στο πεδίο της συχνότητας.

Σύμφωνα με το υφιστάμενο μοντέλο του Glasberg-Moore, για τον υπολογισμό της διέγερσης του ακουστικού σήματος, χρειάζεται ο καθορισμός των ακουστικών περιοχών και η χρήση παράλληλων αναλύσεων γρήγορου μετασχηματισμού Fourier (FFT). Η παραπάνω διαδικασία έχει μεγάλο υπολογιστικό κόστος και καθυστέρηση στην ηχητική επεξεργασία του σήματος.

Στην περίπτωση της ARA, όπου η συγκεκριμένη επεξεργασία χρειάζεται να γίνει σε περισσότερα από ένα σήματα, όπως είναι το αμφιωτικό σήμα του ψευδοακουστικού περιβάλλοντος και έστω σε μια εικονική ηχητική πηγή, έχει σαν φυσικό αποτέλεσμα πολλαπλάσιο υπολογιστικό κόστος και εισαγωγή καθυστέρησης. Για την ομαλή λειτουργία του μοντέλου σε πραγματικό χρόνο, σχεδιάστηκε η πραγματοποίηση μίας ανάλυσης FFT σε κάθε ροή σήματος.

Ωστόσο, η ανάλυση αποτελεί ένα κομβικό στοιχείο του μοντέλου καθώς επηρεάζει τη λειτουργία του, και επηρεάζεται από συγκεκριμένους παράγοντες. Αρχικά, καθώς το μοντέλο λειτουργεί σε πραγματικό χρόνο, η επεξεργασία του σήματος γίνεται με την μέθοδο της παραθυροποίησης των ηχητικών δειγμάτων. Ουσιαστικά το εισερχόμενο σήμα διαιρείται σε παράθυρα τα οποία ανανεώνονται από το μοντέλο ανάλογα με τον ορισμένο ρυθμό δειγματοληψίας.

Ο ρυθμός δειγματοληψίας που επιλέχθηκε για το μοντέλο είναι τα 48 kHz. Με αυτό το ρυθμό δειγματοληψίας επιτυγχάνεται η επιτυχής δειγματοληψία ολόκληρου του ακουστικού συχνοτικού φάσματος (20 Hz — 20 kHz) σύμφωνα με το θεώρημα του Nyquist [82], καθώς η συχνότητα δειγματοληψίας είναι διπλάσια της μέγιστης συχνότητας που εισέρχεται στο μοντέλο.

Είναι αναγκαίο να υπάρχει δειγματοληψία που να υποστηρίζει όλο το συχνοτικό φάσμα στα ηχητικά σήματα του επαυξημένου περιβάλλοντος ήχου. Αυτό στηρίζεται στα ακόλουθα χαρακτηριστικά της επαυξημένης πραγματικότητας ήχου. Η αμφιωτική αναπαραγωγή του ψευδοακουστικού περιβάλλοντος για να είναι ρεαλιστική, διέρχεται από το κατάλληλο φιλτράρισμα (βλπ Εν. 4.3). Η εισερχόμενη φιλτραρισμένη ροή, εκτείνεται σε όλο το συχνοτικό φάσμα της ανθρώπινης ακοής και έχει απρόβλεπτο χαρακτήρα. Η διατήρηση των φασματικών χαρακτηριστικών είναι σημαντική για τον βέλτιστο υπολογισμό της ενέργειας της κάθε συχνοτικής

ζώνης. Το επόμενο χαρακτηριστικό είναι το συχνοτικό περιεχόμενο της εικονική ηχητικής πηγής, το οποίο, για να υποστηρίξει ολόκληρο το φάσμα συχνοτήτων αντίστοιχα, θα πρέπει να χρησιμοποιείται επαρκής συχνότητα δειγματοληψίας.

Ο ρυθμός δειγματοληψίας σε συνθήκες επεξεργασίας πραγματικού χρόνου μέσα στο μοντέλο, επηρεάζει το χρόνο καθυστέρησης στην επεξεργασία ο οποίος μπορεί να εκφραστεί ως συνάρτηση του παραθύρου δεδομένων (βλπ Εξ. 4.12).

$$St = \frac{W}{Sr} \quad (4.12)$$

Όπου  $St$  ο χρόνος ανανέωσης του παραθύρου σε δευτερόλεπτα,  $W$  το μέγεθος του παραθύρου σε δείγματα (*Samples*) και  $Sr$  ο ρυθμός δειγματοληψίας σε δείγματα ανά δευτερόλεπτο (*samples/second*). Στο μοντέλο προσαρμοστικής ARA μείζης, επιλέχθηκε παράθυρο μεγέθους 1024 δειγμάτων. Οπότε σύμφωνα με τον τύπο της Εξίσωσης 4.12, η καθυστέρηση που εισάγεται στο μοντέλο είναι της τάξης των 20 *microseconds*.

Η χρήση μικρότερου παραθύρου, θα βοηθούσε στην ελάττωση της καθυστέρησης όπως για παράδειγμα είναι ένα παράθυρο με 64 δείγματα (*samples*). Αυτό όμως περιορίζει τις διαθέσιμες συχνότητες ανάλυσης του μετασχηματισμού FFT μόλις σε 32 σε ολόκληρο το ακουστικό συχνοτικό φάσμα (20 *Hz* – 20 *kHz*). Το ζητούμενο στην επιλογή του μεγέθους του παραθύρου είναι η διαθέσιμη ανάλυση των κρίσιμων ακουστικών συχνοτήτων που απεικονίζονται στον Πίνακα 4.2.

| Κεντρική Συχνότητα | Εύρος Ζώνης                       | FFT Bins  |
|--------------------|-----------------------------------|-----------|
| 250 <i>Hz</i>      | 0 <i>Hz</i> – 335 <i>Hz</i>       | 1 : 7     |
| 500 <i>Hz</i>      | 335 <i>Hz</i> – 710 <i>Hz</i>     | 7 : 15    |
| 1000 <i>Hz</i>     | 710 <i>Hz</i> – 1420 <i>Hz</i>    | 15 : 30   |
| 2000 <i>Hz</i>     | 1420 <i>Hz</i> – 2840 <i>Hz</i>   | 30 : 61   |
| 4000 <i>Hz</i>     | 2840 <i>Hz</i> – 5680 <i>Hz</i>   | 61 : 121  |
| 8000 <i>Hz</i>     | 5680 <i>Hz</i> – 11360 <i>Hz</i>  | 121 : 242 |
| 16000 <i>Hz</i>    | 11360 <i>Hz</i> – 22720 <i>Hz</i> | 242 : 512 |

Πίνακας 4.3: Ζώνες Φασματικής Ανάλυσης για παράθυρο 1024 δειγμάτων με ρυθμό δειγματοληψίας 48 *kHz*

Η επιλογή του παραθύρου των 1024 δειγμάτων σε φασματική ανάλυση FFT, δίνει διαθέσιμα 512 σημεία (*bins*) σε συχνότητες από 0 *Hz* – 24 *kHz* σύμφωνα με τον ρυθμό δειγματοληψίας των 48 *kHz*. Οπότε για τις κρίσιμες συχνότητες με τα επιλεγμένα εύρη ζώνης, τα FFT σημεία που αποτυπώνονται στον Πίνακα 4.3 επαρκούν για κάθε επιλεγμένο εύρος ζώνης.

Ο υπολογισμός της μέσης τιμής ενέργειας για τις προκαθορισμένες ακουστικές ζώνες φασματικής ανάλυσης για το εκάστοτε περιβάλλον, αποτελούν την ακουστική διέγερση της κάθε ζώνης και από το συνδυασμό όλων των διεγέρσεων από τις ακουστικές ζώνες, συντίθεται ο πίνακας ακουστικής διέγερσης ο οποίος ουσιαστικά συνθέτει την χαμπύλη ακουστικής διέγερσης, του εκάστοτε σήματος εισόδου. Σύμφωνα με τον πίνακα διέγερσης, γίνεται η εκτίμηση

της αμφιωτικής ακουστότητας η οποία παρουσιάζεται στην επόμενη ενότητα. Παρακάτω παρουσιάζεται η μεθοδολογία υπολογισμού της ενέργειας διέγερσης για κάθε ακουστική ζώνη.

Στην Εξίσωση 4.13 ορίζεται ως  $M$  ο αριθμός των ζωνών και ως  $n$  ο αριθμός των διακριτών σημείων συχνότητας της φασματικής ανάλυσης για την εκάστοτε συχνοτική ζώνη:

$$\begin{aligned} M &= 0 \dots M - 1 \\ n &= FFTbins_{(M)} \end{aligned} \quad (4.13)$$

Ο υπολογισμός της ακουστικής διέγερσης προκύπτει από την Εξίσωση 4.14 όπου  $E_{dB_{(M)}}$  είναι η μέση τιμή της κάθε συχνοτικής ζώνης.

$$E_{dB_{(M)}} = \frac{\sum_1^n E_{(n,M)}}{n} \quad (4.14)$$

Για το ψευδοακουστικό περιβάλλον, λόγω της μεταφοράς του από το ανοικτό πεδίο μέσω του αυτιού στον τυμπανικό υμένα, έχει ληφθεί υπόψιν το φιλτράρισμα του έξω και μέσου αυτιού μέσα από την τεχνική αμφιωτικής αναπαραγωγής των σημάτων εισόδου (βλπ Εν. 4.3). Για την περίπτωση του εικονικού περιβάλλοντος, η εικονική ηχητική πηγή, δεν χρειάζεται να φιλτραριστεί καθώς στόχος είναι να προσαρμοστεί μέσα από το προτεινόμενο μοντέλο στο ψευδοακουστικό περιβάλλον.

#### 4.5.3 Αντιμετώπιση ακουστών παραμορφώσεων διεργασίας υπολογισμού καμπύλης διέγερσης

Η μέτρηση της ακουστικής διέγερσης στο μοντέλο, είναι διαρκής. Ο αριθμός των εναλλαγών στην ακουστική ενέργεια είναι ανάλογος του ρυθμού ανανέωσης του μοντέλου. Οι διαρκείς φασματικές μεταβολές του πραγματικού περιβάλλοντος, μπορεί να παραμορφώσουν ή να δημιουργήσουν έντονες διακυμάνσεις στο προσαρμοσμένο εικονικό σήμα και κατ' επέκταση στο τελικό επαυξημένο σήμα. Αυτό οφείλεται στον τρόπο λειτουργίας του αλγόριθμου προσαρμογής, καθώς δεν μπορεί να διακρίνει λόγω της υλοποίησης, για ποιες μεταβολές να μην προχωρήσει σε προσαρμογή. Για την εξάλειψη αυτού του προβλήματος, επιλέχθηκε η διάκριση των φασματικών μεταβολών πριν την εισαγωγή στον αλγόριθμο προσαρμογής σε εκείνες που είναι ωφέλιμες για την προσαρμογή και σε εκείνες που θα το παραμορφώσουν σε μεγάλο βαθμό.

Για τον έλεγχο των φασματικών μεταβολών της ακουστικής διέγερσης κάθε ζώνης, σχεδιάστηκε μέθοδος ελέγχου με εισαγωγή στενών ορίων (κατωφλίων) για την διάκριση των στενών φασματικών μεταβολών στο πεδίο του χρόνου. Για περιπτώσεις έντονων διακυμάνσεων, έχει εισαχθεί και ένα επιπλέον υψηλό όριο, ώστε να βοηθάει τον μηχανισμό να επαναφέρει τον υπολογισμό της ενέργειας της ζώνης στην στιγμιαία τιμή της.

Η μεθοδολογία ελέγχου γίνεται σε ανάλυση ζωνών συχνοτήτων χρονικά με σύγκριση τιμών (βλπ Εξ. 4.15):

$$\begin{cases} S_{(n)} = S_{(n+1)} & S_{(n)} - a \leq S_{(n+1)} \leq S_{(n)} + a \\ S_{(n)} = S_{(n-1)} & S_{(n)} - a \geq S_{(n+1)} \geq S_{(n)} + a \end{cases} \quad (4.15)$$

Όπου  $S_{(n)}$  η εξαγόμενη στάθμη ακουστικής πίεσης κάθε ζώνης συχνοτήτων,  $S_{(n+1)}$  η στάθμη ακουστικής πίεσης σύμφωνα με την τελευταία μέτρηση,  $S_{(n-1)}$  η στάθμη ακουστικής πίεσης σύμφωνα με την προηγούμενη μέτρηση και  $a$  ο συντελεστής καθορισμού του ορίου ελέγχου για τον υπολογισμό της ακουστικής πίεσης.

Σύμφωνα με την Εξίσωση 4.15, εισάγονται δύο όρια τα οποία ισαπέχουν από την στιγμιαία ακουστική στάθμη. Η οποιαδήποτε αλλαγή στην ακουστική πίεση συντελείται μέσα σε αυτά τα όρια, ενσωματώνεται στην τελική μέτρηση της ακουστικής διέγερσης.

Στο προτεινόμενο μοντέλο, έχει σχεδιαστεί η ύπαρξη συνολικά δύο τύπων ορίων.

Στενό ζεύγος ορίων: Βοηθούν στον υπολογισμό μιας εξομαλυμένης στιγμιαίας ενέργειας στο πεδίο του χρόνου, αποτρέπει μεγάλες εναλλαγές και συνεπώς μειώνει τις πιθανές παραμορφώσεις από την επεξεργασία των τιμών στον μείκτη επαυξημένης πραγματικότητας του εικονικού ηχητικού σήματος.

Διευρυμένο ζεύγος ορίων: Σε περιπτώσεις έντονων εναλλαγών, τα διευρυμένα όρια βοηθούν στην επιστροφή της εξομαλυμένης ενέργειας πάλι στην θέση της στιγμιαίας. Αν δεν γίνει αυτό, τότε η εξομαλυμένη ενέργεια, παραμένει σε επίπεδα πολύ διαφορετικά από τη στιγμιαία και μέσα σε στενά όρια, τα οποία καθιστούν αδύνατη τη μετατόπιση της εξομαλυμένης ενέργειας στα επίπεδα της στιγμιαίας. Η ύπαρξη ενός επιπλέον ορίου μεγαλύτερου από το στενό, βοηθάει στην επιστροφή της εξομαλυμένης ενέργειας στα κανονικά επίπεδα της στιγμιαίας.

Παράδειγμα: Η μπάντα των 250 Hz έχει στιγμιαία στάθμη 55 dB SPL. Το ζεύγος ορίων έχει τεθεί στα  $\pm 3$  dB SPL, οπότε η επόμενη ενέργεια που μπορεί να δεχτεί είναι από 52 – 58 dB SPL. Στην περίπτωση έντονης βαθμιαίας αύξησης και απότομης ύφεσης υπάρχει ο κίνδυνος η στιγμιαία στάθμη να μείνει σε υψηλό επίπεδο δηλαδή για παράδειγμα στην περίπτωση που φτάσει στα 78 dB SPL τότε το μικρό όριο των  $\pm 3$  dB SPL επιτρέπει να πάρει τιμές από 75 – 81 dB SPL. Αν η πραγματική στιγμιαία τιμή βρίσκεται στο 58 dB SPL και συνεχίσει στο ίδιο επίπεδο, δεν θα καταφέρει να επανέλθει το σύστημα μέτρησης αν δεν έχει τεθεί και ένα πιο ευρύ όριο της τάξης για παράδειγμα των  $\pm 20$  dB SPL.

#### 4.5.4 Τεχνική Εκτίμησης Αμφιωτικής Ακουστότητας Σημάτων Εισόδου

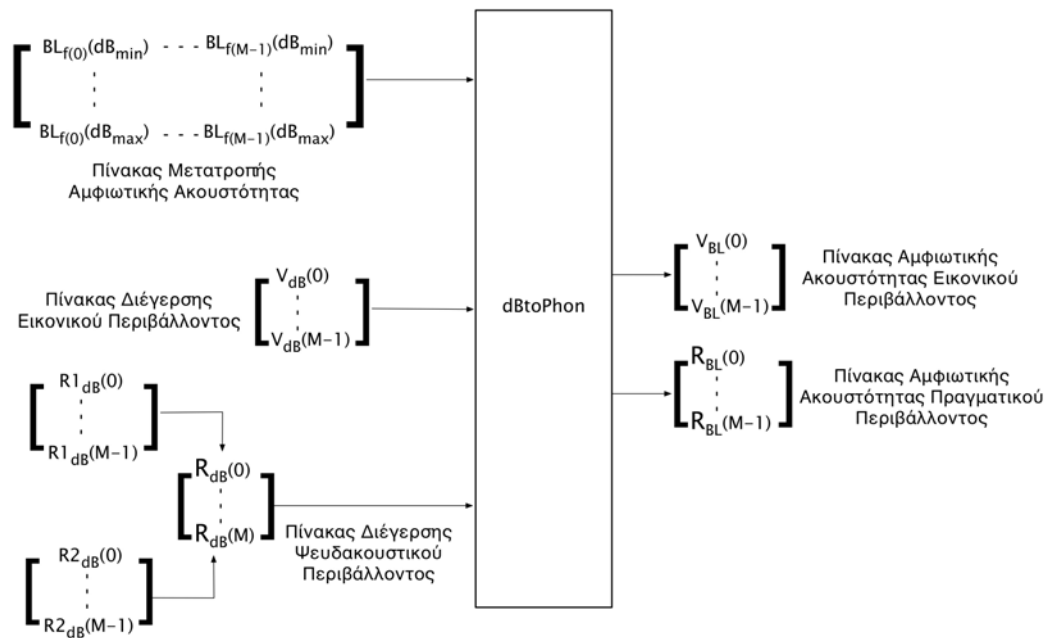
Η εκτίμηση της αμφιωτικής ακουστότητας, όπως και στα υφιστάμενα μοντέλα υπολογισμού της, προέρχεται από την χαμπύλη ακουστικής διέγερσης. Στην περίπτωση του προσαρμοστικού μοντέλου ARA μείζης, εφαρμόζεται τεχνική στην οποία χρησιμοποιείται μήτρα προϋπολογισμένων τιμών αμφιωτικής ακουστότητας για την συσχέτιση τριών μεγεθών: Συχνότητα (Hz), Στάθμη Ακουστικής Πίεσης (dB SPL), Αμφιωτική Ακουστότητα (Phon). Η ακολουθούμενη τεχνική παρουσιάζεται παρακάτω.

Αρχικά, προϋπόθεση για τη χρήση της προϋπολογισμένης μήτρας με βάση την αμφιωτική ακουστότητα που περιγράφηκε στην Ενότητα 4.4.2, είναι η βαθμονόμηση των σημάτων για τα οποία έχει υπολογιστεί η χαμπύλη ακουστικής διέγερσης, σε στάθμες ακουστικής πίεσης. Ο υπολογισμός της αμφιωτικής ακουστότητας από το διάγραμμα διέγερσης βασίζεται στην υπόθεση ότι η μέση τιμή της ηχητικής ενέργειας της κάθε ζώνης συχνοτήτων αντιπροσωπεύει

ύει την ενέργεια ολόκληρης της ζώνης συχνοτήτων. Η υπόθεση αυτή, περιέχει τον κίνδυνο υποβάθμισης της τελικής τιμής και επίδρασης ακραίων τιμών.

Στην Ενότητα 6 θα παρουσιαστούν μετρήσεις του μοντέλου σε σχέση με υφιστάμενα μοντέλα αμφιωτικής ακουστότητας και αν αυτό επηρεάζει τη λειτουργία του συνολικού μοντέλου.

Η καμπύλη διέγερσης του κάθε ηχητικού περιβάλλοντος συνοψίζεται σε έναν πίνακα με τις στάθμες ηχητικής πίεσης για την κάθε συχνοτική ζώνη. Το εικονικό περιβάλλον αντιπροσωπεύεται από τον αντίστοιχο πίνακα. Το ψευδακουστικό περιβάλλον λόγω της ιδιαιτερότητάς του να αποτελείται από δύο ηχητικά σήματα καθώς χρησιμοποιείται αμφιωτικής τεχνολογία αναπαραγωγής του πραγματικού περιβάλλοντος, υπολογίζεται ως άθροισμα των τιμών από τους αντίστοιχους πίνακες των διαγραμμάτων διέγερσης του κάθε καναλιού (βλπ Σχ. 4.9).



Σχήμα 4.9: Διάγραμμα μεθοδολογίας υπολογισμού αμφιωτικής ακουστότητας

Ο μετασχηματισμός του πίνακα διέγερσης του κάθε περιβάλλοντος σε αμφιωτική ακουστικότητα πραγματοποιείται ακολούθως. Αρχικά εισάγεται ο πίνακας μετατροπής στάθμης ακουστικής πίεσης σε αμφιωτική ακουστότητα ανά κεντρική συχνότητα. Στον πίνακα της Εξίσωσης 4.16 παρουσιάζεται η μαθηματική δομή του πίνακα συσχέτισης. Σε αυτόν όπου  $a$ , είναι η υπολογισμένη αμφιωτική ακουστότητα όπως έχει περιγραφεί στην Ενότητα 4.4.2 για την αντίστοιχη συχνοτική ζώνη με κεντρική συχνότητα  $M$  και στάθμη ηχητικής πίεσης  $E_{dB}[M]$ .

$$a_{(M,E_{dB(M)})} \in \begin{bmatrix} a_{(0,E_{dB}[\text{Min}])} \dots a_{(M-1,E_{dB}[\text{Min}])} \\ \vdots \\ a_{(0,E_{dB}[\text{Max}])} \dots a_{(M-1,E_{dB}[\text{Max}])} \end{bmatrix} \quad (4.16)$$

$$E_{BL(M)} = a_{(M,E_{dB(M)})}$$

Η τελική τιμή αμφιωτικής ακουστότητας  $E_{BL(M)}$  (βλπ Εξ. 4.16) προκύπτει από την εύρεση της αμφιωτικής ακουστότητας σύμφωνα με την συχνотική ζώνη  $M$  και της αντίστοιχης τιμής της στάθμης ηχητικής πίεσής της  $E_{dB}[M]$  (βλπ Εξ. 4.16).

Η συνάρτηση μετατροπής της ακουστικής πίεσης σε αμφιωτική ακουστότητα, εξάγει δύο πίνακες. Τον πίνακα αμφιωτικής ακουστότητας της κάθε ζώνης για το εικονικό περιβάλλον και τον πίνακα αμφιωτικής ακουστότητας για τον ψευδοακουστικό περιβάλλον. Οι δύο πίνακες αποτελούν μεταβλητές εισόδου στην συνάρτηση προσαρμογής που περιγράφεται στην επόμενη ενότητα.

## 4.6 Ανάλυση και Σύνθεση εικονικού περιβάλλοντος

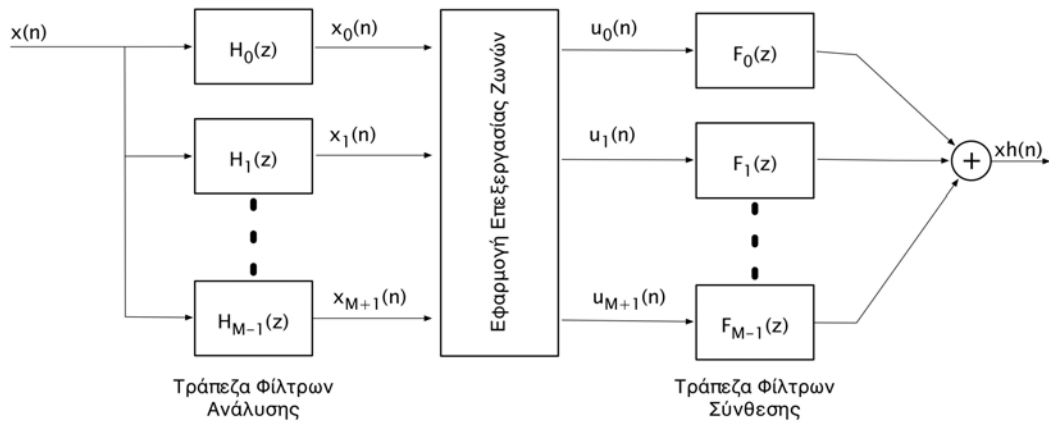
Το εικονικό περιβάλλον μέσα στο μοντέλο προσαρμοστικής ARA μείζης, είναι εκείνο στο οποίο θα πρέπει να εφαρμοστούν οι εξαγόμενοι συντελεστές του αλγόριθμου προσαρμογής. Η εφαρμογή των αλλαγών στο σήμα του εικονικού περιβάλλοντος, γίνεται σε επίπεδο συχνотικών ζωνών. Το μικρό μέγεθος του παραθύρου δεν επιτρέπει την άμεση επεξεργασία της FFT ανάλυσης και επανασύνθεσή της χωρίς την δημιουργία παραμορφώσεων στην κάθε ζώνη.

Η μέθοδος ανάλυσης του εικονικού ηχητικού σήματος στο πεδίο του χρόνου, γίνεται με τη χρήση ψηφιακών φίλτρων πεπερασμένης χρονικής απόκρισης FIR. Τα παραπάνω φίλτρα αποτελούν την τράπεζα φίλτρων ανάλυσης. Το μέγεθος των φίλτρων δεν είναι ομοιόμορφο καθώς ακολουθείται η ακουστική κλίμακα οκτάβας για τον καταμερισμό των φίλτρων.

Στο Σχήμα 4.10 φαίνεται η αρχιτεκτονική ανάλυσης και σύνθεσης του εικονικού περιβάλλοντος με τις αντίστοιχες τράπεζες φίλτρων. Όπου  $x(n)$  ορίζεται το εισερχόμενο σήμα,  $H_M(z)$  είναι το φίλτρο ανάλυσης,  $x_M(n)$  το φιλτραρισμένο σήμα που εισέρχεται τον αλγόριθμο προσαρμογής αμφιωτικής ακουστότητας,  $u_M(n)$  ορίζεται η κάθε εξερχόμενη ροή από τον αλγόριθμο προσαρμογής,  $F_M(z)$  είναι το φίλτρο σύνθεσης της κάθε ηχητικής ροής. Τέλος γίνεται η πρόσθεση των φιλτραρισμένων σημάτων και παράγεται το τελικό ηχητικό σήμα.

Για το μοντέλο προσαρμοστικής ARA μείζης, χρειάζονται τόσα φίλτρα όσες και οι ζώνες οκτάβας που επιλέχθηκαν για την ανάλυση των ηχητικών περιβαλλόντων. Πιο συγκεκριμένα, η κάθε ζώνη οκτάβας, αντιστοιχεί σε ένα ένα ψηφιακό φίλτρο. Για την κάλυψη όλου του συχνотικού εύρους χρειάζεται να ορισθούν επτά FIR φίλτρα.

Τα φίλτρα που χρησιμοποιούνται είναι τύπου ζωνοπερατού, χαμηλοδιαβατού και υψηλοπερατού. Το χαμηλοπερατό φίλτρο χρησιμοποιείται για τη ζώνη φίλτρου με κεντρική συχνότητα  $250 \text{ Hz}$ , τα ζωνοπερατά φίλτρα για τις ενδιάμεσες ζώνες με κεντρικές συχνότητες από  $500 \text{ Hz}$



Σχήμα 4.10: Αρχιτεκτονική Τράπεζας Φίλτρων

μέχρι  $8\text{ kHz}$  και το υψηλοπερατό φίλτρο στη ζώνη φίλτρου με κεντρική συχνότητα τα  $16\text{ kHz}$ . Στην περίπτωση του προσαρμοστικού μοντέλου ARA μείζης, τα φίλτρα υπολογίστηκαν να έχουν μεταξύ τους επικάλυψη κατά  $50\%$  ώστε να ελαττωθούν τυχόν έντονες παραμορφώσεις λόγω της επεξεργασίας. Η επικάλυψη παρουσιάζεται στον Πίνακα 4.4.

| Κεντρική Συχνότητα Ζώνης Φίλτρου | Ζώνη Φίλτρου                       |
|----------------------------------|------------------------------------|
| $250\text{ Hz}$                  | $0\text{ Hz} - 500\text{ Hz}$      |
| $500\text{ Hz}$                  | $250\text{ Hz} - 1000\text{ Hz}$   |
| $1000\text{ Hz}$                 | $625\text{ Hz} - 1750\text{ Hz}$   |
| $2000\text{ Hz}$                 | $1250\text{ Hz} - 3500\text{ Hz}$  |
| $4000\text{ Hz}$                 | $2500\text{ Hz} - 7000\text{ Hz}$  |
| $8000\text{ Hz}$                 | $5000\text{ Hz} - 14000\text{ Hz}$ |
| $16000\text{ Hz}$                | $8000\text{ Hz} - 24000\text{ Hz}$ |

Πίνακας 4.4: Πίνακας Ζώνης συχνοτήτων φίλτρων για Τράπεζα Φίλτρων με επικάλυψη  $50\%$ 

Για το ζωνοπερατό φίλτρο, υπολογίζεται ότι για κεντρική συχνότητα  $f(n)$  ισχύει η Εξίσωση 4.17.

$$Sb_a = f(n - 1)$$

$$Pb_a = \frac{f(n) - f(n - 1)}{2} \quad (4.17)$$

$$Sb_b = f(n + 1)$$

$$Pb_b = \frac{f(n + 1) - f(n)}{2}$$

Όπου  $Sb_a$  είναι η ελάχιστη συχνότητα αποκοπής,  $Pb_a$  η ελάχιστη συχνότητα διέλευσης,  $Sb_b$  είναι η μέγιστη συχνότητα αποκοπής, και  $Pb_b$  η μέγιστη συχνότητα διέλευσης.

Για το χαμηλοπερατό φίλτρο (βλπ Εξ. 4.18) υπολογίζεται η μεταβλητή  $Slp$  ως η συχνότητα αποκοπής και  $Plp$  η συχνότητα διέλευσης (βλπ Σχ. 4.11).

$$Slp = f(n) \quad (4.18)$$

$$Plp = \frac{f(n) - f(n - 1)}{2}$$

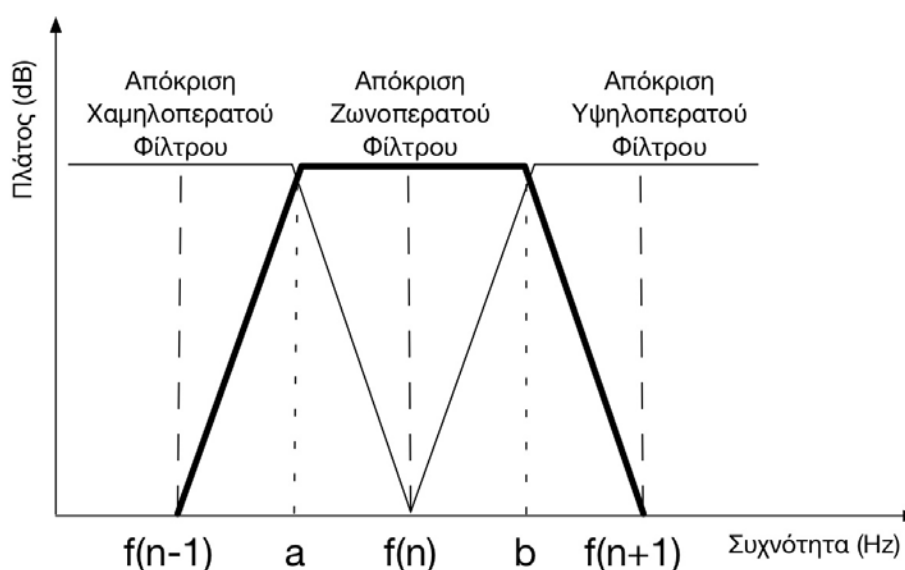
Για το υψηλοπερατό φίλτρο, υπολογίζεται ότι για κεντρική συχνότητα  $f(n - 1)$  όπως φαίνεται στο Σχήμα 4.11, ισχύει η Εξίσωση 4.19, όπου  $Shp$  είναι η συχνότητα αποκοπής και  $Php$  είναι η συχνότητα διέλευσης.

$$Shp = f(n) \quad (4.19)$$

$$Php = \frac{f(n + 1) - f(n)}{2}$$

Το ηχητικό σήμα μετατρέπεται σε επτά ηχητικά σήματα φιλτραρισμένα διατηρώντας το συχνοτικό περιεχόμενο ανάλογα με την ακουστική ζώνη που ανήκουν. Η πραγματική συχνοτική απόκριση των υλοποιημένων φίλτρων παρουσιάζεται στο Σχήμα 5.11 της Ενότητας 5.2.4. Στο κάθε φιλτραρισμένο σήμα, θα εφαρμοστεί ο κατάλληλος συντελεστής που θα εξαχθεί από τον αλγόριθμο προσαρμογής οποίος περιγράφεται στην επόμενη ενότητα.

Η σύνθεση του εικονικού ηχητικού σήματος διέρχεται μέσα από το φιλτράρισμα των εξαγόμενων ηχητικών σημάτων από τον αλγόριθμο προσαρμογής αμφιωτικής ακουστότητας. Για το φιλτράρισμα, σχεδιάστηκε η τράπεζα φίλτρων σύνθεσης, η οποία εξαλείφει τις παραμορφώσεις στα όρια των συχνοτικών ζωνών. Τα φίλτρα που χρησιμοποιούνται είναι όμοια με της τράπεζας φίλτρων ανάλυσης (βλπ Πίν. 4.2) και έχουν στόχο τη μείωση παραμορφώσεων του πλάτους και της φάσης καθώς και των παραμορφώσεων λόγω αναδίπλωσης συχνοτήτων. Τα εξερχόμενα φιλτραρισμένα ηχητικά σήματα αθροίζονται για να παραχθεί το τελικό εικονικό περιβάλλον, το οποίο είναι προσαρμοσμένο στη μείξη με το ψευδοακουστικό περιβάλλον.



Σχήμα 4.11: Συχνотική Απόκριση Τράπεζας Φίλτρων

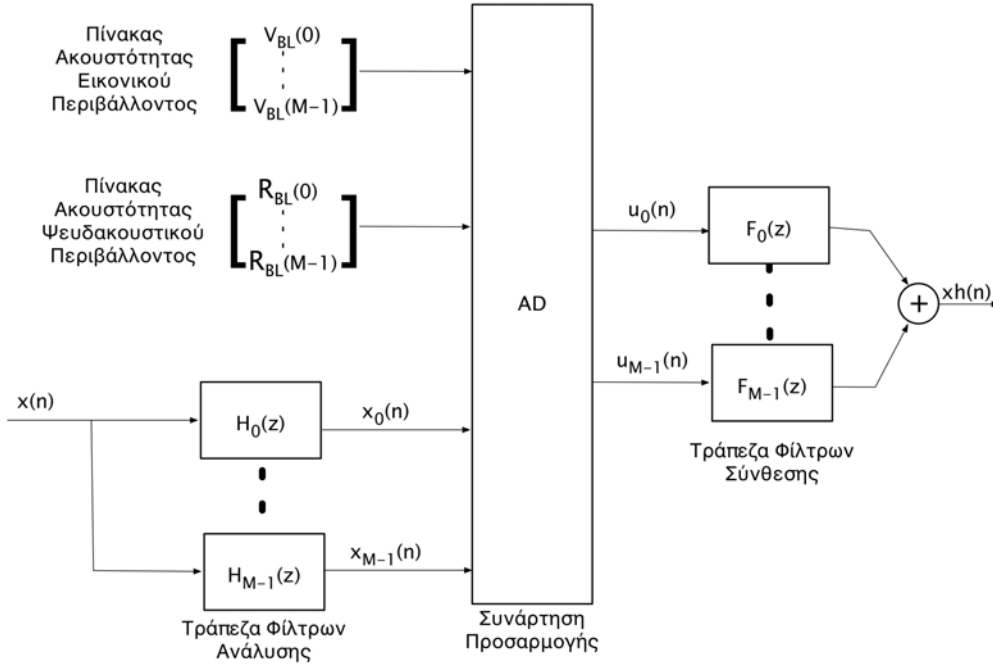
## 4.7 Αλγόριθμος προσαρμογής περιβαλλόντων τύπου ARA

Ο πυρήνας του μοντέλου προσαρμοστικής ARA μείξης, αποτελείται από τον αλγόριθμο προσαρμογής της μείξης με όρους αμφιωτικής ακουστότητας. Η μεθοδολογία της προσαρμοστικής μείξης του επαυξημένου περιβάλλοντος αναλύεται στην Ενότητα 4.7.1. Η τεχνική εκτίμηση της αμφιωτικής ακουστότητας από τον πίνακα διέγερσης ζωνών των σημάτων εισόδου του αλγόριθμου προσαρμογής παρουσιάζεται στην Ενότητα 4.5.4. Μέσα από την ανάλυση του αλγόριθμου αναδεικνύονται οι σκοποί της προσαρμοστικής μείξης στα περιβάλλοντα τύπου ARA καθώς και ειδικά ζητήματα που προκύπτουν σε σχέση με την μεθοδολογία που ακολουθείται.

### 4.7.1 Αρχιτεκτονική αλγορίθμου προσαρμογής Αμφιωτικής Ακουστότητας

Σε αυτή την ενότητα θα παρουσιαστεί η μεθοδολογία της λειτουργίας του αλγορίθμου προσαρμογής αμφιωτικής ακουστότητας. Το πρόβλημα που καλείται να επιλύσει ο αλγόριθμος προσαρμογής είναι η συσχέτιση του εικονικού περιβάλλοντος ήχου με το ψευδοακουστικό. Η συσχέτιση, αποτελεί επεξεργασία στο πεδίο των συχνотικών ζωνών του κάθε ηχητικού περιβάλλοντος. Πιο αναλυτικά, για την επαυξημένη μείξη ήχου, και την εμπύθιση του ακροατή, ζητούμενο είναι η ανάμειξη των ηχητικών περιβαλλόντων. Το εικονικό πρέπει να αποτελεί μέρος του πραγματικού. Η επεξεργασία στο πεδίο των συχνотικών ζωνών στοχεύει στο να προσδώσει συχνотικά χαρακτηριστικά όπως είναι η χροιά ή το χρώμα του πραγματικού

περιβάλλοντος στο εικονικό χωρίς όμως να το αλλοιώσει.



Σχήμα 4.12: Διάγραμμα Εισόδων-Εξόδων Συνάρτησης Προσαρμογής

Το πρώτο στάδιο της μεθοδολογίας προσαρμογής είναι η εύρεση της ακουστικής διαφοράς σε επίπεδο ζώνης συχνοτήτων. Στη συνάρτηση προσαρμογής, εισάγεται ο πίνακας αμφιωτικής ακουστότητας των συχνοτικών ζωνών του εικονικού περιβάλλοντος και του πραγματικού. Από τα δύο περιβάλλοντα, βρίσκεται η μεταξύ τους διαφορά στην κάθε ζώνη. Στην Εξίσωση 4.20, ορίζεται ως  $V_{dif}(M)$  η διαφορά της συχνοτικής ζώνης  $M$  του ψευδοακουστικού περιβάλλοντος  $R_{BL}(M)$  από το εικονικό  $V_{BL}(M)$ .

$$V_{dif}(M) = R_{BL}(M) - V_{BL}(M) \quad (4.20)$$

Η διαφορά αυτή φαίνεται ικανή να εφαρμοστεί στην εξισορρόπηση του εικονικού περιβάλλοντος με το πραγματικό. Ωστόσο, αυτό μπορεί να δημιουργήσει προβλήματα αλλοίωσης του εικονικού περιβάλλοντος ανάλογα με το μέγεθος της διαφοράς μεταξύ των περιβαλλόντων. Για παράδειγμα, αν η ζώνη των  $1000 \text{ Hz}$  έχει αμφιωτική ακουστότητα της τάξης των  $50 \text{ Phons}$  για το εικονικό περιβάλλον και  $70 \text{ Phons}$  για το πραγματικό, η διαφορά μεταξύ τους είναι  $20 \text{ Phons}$ . Η προσθήκη  $20 \text{ Phons}$  θα δημιουργήσει μεγάλη αύξηση στο εικονικό περιβάλλον, αλλοιώνοντας πιθανόν σε μεγάλο βαθμό το περιεχόμενό του, παραμορφώνοντας το συχνοτικό περιεχόμενο της συγκεκριμένης ζώνης και δημιουργώντας δυσάρεστη αίσθηση για το συνολικό εικονικό περιβάλλον για τον ακροατή. Γι' αυτό το λόγο, εισάγεται ο συντελεστής εξομάλυνσης της παραπάνω διαφοράς λαμβάνοντας υπόψιν την αμφιωτική ακουστότητα από

όλες τις ζώνες συχνοτήτων της εικονικής ηχητικής πηγής.

Παρακάτω δίνεται η μεθοδολογία υπολογισμού του συντελεστή εξομάλυνσης της διαφοράς σε σχέση με τη συνολική αμφιωτική ακουστότητα του εικονικού περιβάλλοντος.

Αρχικά υπολογίζεται η μέση τιμή από τις ακουστότητες των συχνοτικών ζωνών της εικονικής ηχητικής πηγής (βλπ Εξ. 4.21), όπου  $AVG_V$  είναι η μέση τιμή,  $M$  ο αριθμός των ζωνών και  $V_{BL}(M)$  η αμφιωτική ακουστότητα κάθε συχνοτικής ζώνης  $M$ .

$$AVG_V = \frac{\sum_0^{M-1} V_{BL}(M)}{M} \quad (4.21)$$

Η παραπάνω μέση τιμή χρησιμοποιείται για τον υπολογισμό του λόγου αμφιωτικής ακουστότητας της κάθε ζώνης προς τη μέση τιμή αμφιωτικής ακουστότητας από όλες τις ζώνες (βλπ Εξ. 4.22).

$$V_{Ratio}(M) = \frac{V_{BL}(M)}{AVG_V} \quad (4.22)$$

Όπου  $V_{Ratio}(M)$  είναι ο λόγος της αμφιωτικής ακουστότητας  $V_{BL}(M)$  για τη ζώνη  $M$  και  $AVG_V$  η μέση τιμή αμφιωτικής ακουστότητας του συνόλου των συχνοτικών ζωνών. Με τον υπολογισμό του παραπάνω λόγου, σταθμίζεται η διαφορά της κάθε ζώνης  $V_{dif}(M)$  σύμφωνα με τη μέση τιμή ακουστότητας όλων των ζωνών. Έτσι λαμβάνεται υπόψιν η συνολική αμφιωτική ακουστότητα της εικονικής ηχητικής πηγής.

$$V'_{dif}(M) = V_{dif}(M) \cdot V_{Ratio}(M) \quad (4.23)$$

Η τελική τιμή της διαφοράς  $V'_{dif}(M)$  προστίθεται στην αρχική αμφιωτική ακουστότητα της ζώνης  $V_{BL}(M)$  και παράγεται η νέα αμφιωτική ακουστότητα  $V'_{BL}(M)$  του εικονικού σήματος για τη συγκεκριμένη ζώνη:

$$V'_{BL}(M) = V_{BL}(M) + V'_{dif}(M) \quad (4.24)$$

Ο αλγόριθμος προσαρμογής, λαμβάνει ως είσοδο τις ηχητικές ροές της εικονικής πηγής που διέρχονται από την τράπεζα φίλτρων ανάλυσης, η οποία αναλύει το εικονικό σήμα σε αντίστοιχο αριθμό ροών ώστε να υπάρχει μια αντιπροσωπευτική για κάθε συχνοτική ζώνη. Στην κάθε ηχητική ροή θα εφαρμοστεί η αντίστοιχη σταθμισμένη διαφορά  $V'_{dif}(M)$  όπως υπολογίστηκε παραπάνω.

Για την μετατροπή της διαφοράς σε λόγο, χρησιμοποιείται η Εξίσωση 4.25 όπου  $V'_{dif}(M)$  είναι η τελική διαφορά αμφιωτικής ακουστότητας στην ζώνη  $M$  και  $BL_{difRatio}$  είναι ο λόγος ο οποίος θα πρέπει να εφαρμοστεί στις εισαγόμενες ηχητικές ροές:

$$BL_{difRatio} = 2^{\frac{V'_{dif}(M)}{10}} \quad (4.25)$$

Στην Εξίσωση 4.26 ο λόγος της διαφοράς  $BL_{difRatio}$  εφαρμόζεται στο αντίστοιχο φιλτραρισμένο σήμα  $x_M(n)$  και εξάγεται η τελική ηχητική ροή  $u_M(n)$ :

$$u_M(n) = x_M(n) \cdot BL_{difRatio}(M) \quad (4.26)$$

Οι εξαγόμενες ηχητικές ροές διέρχονται από την τράπεζα φίλτρων σύνθεσης, η οποία έχει σκοπό την απαλοιφή τυχόν παραμορφώσεων εκτός της συχνοτικής ζώνης λόγω της επεξεργασίας και την σύνθεση του τελικού σήματος. Στην Εξίσωση 4.27, το φίλτρο της τράπεζας φίλτρων  $F_M(n)$  εφαρμόζεται στο εξερχόμενο σήμα από τον αλγόριθμο προσαρμογής  $U_M(n)$  για τη συχνοτική ζώνη  $M$ . Το άθροισμα όλων των φιλτραρισμένων ροών συνθέτουν το τελικό σήμα της εικονικής πηγής ήχου όπως απεικονίζεται στο Σχήμα 4.12).

$$xh(n) = \sum_0^{M-1} U_M(n) \cdot F_M(n) \quad (4.27)$$

Η προσαρμοστικότητα στην ARA μείζη, προσβλέπει στην αυτοματοποιημένη μέθοδο μείζης του εικονικού στοιχείου με το πραγματικό. Το μοντέλο δεν απαιτεί χειρισμό από τον ακροατή. Παρ' όλα αυτά, υπάρχει προαιρετικά η μεθοδολογία εισαγωγής μιας παραμέτρου διαβάθμισης της επαυξημένης μείζης η οποία μπορεί να ρυθμιστεί από τον ακροατή.

Για τον παραπάνω λόγο, εισάγεται η παράμετρος ως σταθερά  $c$  (βλπ Εξ. 4.28) κατά τον υπολογισμό της διαφοράς μεταξύ του εικονικού περιβάλλοντος  $R_{BL}(M)$  και του ψευδοακουστικού  $(M) - V_{BL}(M)$ .

$$V_{dif}(M) = R_{BL}(M) - V_{BL}(M) + c \quad (4.28)$$

Το τελικό ηχητικό σήμα του εικονικού περιβάλλοντος  $xh(n)$ , μετά από τις διεργασίες προσαρμογής είναι έτοιμο να αναπαραχθεί στον τρισδιάστατο επαυξημένο χώρο μέσα από την επιθυμητή τεχνική χωρικής τοποθέτησής του. Ουσιαστικά, ο αλγόριθμος προσαρμογής λειτουργεί ως μια προεπεξεργασία του εικονικού περιβάλλοντος λαμβάνοντας χαρακτηριστικά του πραγματικού περιβάλλοντος στο συχνοτικό πεδίο. Η τοποθέτηση της συνάρτησης προσαρμογής στο μοντέλο μετά από τη διαδικασία της χωρικής τοποθέτησής της εικονικής πηγής στον τρισδιάστατο εικονικό ηχητικό χώρο, θα αλλοίωνε τα φασματικά χαρακτηριστικά που δημιουργεί η μέθοδος χωρικής τοποθέτησης, όπως είναι η τεχνική ενσωμάτωση χρουστικών αποκρίσεων τύπου HRIR και θα δημιουργούσε σύγχυση στον ακροατή.

## 4.8 Σύνοψη Κεφαλαίου

Σε αυτό το κεφάλαιο δόθηκε η μεθοδολογία για την ενσωμάτωση της προσαρμοστικότητας στην ARA μείζη. Αρχικά παρουσιάστηκε το γενικευμένο μοντέλο προσαρμοστικής ARA μείζης, αποτυπώνοντας και περιγράφοντας τα επίπεδα του προτεινόμενου μοντέλου. Κατόπιν, παρουσιάστηκε η τεχνική επεξεργασίας των αμφιωτικών σημάτων εισόδου του ψευδοακουστικού περιβάλλοντος μέσα από τα ιδιαίτερα χαρακτηριστικά του αισθητήριου της ακοής. Στη συνέχεια, αναλύθηκε η μεθοδολογία εισαγωγής δεδομένων ακουστικής αντίληψης μέσα από μια επισκόπηση των υφιστάμενων μοντέλων πρόβλεψης αμφιωτικής ακουστότητας, εστιάζοντας στο μοντέλο του Zwicker και Glasberg - Moore. Η μέθοδος δημιουργίας κανόνων μείζης για το προτεινόμενο μοντέλο δίνεται μέσα από την παρουσίαση του προϋπολογισμένου πίνακα μετατροπής στον οποίο συνδέονται η αμφιωτική ακουστότητα με τη στάθμη ακουστικής πίεσης για κάθε επιλεγμένη ζώνη συχνότητων. Επίσης, δίνεται έμφαση στη μέθοδο υπολογισμού της

διέγερσης των σημάτων εισόδου μέσα από τεχνική βαθμονόμησης τους, και τη μέθοδο υπολογισμού της ενέργειας σύμφωνα με την φασματική ανάλυση τύπου FFT που έχει επιλεγεί. Επιπροσθέτως, επισημαίνεται η μέθοδος αντιμετώπισης παραμορφώσεων λόγω της διεργασίας του υπολογισμού ενέργειας του κάθε περιβάλλοντος και η τεχνική ενσωμάτωσης των ηχητικών μετρήσεων θέτοντας αντίστοιχα ηχητικά όρια. Για την εκτίμηση της αμφιωτικής ακουστότητας, υλοποιείται η μεθοδολογία μετασχηματισμού του πίνακα διέγερσης σε πίνακα αμφιωτικής ακουστότητας με την χρήση της προϋπολογισμένης μήτρας αμφιωτικής ακουστότητας. Για την ανάλυση και επανασύνθεση του εικονικού ηχητικού περιβάλλοντος, σχεδιάστηκε η αντίστοιχη τράπεζα φίλτρων με φίλτρα που αντιστοιχούν στην εκάστοτε συχνοτική ζώνη. Στην τελευταία ενότητα, παρουσιάστηκε η μεθοδολογία λειτουργίας του αλγόριθμου προσαρμογής αμφιωτικής ακουστότητας. Σε αυτήν, αναλύεται η αρχιτεκτονική του αλγόριθμου προσαρμογής μέσα από τη μεθοδολογία υπολογισμού του τελικού εικονικού σήματος ήχου. Λαμβάνοντας υπόψη τη συνολική μεθοδολογία, στο επόμενο κεφάλαιο παρουσιάζεται η πρακτική υλοποίηση του μοντέλου προσαρμοστικής επαυξημένης μείξης ήχου.



## Κεφάλαιο 5

# Υλοποίηση Προσαρμοστικής ARA Μείξης Πραγματικού Χρόνου

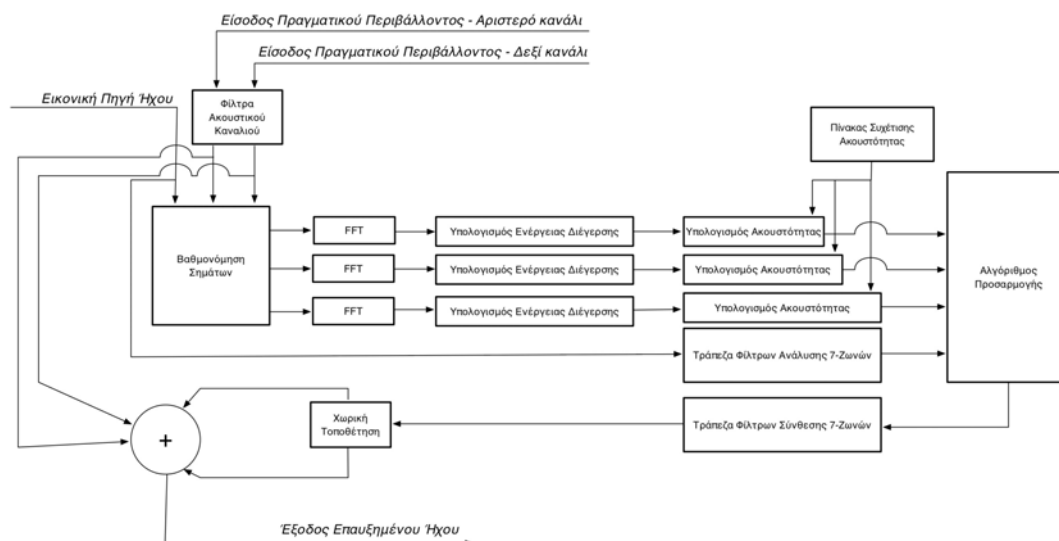
### 5.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζεται η πειραματική υλοποίηση του μοντέλου προσαρμοστικής ARA μείξης σε συνθήκη πραγματικού χρόνου. Η υλοποίηση, έδωσε τη δυνατότητα για τη μετέπειτα εισαγωγή του σε ένα πρωτότυπο ηχητικό παιχνίδι επαυξημένης πραγματικότητας καθώς και την πειραματική αξιολόγησή του. Για να γίνει το παραπάνω εφικτό, υλοποιήθηκε η μεθοδολογία που περιγράφεται στο προηγούμενο κεφάλαιο. Η παράμετρος της υλοποίησης σε χρόνο, ήταν κρίσιμη καθώς το περιβάλλον επαυξημένης πραγματικότητας είναι συνεχώς μεταβαλλόμενο στο πεδίο του χρόνου, οπότε στην διαδικασία αξιολόγησης, η λειτουργία του μοντέλου σε πραγματικό χρόνο βοήθησε στην άμεση ρύθμιση παραμέτρων του, και στην κατανόηση της επίδρασης της πρωτότυπης εφαρμογής προσαρμοστικής μείξης σε σενάριο επαυξημένης πραγματικότητας ήχου.

Σε πρώτη φάση, στο παρών κεφάλαιο δίνεται η υλοποίηση της αμφιωτικής αναπαραγωγής του πραγματικού περιβάλλοντος προς τον ακροατή και η εισαγωγή του στο μοντέλο. Ακολούθως αναλύεται η υλοποίηση των διεργασιών βαθμονόμησης των σημάτων εισόδου στο μοντέλο. Παράλληλα, παρουσιάζεται η υλοποίηση της μεθόδου υπολογισμού της ακουστικής διέγερσης των σημάτων εισόδου, καθώς και η ανάλυση και επανασύνθεση του εικονικού σήματος ήχου. Επίσης, παρουσιάζεται η υλοποίηση της μήτρας μετατροπής αμφιωτικής ακουστικότητας και η υλοποίηση της μεθόδου εκτίμησης της αμφιωτικής ακουστικότητας. Τέλος, περιγράφεται η υλοποίηση του αλγόριθμου προσαρμογής και ο παράγοντας διαβάθμισης της προσαρμοστικής ARA μείξης.

## 5.2 Πειραματική Υλοποίηση

Η πειραματική υλοποίηση του μοντέλου σχεδιάστηκε σύμφωνα με τη μεθοδολογία που περιγράφηκε στο προηγούμενο κεφάλαιο. Για την υλοποίηση λήφθηκαν υπόψη όλες οι παράμετροι που συναντώνται σε περιβάλλοντα επαυξημένης πραγματικότητας ήχου. Πιο αναλυτικά, στο Σχήμα 5.1 δίνεται το διάγραμμα αρχιτεκτονικής του συνολικού μοντέλου προσαρμοστικής ARA μείξης, το οποίο ακολουθήθηκε σε όλη τη διαδικασία υλοποίησης.



Σχήμα 5.1: Αρχιτεκτονική μοντέλου προσαρμοστικής μείξης τύπου ARA

Σύμφωνα με το διάγραμμα, τα επιμέρους δομικά μέρη του μοντέλου είναι διακριτά σύμφωνα με το ρόλο τους. Αρχικά, τα ηχητικά σήματα του πραγματικού και του εικονικού περιβάλλοντος εισάγονται στο μοντέλο. Τα σήματα του πραγματικού περιβάλλοντος διέρχονται από τη μονάδα αντιστάθμισης της επίδρασης του ακουστικού καναλιού και μετατρέπονται σε σήματα του ψευδοακουστικού περιβάλλοντος. Η εικονική πηγή ήχου δεν ακολουθεί την ίδια διαδικασία, καθώς αποτελεί ανεξάρτητη πηγή ήχου από το πραγματικό ηχητικό περιβάλλον.

Ο συνολικός αριθμός των εισερχόμενων σημάτων είναι τρία. Τα δύο αντιπροσωπεύουν τις αμφωτικές συνιστώσες του πραγματικού περιβάλλοντος ήχου και το τρίτο σήμα αντιστοιχεί στην εικονική πηγή ήχου. Όλα τα σήματα στη συνέχεια διέρχονται από το υποσύστημα βαθμονόμησης. Ακολούθως, τα σήματα αναλύονται με τη χρήση του γρήγορου μετασχηματισμού Fourier και κατόπιν υπολογίζεται η καμπύλη ακουστικής διέγερσης. Στη συνέχεια, οι μετρήσεις εισέρχονται στο σύστημα υπολογισμού αμφωτικής ακουστότητας μέσα στο οποίο ενσωματώνονται και οι τιμές του πίνακα μετατροπής αμφωτικής ακουστότητας.

Οι τελικές εκτιμήσεις αμφωτικής ακουστότητας εισάγονται στον αλγόριθμο προσαρμογής. Επιπροσθέτως, το ηχητικό σήμα της εικονικής πηγής, διέρχεται από την τράπεζα φίλτρων ανάλυσης μέσω της οποίας αναλύεται σε επτά συχνοτικές ζώνες. Το αποτέλεσμα του αλγόριθμου προσαρμογής εφαρμόζεται στο αναλυμένο εικονικό σήμα και ακολούθως εξέρχεται από

τον αλγόριθμο προσαρμογής και εισάγεται στην αντίστοιχη τράπεζα φίλτρων σύνθεσης επτά ζωνών. Κατόπιν, το εικονικό σήμα είναι διαθέσιμο για την οποιαδήποτε χωρική τοποθέτηση στον τρισδιάστατο εικονικό ηχητικό χώρο. Τέλος, το ψευδοακουστικό σήμα αναμειγνύεται με το εικονικό σήμα και παράγεται το προσαρμοσμένο επαυξημένο περιβάλλον ήχου.

Η πειραματική υλοποίηση, πραγματοποιείται στο περιβάλλον εξομοίωσης Simulink [28] του προγραμματιστικού περιβάλλοντος της Matlab [79]. Η επεξεργασία των δειγμάτων κατά την εξομοίωση του μοντέλου, γίνεται με τη μέθοδο των παραθύρων. Για την εξομοίωση αρχικά ορίστηκε η λειτουργία του μοντέλου με ρυθμό δειγματοληψίας ήχου στα 48 kHz. Το μέγεθος του παραθύρου επιλέχθηκε να είναι τα 1024 δείγματα καθώς αυτό το μέγεθος παρέχει 512 διαθέσιμα σημεία (bins) τα οποία επαρκούν για την ανάλυση των κρίσιμων συχνοτήτων, με τα επιλεγμένα εύρη ζωνών τους, σύμφωνα με τη μεθοδολογία υπολογισμού ακουστικής διέγερσης που αναλύθηκε στο προηγούμενο κεφάλαιο.

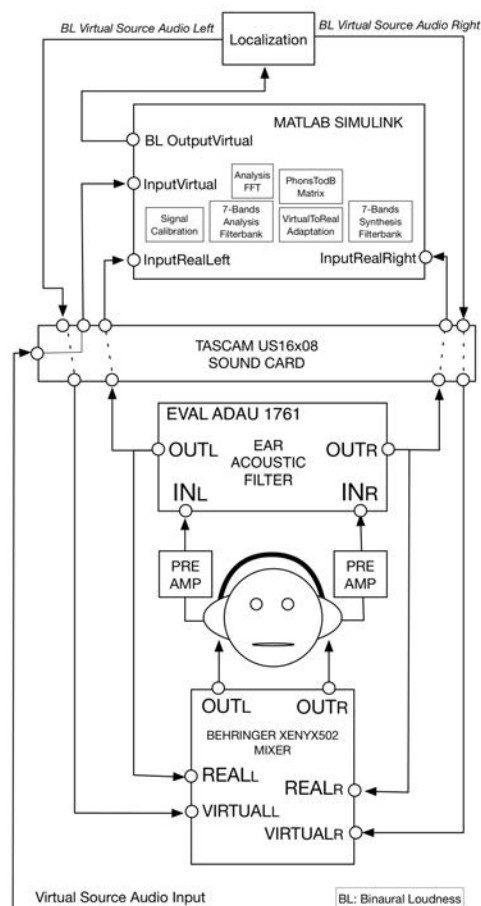
Εκτός από το προγραμματιστικό περιβάλλον εξομοίωσης, μια εξωτερική συστοιχία μονάδων επεξεργασίας και δρομολόγησης σημάτων χρησιμοποιήθηκε (βλπ Σχ. 5.2). Συγκεκριμένα, χρησιμοποιήθηκε μια κάρτα ήχου τύπου Tascam US16x08 για τη δρομολόγηση των σημάτων στο πρόγραμμα εξομοίωσης. Για την αμφιωτική εγγραφή του ψευδοακουστικού περιβάλλοντος χρησιμοποιήθηκε μια μονάδα ψηφιακής επεξεργασίας ήχου με χαμηλή καθυστέρηση καθώς και ένας μείκτης για την δρομολόγηση του επαυξημένου ήχου. Επιπλέον χρησιμοποιήθηκε ένα ζεύγος ακουστικών τύπου In-Ear κατάλληλα για χρήση σε περιβάλλοντα τύπου ARA.

Τα παραπάνω θα παρουσιαστούν αναλυτικά μέσα από τις επόμενες ενότητες του κεφαλαίου. Πιο ειδικά στη συνέχεια, αρχικά δίνεται η περιγραφή της υλοποίησης της αμφιωτικής αναπαραγωγής του πραγματικού περιβάλλοντος. Κατόπιν, περιγράφονται οι διεργασίες βαθμολόγησης των σημάτων εισόδου. Ακολούθως παρουσιάζεται η μέτρηση της ενέργειας διέγερσης των σημάτων εισόδου. Στην Ενότητα 5.2.4 δίνεται η υλοποίηση της ανάλυσης και σύνθεσης του εικονικού περιβάλλοντος. Στις επόμενες ενότητες περιγράφεται η υλοποίηση του πίνακα μετατροπής αμφιωτικής ακουστότητας και η εκτίμηση της αμφιωτικής ακουστότητας των σημάτων εισόδου. Τέλος δίνεται αναλυτικά η υλοποίηση του αλγόριθμου προσαρμοστικής ARA μείξης και η εισαγωγή της διαβάθμισης στην προσαρμοστική ARA μείξη.

### 5.2.1 Υλοποίηση Αμφιωτικής Αναπαραγωγής Πραγματικού Περιβάλλοντος

Το πραγματικό περιβάλλον ήχου αποτελεί τη μία εκ των δύο συνιστωσών του επαυξημένου περιβάλλοντος ήχου. Η λήψη του και η αναπαραγωγή του ως ψευδοακουστικό περιβάλλον αποτελεί κοινό στοιχείο στο πεδίο της επαυξημένης πραγματικότητας ήχου. Η υλοποίηση του συστήματος εγγραφής και αναπαραγωγής του πραγματικού περιβάλλοντος ως ψευδοακουστικό θα αναλυθεί στο παρόν υποκεφάλαιο. Για την υλοποίηση, ακολουθείται η μεθοδολογία που αναλύεται στην Ενότητα 4.3.

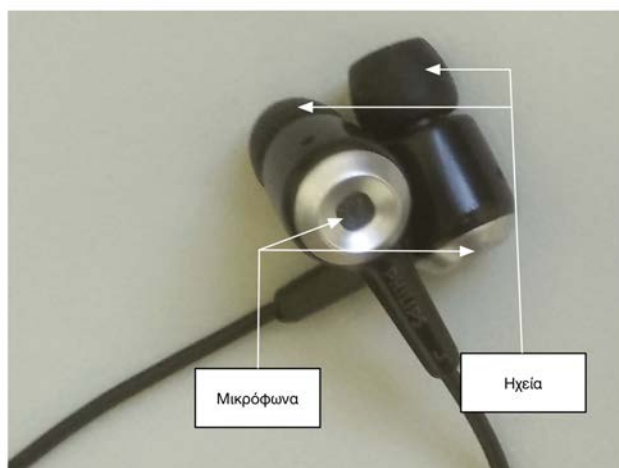
Η βάση της ανάπτυξης του συστήματος αναπαραγωγής του πραγματικού περιβάλλοντος είναι η χρήση της αμφιωτικής τεχνολογίας. Σύμφωνα με αυτή την τεχνολογία, για την εγγραφή και αναπαραγωγή του ήχου, απαιτείται η τοποθέτηση μικροφώνων στην είσοδο του αυτιού στην



Σχήμα 5.2: Επισκόπηση Υλοποίησης μοντέλου προσαρμοστικής μείξης τύπου ARA

αρχή του ακουστικού πόρου, για να είναι δυνατόν στο ληφθέν σήμα να ενσωματώνεται κατά την εγγραφή του πραγματικού περιβάλλοντος, όσο το δυνατόν τη συνολικότερη επίδραση του ανθρώπινου σώματος και του έξω αυτιού. Το εισερχόμενο σήμα αναπαράγεται από ακουστικά τα οποία βρίσκονται σχεδόν στην ίδια θέση με εκείνη των μικροφώνων. Για την παρούσα υλοποίηση, χρησιμοποιήθηκε ένα ζεύγος ακουστικών με σύστημα ακύρωσης θορύβου το οποίο ήταν απενεργοποιημένο (βλπ Σχ. 5.3).

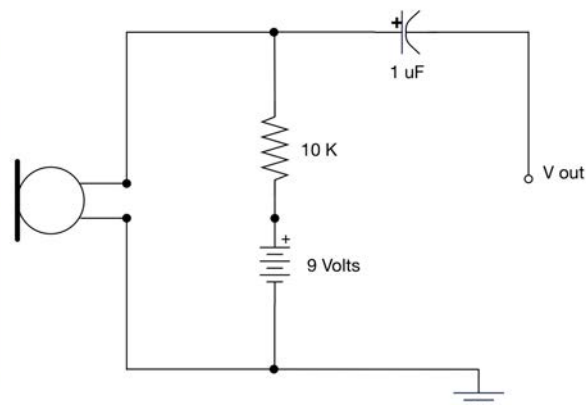
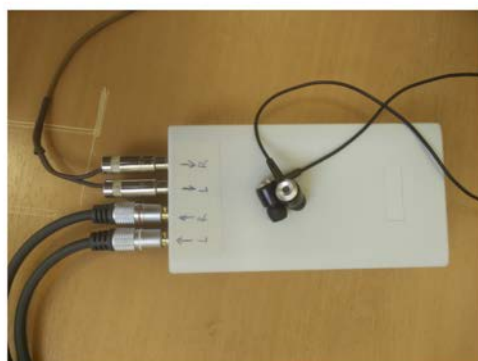
Σε αυτό τον τύπο ακουστικών, υπάρχει ένα μικρόφωνο στην εξωτερική πλευρά κάθε ακουστικού το οποίο καταγράφει τις ανακλάσεις του ανθρώπινου σώματος καθώς και την ακουστική επίδραση του εξωτερικού αυτιού. Επιπροσθέτως, στην εσωτερική πλευρά του κάθε ακουστικού, υπάρχει το αντίστοιχο ηχείο του. Επιλέχθηκε ο συγκεκριμένος τύπος ακουστικών και όχι για παράδειγμα ακουστικά τύπου Closed-back, λόγω της υψηλότερης πιστότητας αναπαραγωγής του πραγματικού περιβάλλοντος που μπορεί να επιτευχθεί χωρίς για παράδειγμα επιπλέον διεργασίες επανασύστασης της επίδρασης του εξωτερικού αυτιού. Επίσης, άλλο ένα στοιχείο είναι η μηχανική απόσβεση που προσφέρουν στον ήχο η οποία ισοδυναμεί με ένα χαμηλοδιαβατό φίλτρο το οποίο εισάγει μια μικρή φυσική καθυστέρηση στη μετάδοση. Αυτή η καθυστέρηση είναι πολύ χρήσιμη ώστε να είναι εφικτή η επεξεργασία του ήχου από τα



Σχήμα 5.3: Ακουστικά Αμφιωτικής Αναπαραγωγής - Philips SHN2500

μικρόφωνα πριν αναπαραχθεί στα αυτιά του ακροατή [102].

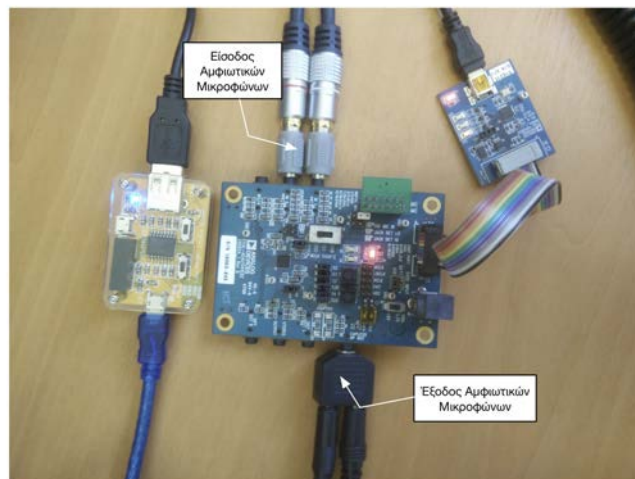
Το εισερχόμενο σήμα των μικροφώνων, λόγω του τύπου τους (Electret Condenser Mic), χρειάζεται να προενισχυθεί. Για το λόγο αυτό, υλοποιήθηκε μια μονάδα προενίσχυσης για το κάθε μικρόφωνο σύμφωνα με το Σχήμα 5.4. Η προενίσχυση, ενισχύει την έξοδο του κάθε μικροφώνου ώστε να έχει στην απαιτούμενη τάση εισόδου στο σύστημα.



Σχήμα 5.4: Μονάδα Προενίσχυσης Μικροφώνων Αμφιωτικής λήψης

Όπως προαναφέρθηκε, η θέση των μικροφώνων, βρίσκεται στην είσοδο του ακουστικού καναλιού. Για την εξάλειψη των ακουστικών φαινομένων λόγω του μπλοκαρισμένου ακουστικού πόρου, που οφείλονται στην χρήση των In-Ear ακουστικών, υλοποιείται κατάλληλο φιλτράρισμα. Η επεξεργασία του φιλτραρίσματος, υλοποιείται με την χρήση εξωτερικής μονάδας ψηφιακής επεξεργασίας ήχου χαμηλής καθυστέρησης όπως φαίνεται στο Σχήμα 5.5. Σε αυτή την μονάδα, εισέρχεται το προενισχυμένο σήμα του κάθε μικροφώνου. Η έξοδος της μονάδας στέλνει το φιλτραρισμένο σήμα των μικροφώνων στην είσοδο της κάρτας ήχου

που χρησιμοποιείται ως ηχητική διεπαφή του μοντέλου προσαρμογής και στην είσοδο του με-  
 ίκτη που χρησιμοποιείται για την δρομολόγηση του σήματος κατευθείαν στα ακουστικά (βλπ  
 Σχ. 5.2). Το σήμα αυτό αποτελεί το ψευδοακουστικό περιβάλλον του μοντέλου επαυξημένης  
 πραγματικότητας ήχου.



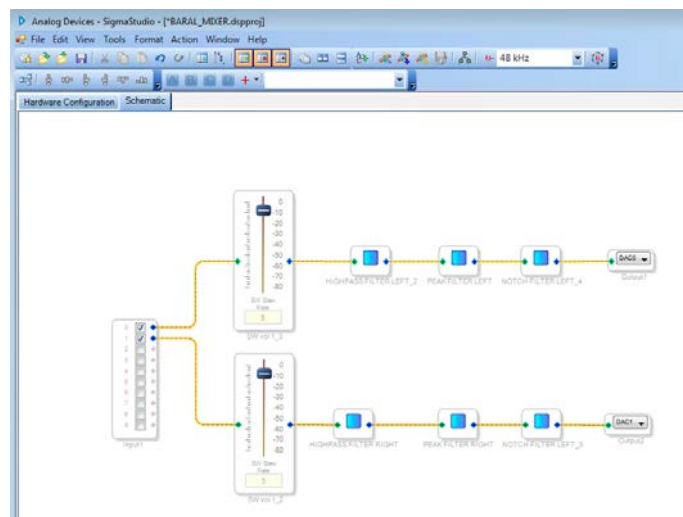
Σχήμα 5.5: Μονάδα Ψηφιακής επεξεργασίας Ήχου Αμφιωτικής λήψης «Analog Devices EVAL ADAU1761»

Το πραγματικό περιβάλλον ήχου για να αναπαραχθεί ως ψευδοακουστικό, φιλτράρεται κατάλληλα ώστε να μετριασθεί η επίδραση του κλειστού ακουστικού πόρου. Για την υλοποίηση του φιλτραρίσματος, χρησιμοποιήθηκε το προγραμματιστικό περιβάλλον του Sigma Studio (βλπ Σχ. 5.6) [21]. Μέσα από το Sigma Studio, τα υλοποιημένα φίλτρα, εισάγονται στην μονάδα ψηφιακής επεξεργασίας. Επιπροσθέτως, για την εξασθένιση του θορύβου που εισάγει η σύνδεση της μονάδας ψηφιακής επεξεργασίας ήχου με τον υπολογιστή μέσω σύνδεσης USB, παρεμβάλλεται μια μονάδα εξασθένισης θορύβου τύπου Analogue Devices ADUM4160 USB to USB Isolator (βλπ Σχ. 5.5), για την εξάλειψη του θορύβου τάσης κατά τη λειτουργία της.

Αναφορικά με τα φίλτρα αντιστάθμισης, υλοποιήθηκαν τρία φίλτρα όμοια για το κάθε ηχητικό κανάλι των μικροφώνων σύμφωνα με τη μεθοδολογία που περιγράφηκε στην Ενότητα 4.3. Ειδικότερα, υλοποιήθηκε ένα υψηλοπερατό φίλτρο άπειρης χροστικής απόκρισης πρώτου βαθμού (IIR 1 Order Filter) όπως φαίνεται η απόκρισή του στο Σχήμα 5.7α'. Η συχνότητα αποκοπής του, είναι στα  $1.5\text{ kHz}$  με σκοπό τη μείωση της ενέργειας που συσσωρεύεται στον ακουστικό πόρο από την εφαρμογή του ακουστικού.

Για την ακύρωση της αντήχησης του ακουστικού πόρου με μήκος κύματος  $l/2$ , σχεδιάστηκε το φίλτρο εγχοπής τύπου άπειρης χροστικής απόκρισης πρώτου βαθμού με απόκριση συχνότητας όπως φαίνεται στο Σχήμα 5.7β' με συχνότητα εγχοπής στα  $7\text{ kHz}$ .

Η αποκατάσταση της πρώτης αντήχησης του ακουστικού καναλιού με μήκος κύματος  $l/4$  υλοποιείται με την εισαγωγή του φίλτρου κορυφής τύπου άπειρης χροστικής απόκρισης πρώτου βαθμού με απόκριση συχνότητας σύμφωνα με το Σχήμα 5.7γ' και συχνότητα κορυφής στα  $2.2\text{ kHz}$ .



Σχήμα 5.6: Περιβάλλον Υλοποίησης Φιλτραρίσματος Ακουστικού Καναλιού «Sigma Studio»



(α') Υψηλοπερατό Φίλτρο



(β') Φίλτρο εγκοπής



(γ') Φίλτρο Κορυφής

Σχήμα 5.7: Υλοποίηση φίλτρων αντιστάθμισης επίδρασης ακουστικού καναλιού: α)Υψηλοπερατό Φίλτρο, β)Φίλτρο εγκοπής, γ)Φίλτρο Κορυφής

Η συνολική απόκριση των φίλτρων ρυθμίστηκε σύμφωνα με εμπειρικές δοκιμές των ακουστικών στο εργαστήριο. Στόχος ήταν η μέγιστη αντιστάθμιση της επίδρασης των ακουστικών κατά την αναπαραγωγή του πραγματικού περιβάλλοντος, με σκοπό τη βέλτιστη εμπειρία κατά την αξιολόγηση του μοντέλου.

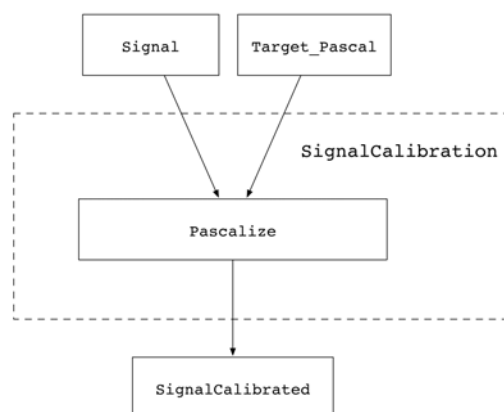
### 5.2.2 Διεργασίες Βαθμονόμησης σημάτων εισόδου

Το μοντέλο δέχεται ως εισερχόμενες ηχητικές ροές το αμφιωτικό σήμα του ψευδοακουστικού περιβάλλοντος και την εικονική πηγή ήχου. Τα σήματα του ψευδοακουστικού περιβάλλοντος εισάγονται στο μοντέλο μέσα από την κάρτα ήχου που λειτουργεί ως συσκευή διεπαφής για το μοντέλο, ενώ η εικονική πηγή ήχου βρίσκεται αποθηκευμένη στον υπολογιστή.

Όπως αναφέρθηκε προηγουμένως, το μοντέλο αναπτύχθηκε στο περιβάλλον εξομίωσης Simulink όπου ανήκει στο ευρύτερο περιβάλλον της Matlab. Μέσα σε αυτό το περιβάλλον, το ηχητικό σήμα λαμβάνει τιμές που κυμαίνονται μέσα στο εύρος από  $-1$  μέχρι  $1$ . Οι τιμές αυτές είναι δύσχρηστες για εφαρμογή στο ακουστικό μοντέλο και κατά συνέπεια θα πρέπει να βαθμονομηθούν σε στάθμες ακουστικής πίεσης  $dB SPL$ .

Η βαθμονόμηση των εισερχόμενων σημάτων επιτελείται μέσα από τον ομώνυμο αλγόριθμο ο οποίος δομείται στα ακόλουθα μέρη (βλπ Σχ. 5.8) και χρησιμοποιείται για κάθε εισερχόμενο σήμα.

Στόχος της βαθμονόμησης είναι ο ορισμός της στάθμης του εισερχόμενου σήματος στο σύστημα σύμφωνα με τη στάθμη του στον πραγματικό κόσμο. Με αυτό τον τρόπο υπολογίζεται η επίδραση των ενδιάμεσων κυκλωμάτων όπως είναι η κάρτα ήχου και η μονάδα ψηφιακής επεξεργασίας ήχου σχετικά με το φιλτράρισμα του αμφιωτικού σήματος και εξάγεται ο αντίστοιχος συντελεστής βαθμονόμησης.



Σχήμα 5.8: Διάγραμμα Αλγόριθμου Βαθμονόμησης Ηχητικού Σήματος

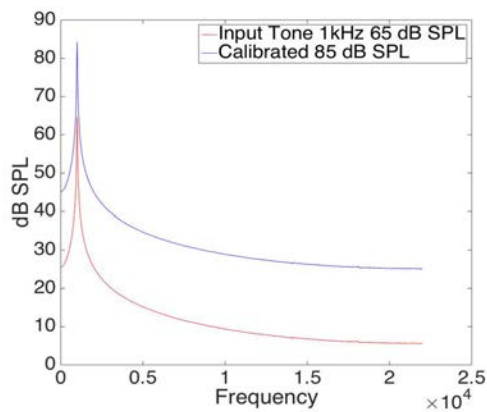
α) Εισαγωγή του σήματος (Signal) στον αλγόριθμο βαθμονόμησης και ορισμός της στάθμης στόχου (Target\_Pascal) η οποία αντιστοιχεί την στάθμη σήματος στον πραγματικό κόσμο. Σε αυτό το επίπεδο το ψηφιοποιημένο αμφιωτικό σήμα του πραγματικού περιβάλλοντος εισάγεται στον αλγόριθμο και ορίζεται η επιθυμητή στάθμη του κάθε καναλιού στον πραγματικό κόσμο. Στο αμφιωτικό σήμα, επιλέχθηκε να μοιραστεί ισομερώς η συνολική ενέργεια στο κάθε κανάλι σύμφωνα με την θεωρία του λόγου των μονοωτικών σημάτων ως προς το αμφιωτικό όπου δεν ξεπερνά την τιμή 2 [115]. Για το μονοφωνικό σήμα της εικονικής ηχητικής πηγής επιλέχθηκε η στάθμη να είναι ίση με τη συνολική στάθμη του αμφιωτικού σήματος.

β) Στο δεύτερο επίπεδο γίνεται ο υπολογισμός του ηχητικού σήματος σε στάθμη ακουστι-

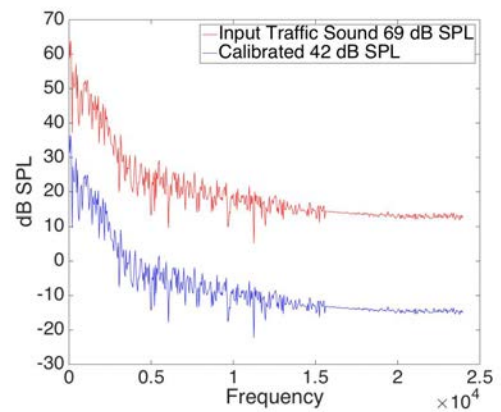
κής πίεσης  $dB SPL$ . Πιο αναλυτικά, σε αυτό το επίπεδο γίνεται η βαθμονόμηση των τιμών πλάτους που λαμβάνει το σήμα από το περιβάλλον εξομοίωσης Simulink οι οποίες κυμαίνονται στο εύρος  $[-1..1]$  σε τιμές Pascal με στάθμη αναφοράς  $20e - 6$ . Ακολούθως, γίνεται ο μετασχηματισμός της τιμής από Pascal σε στάθμη ακουστικής πίεσης  $dB SPL$ .

γ) Στο τρίτο επίπεδο, γίνεται η εύρεση του συντελεστή βαθμονόμησης  $c$  (βλπ Πηγ. Κώδ. B.1, Παράρτημα Β'), ο οποίος είναι ο λόγος της διαφοράς σε κλίμακα  $dB SPL$  ο οποίος εφαρμόζεται στο σήμα για να λάβει την επιθυμητή στάθμη που έχει τεθεί στο μοντέλο (Target\_Pascal) σύμφωνα με τη στάθμη που έχει στο πραγματικό περιβάλλον.

Στο Σχήμα 5.9α' απεικονίζεται η περίπτωση βαθμονόμησης ενός τόνου  $1 kHz$  ο οποίος έχει ηχητική στάθμη καθώς εισέρχεται στο σύστημα στα  $65 dB SPL$  και κατά την βαθμονόμηση του έχει τεθεί ο στόχος της στάθμης στα  $85 dB SPL$  που είναι η στάθμη αναπαραγωγής στο πραγματικό περιβάλλον όπου και επιτυγχάνεται. Εκτός από την παραπάνω περίπτωση, στο Σχήμα 5.9β' παρουσιάζεται το παράδειγμα βαθμονόμησης ενός ηχητικού δείγματος από κίνηση δρόμου. Η εισερχόμενη στάθμη του ηχητικού δείγματος είναι  $69 dB SPL$  και η στοχευμένη στάθμη είναι στα  $42 dB SPL$ . Σύμφωνα με τη φασματική παράσταση και τον υπολογισμό της RMS τιμής επιτυγχάνεται ο στόχος των  $42 dB SPL$ .



(α') Καθαρός Τόνος  $1 kHz$



(β') Ηχητικό Δείγμα κίνησης δρόμου

Σχήμα 5.9: Περιπτώσεις χρήσης Αλγορίθμου Βαθμονόμησης : α)Καθαρός Τόνος  $1 kHz$ , β)Δοκιμαστικός Ήχος από Θόρυβο κίνησης σε δρόμο

Η επιλογή της παραμέτρου Target\_Pascal για τον ορισμό του επιπέδου μέτρησης της ενέργειας στο πραγματικό περιβάλλον είναι αρκετά σημαντική στο μοντέλο. Πρέπει να αναφερθούν δύο στοιχεία ως προς την επιλογή της στοχευμένης στάθμης. Αρχικά, η στάθμη μεταξύ των περιβαλλόντων επιλέχθηκε να είναι όμοια. Δηλαδή, για παράδειγμα, αν η στάθμη στο εικονικό περιβάλλον είναι  $80 dB SPL$ , τότε και το αμφιωτικό σήμα θα έχει συνολική στοχευμένη στάθμη στα  $80 dB SPL$ . Αυτό έγινε για να παρουσιαστεί η διαφορά μεταξύ της αμφιωτικής ακουστότητας με την ενέργεια ακόμα και σε περιπτώσεις που θεωρείται ότι η ενέργεια είναι ίση. Ωστόσο η τιμή θα μπορούσε στο μέλλον να ορίζεται αυτόματα από ένα μετρητικό σύστημα του πραγματικού περιβάλλοντος.

Το τελικό βαθμονομημένο σήμα του κάθε ηχητικού περιβάλλοντος, ακολούθως θα εισαχθεί για ανάλυση και εκτίμηση της ενέργειας διέγερσης, όπως αναλύεται στην επόμενη ενότητα.

### 5.2.3 Μέτρηση ενέργειας διέγερσης σημάτων εισόδου

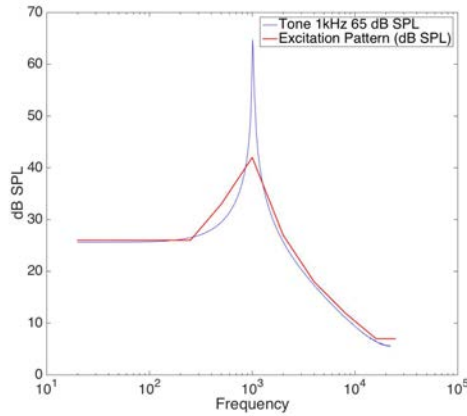
Για τη μέτρηση της ενέργειας διέγερσης των σημάτων εισόδου, ακολουθείται η μεθοδολογία που περιγράφεται στην Ενότητα 4.5. Ο αλγόριθμος που αναπτύχθηκε παρουσιάζεται στον Πηγαίο Κώδικα B.2 στο Παράρτημα Β'.

Αρχικά, το βαθμονομημένο σήμα εισάγεται στη συνάρτηση υπολογισμού της φασματικής ενέργειας (βλπ Πηγ. Κώδ. B.2, Παρ.Β'). Για την ανάλυση του σήματος χρησιμοποιείται η συνάρτηση γρήγορου μετασχηματισμού Fourier (FFT).

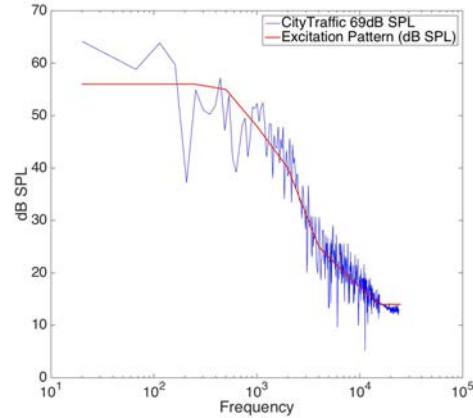
Το αποτέλεσμα της φασματικής ανάλυσης είναι η ενέργεια της κάθε συχνότητας σε κλίμακα *dB SPL*. Ο αριθμός των δειγμάτων για τον οποίο γίνεται η παραπάνω ανάλυση είναι ίδιος με το μέγεθος του παραθύρου που χρησιμοποιείται από το μοντέλο δηλαδή τα 1024 δείγματα. Για τον υπολογισμό του πίνακα διέγερσης, το παράθυρο των δειγμάτων διαιρείται σε αντίστοιχες περιοχές με τις επιλεγμένες ακουστικές ζώνες συχνοτήτων. Οι προεπιλεγμένες ακουστικές ζώνες έχουν κεντρικές συχνότητες τις ακόλουθες και εκτείνονται στη λογαριθμική κλίμακα από 250 *Hz* μέχρι 16000 *Hz* (250 *Hz*, 500 *Hz*, 1000 *Hz*, 2000 *Hz*, 4000 *Hz*, 8000 *Hz*, 16000 *Hz*).

Για τις παραπάνω ζώνες συχνοτήτων, υπολογίζεται ο μέσος όρος της ενέργειας για κάθε μπάντα συχνοτήτων (βλπ Πίν. 4.3). Στις φασματικές αναλύσεις του Σχήματος 5.10 παρουσιάζονται τέσσερα παραδείγματα ηχητικών δειγμάτων σε σύγκριση με την καμπύλη διέγερσης, η οποία υπολογίζεται από τον αλγόριθμο. Στο παράδειγμα του καθαρού τόνου (βλπ Σχ. 5.10α'), η καμπύλη διέγερσης που έχει υπολογιστεί, στη συχνότητα 1 *kHz*, παρουσιάζει κορυφή στα 42 *dB SPL* σε σύγκριση με την κορυφή των 65 *dB SPL* της φασματικής ανάλυσης. Στη δεύτερη περίπτωση του δοκιμαστικού ηχητικού δείγματος από ήχο πόλης σε στάθμη 65 *dB SPL*, φαίνεται πως η καμπύλη διέγερσης ακολουθεί καλύτερα τα αποτελέσματα της φασματικής ανάλυσης του ηχητικού δείγματος. Παρόμοια συμβαίνει και στον τρίτο δοκιμαστικό ήχο που αντικατοπτρίζει έναν ηλεκτρονικό τεχνητό ήχο (βλπ Σχ. 5.10γ'). Στο παράδειγμα του λευκού θορύβου (βλπ Σχ. 5.10δ'), φαίνεται πως υπάρχει μια διαφοροποίηση με την φασματική ανάλυση του αρχικού σήματος σε σχέση και με τις υπόλοιπες περιπτώσεις.

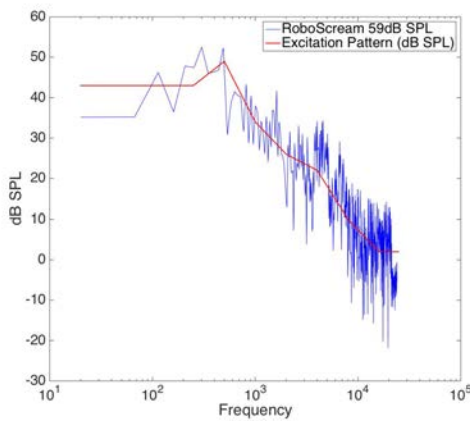
Σύμφωνα με τις παραπάνω περιπτώσεις, η μεθοδολογία υπολογισμού της καμπύλης διέγερσης, υστερεί την αποτύπωση των φασματικών αλλαγών σε ήχους όπως είναι οι καθαροί τόνοι λόγω της φύσης του υπολογισμού του μέσου όρου ο οποίος τείνει να μειώνει και να εξομαλύνει αριθμητικά ένα αποτέλεσμα το οποίο προέρχεται από έντονες μεταβολές, όπως στην περίπτωση του καθαρού τόνου, όπου η σημαντική πληροφορία τείνει να εξομοιωθεί με την υπόλοιπη της ακουστικής ζώνης των 1000 *Hz*. Ωστόσο, στις περιπτώσεις των ήχων με ευρύτερο φάσμα, όπως είναι οι ήχοι των Σχημάτων 5.10β' και 5.10γ', φαίνεται πως εκτιμάται καλύτερα η καμπύλη διέγερσης. Στην περίπτωση του λευκού θορύβου, λόγω των έντονων φασματικών εναλλαγών ενέργειας, η τιμή διέγερσης παρουσιάζει αποκλίσεις από τις τιμές της φασματικής ανάλυσης



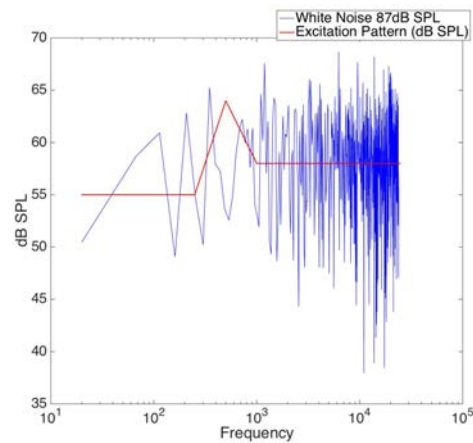
(α') Τόνος 1 kHz



(β') Ηχητικό Δείγμα από Κίνηση Πόλης



(γ') Δοκιμαστικός ήχος Roboscream



(δ') Λευκός Θόρυβος

Σχήμα 5.10: Διαγράμματα υπολογισμού Καμπύλης Διέγερσης για: α) Τόνο 1 kHz, β) Ηχητικό δείγμα Κίνησης Πόλης, γ) Ηχητικό δείγμα «Roboscream», δ) Ηχητικό δείγμα Λευκού θορύβου

με τάση να μειώνεται η υπολογισμένη ενέργεια όπως γίνεται και με τον καθαρό τόνο.

Η παραπάνω αδυναμία, λαμβάνεται υπόψιν για την επιλογή των ήχων που θα χρησιμοποιηθούν για την αξιολόγηση του μοντέλου από την εφαρμογή στόχου που είναι τα ηχητικά παιχνίδια επαυξημένης πραγματικότητας.

Στην κατηγορία των ηχητικών παιχνιδιών επαυξημένης πραγματικότητας τα οποία αποτελούν την εφαρμογή στόχου, είναι συνήθης η χρήση συνθετικών ήχων οι οποίοι εξαρτώνται από την τεχνική ηχοποίησης που έχει επιλεγεί [66]. Συνεπώς θα γίνει σχεδίαση ήχων με πιο ομαλές ενεργειακές μεταβολές στο πεδίο της συχνότητας αποκλείοντας καθαρούς θορύβους καθώς και ήχους με πολύ στενό εύρος όπως είναι οι τόνοι καθώς θα υπάρχει έλλειμμα στην εκτίμηση της καμπύλης διέγερσης.

Επιπλέον, λόγω της χρήσης των εξαγόμενων τιμών ως μεταβλητών για την άμεση προσαρμογή της μείξης, υλοποιήθηκε η μέθοδος για έλεγχο των τιμών στις ζώνες της καμπύλης διέγερσης στο πεδίο του χρόνου με σκοπό την αποφυγή ενσωμάτωσης στο αποτέλεσμα της μείξης έντονων ηχητικών εναλλαγών παραμορφώνοντας το ακουστικό αποτέλεσμα. Η μέθοδος σύμφωνα με την Ενότητα 4.5.3, εισάγει ζεύγος ορίων τα οποία ελέγχουν τη στιγμιαία τιμή ενέργειας για κάθε ζώνη σε σχέση με προκαθορισμένα ζεύγη ορίων.

Στο μοντέλο δοκιμάστηκε η χρήση ενός μικρού ορίου  $\pm 3 \text{ dB SPL}$  για την σταδιακή αυξομείωση της ενέργειας με το βέλτιστο ακουστικό αποτέλεσμα, και ένα μεγαλύτερο όριο της τάξης των  $\pm 28 \text{ dB SPL}$  για την περίπτωση που υπάρχει έντονη εναλλαγή στην εκάστοτε ζώνη και χρειάζεται να υπάρξει επαναφορά της ενέργειας στο στιγμιαίο επίπεδο το οποίο διατηρείται στη συνέχεια. Τα παραπάνω όρια τέθηκαν κατόπιν εμπειρικής δοκιμής του μοντέλου. Οι τελικές ελεγχόμενες τιμές της συνάρτησης υπολογισμού της ενέργειας διέγερσης, στη συνέχεια εισάγονται στον αλγόριθμο προσαρμογής για την υλοποίηση της τελικής μείξης.

#### 5.2.4 Υλοποίηση Ανάλυσης και Σύνθεσης σήματος Εικονικού περιβάλλοντος

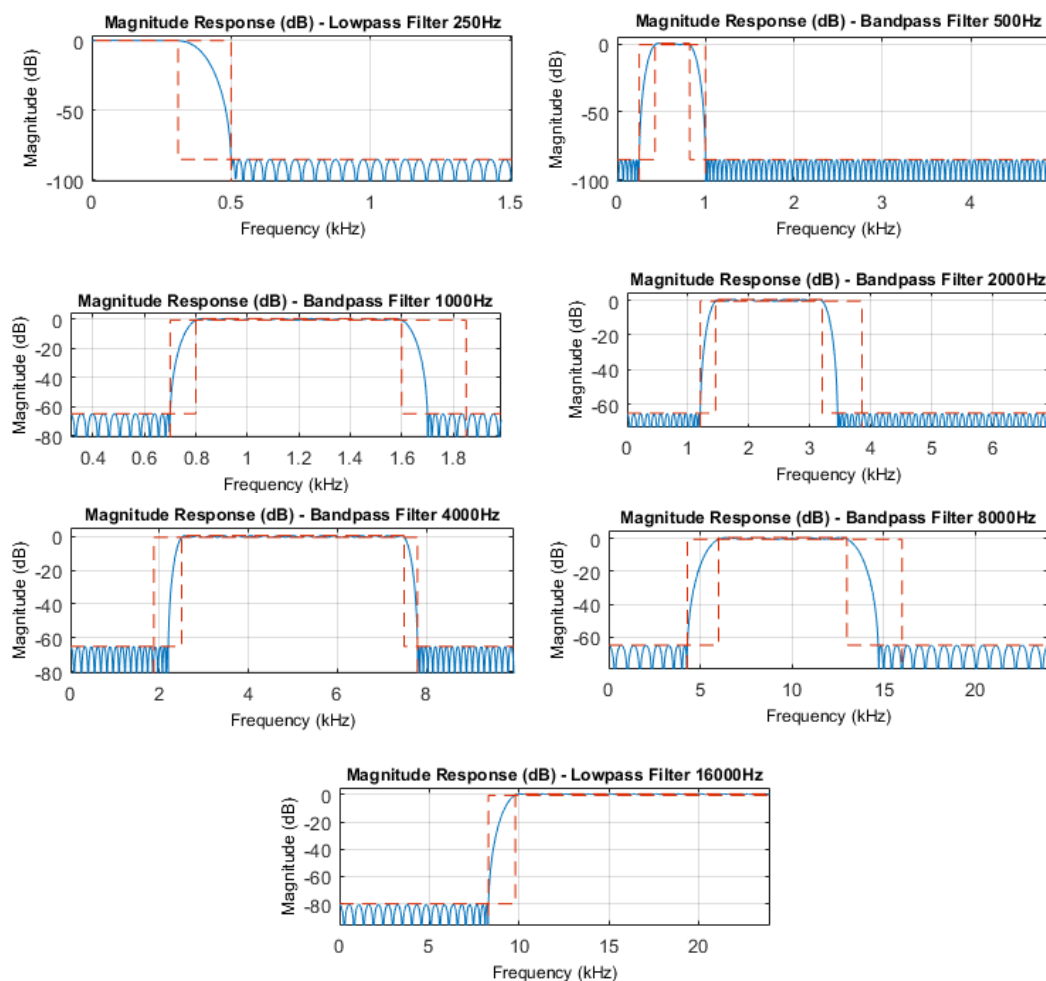
Σε αυτή την ενότητα θα περιγραφεί η υλοποίηση των φίλτρων που χρησιμοποιούνται από το μοντέλο για την ανάλυση και την επανασύνθεση του εικονικού σήματος ήχου. Η χρήση μικρού παραθύρου της τάξης των 1024 δειγμάτων, αποτελεί ανασταλτικό παράγοντα για την επεξεργασία του σήματος στο πεδίο της συχνότητας μέσα από τον γρήγορο μετασχηματισμό Fourier και την επανασύνθεση του σήματος με τη χρήση αντίστροφου μετασχηματισμού Fourier. Ο μικρός αριθμός διακριτών σημείων (Bins) της επιλεγμένης FFT ανάλυσης κάνει αδύνατη την επεξεργασία και ενσωμάτωση συντελεστών στο πεδίο της συχνότητας. Αυτό, έχει ως αποτέλεσμα τη χρήση τράπεζας φίλτρων για την επεξεργασία του εικονικού ηχητικού σήματος, ώστε να επιτευχθεί η καλύτερη ανάλυση του σήματος και εφαρμογή σε αυτό των αποτελεσμάτων της προσαρμοστικής μείξης. Για την επανασύνθεση του σήματος, υλοποιείται μια αντίστοιχη τράπεζα φίλτρων σύνθεσης η οποία εξομαλύνει τυχόν παραμορφώσεις στα άκρα των επεξεργασμένων σημάτων.

| Κεντρική Συχνότητα | Ζώνη Φίλτρου       | Τύπος    | Αριθμός Συντελεστών |
|--------------------|--------------------|----------|---------------------|
| 250 Hz             | 0 Hz - 500 Hz      | Lowpass  | 739                 |
| 500 Hz             | 250 Hz - 1000 Hz   | Bandpass | 620                 |
| 1000 Hz            | 625 Hz - 1750 Hz   | Bandpass | 1070                |
| 2000 Hz            | 1250 Hz - 3500 Hz  | Bandpass | 432                 |
| 4000 Hz            | 2500 Hz - 7000 Hz  | Bandpass | 348                 |
| 8000 Hz            | 5000 Hz - 14000 Hz | Bandpass | 65                  |
| 16000 Hz           | 8000 Hz - 24000 Hz | Highpass | 83                  |

Πίνακας 5.1: Πίνακας χαρακτηριστικών Τράπεζας Φίλτρων

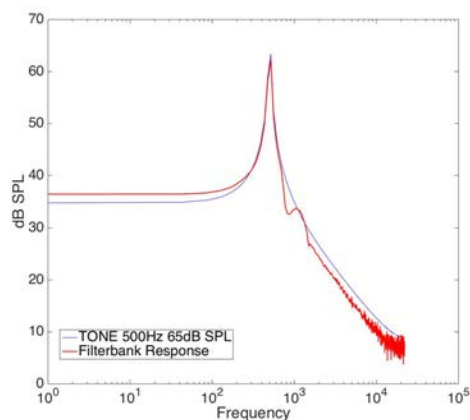
Σύμφωνα με την μεθοδολογία που περιγράφεται στην Ενότητα 4.6, δημιουργήθηκαν επτά

φίλτρα σύμφωνα με τον αριθμό των ζωνών διέγερσης που έχει επιλεχθεί. Για την υλοποίηση του βέλτιστου δυνατού φίλτρου πεπερασμένης χρονικής απόκρισης με το οποίο να παρέχεται η βέλτιστη απόκριση συχνότητας σύμφωνα με τον Πίνακα 5.1, και με το μικρότερο πλήθος συντελεστών, χρησιμοποιήθηκε ο αλγόριθμος Parks-McClellan [80]. Η υλοποίηση του βέλτιστου FIR ισοκυματικού φίλτρου (equiripple filter) έγινε στο περιβάλλον της Simulink με δοκιμές σύμφωνα με τη συχνотική απόκριση του φίλτρου. Στο Σχήμα 5.11 παρουσιάζονται οι συχνотικές αποκρίσεις του κάθε φίλτρου.

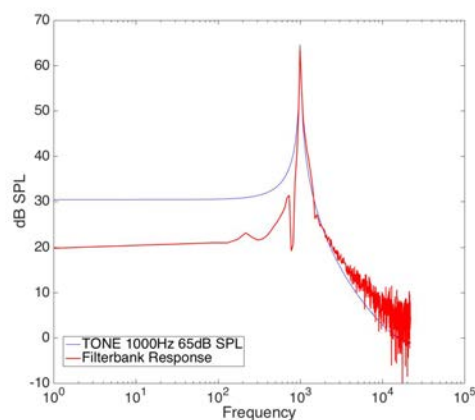


Σχήμα 5.11: Συχνотική Απόκριση Τράπεζας Φίλτρων

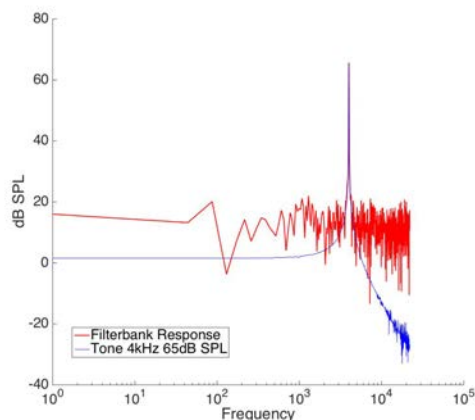
Στο Σχήμα 5.12 αποτυπώνονται τα παραδείγματα εφαρμογής της τράπεζας φίλτρων σε παράθυρο 1024 δειγμάτων. Στις αναλύσεις του συνόλου των δοκιμαστικών ήχων δεν υπάρχει τέλεια σύνθεση του αρχικού σήματος. Στην κατηγορία των τόνων υπάρχει σχετικά μεγαλύτερη απόκλιση στην απόκριση όσο αυξάνεται η συχνότητα ως προς την αντίστοιχη του δοκιμαστικού ήχου. Στην περίπτωση του τεχνητού ήχου Roboscream (βλπ Σχ. 5.12ε') και του λευκού θορύβου στο Σχήμα 5.12ζ', οι αποκρίσεις σε σχέση με το εισερχόμενο σήμα είναι μικρότερες.



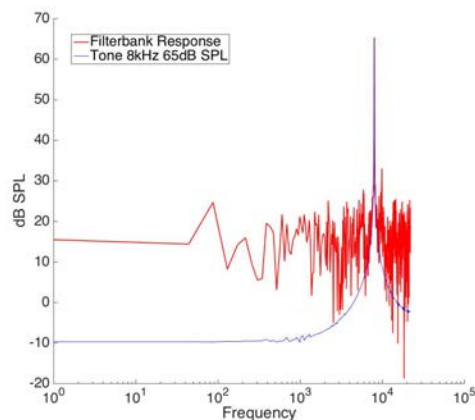
(α') Τόνος 500 Hz



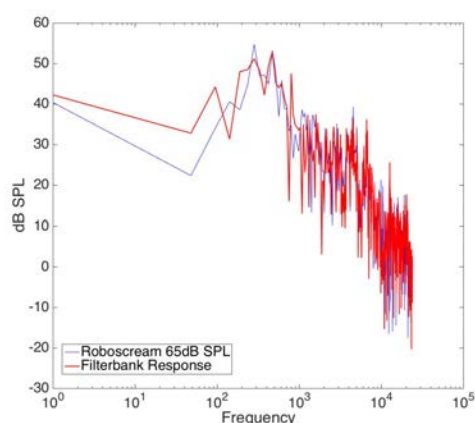
(β') Τόνος 1 kHz



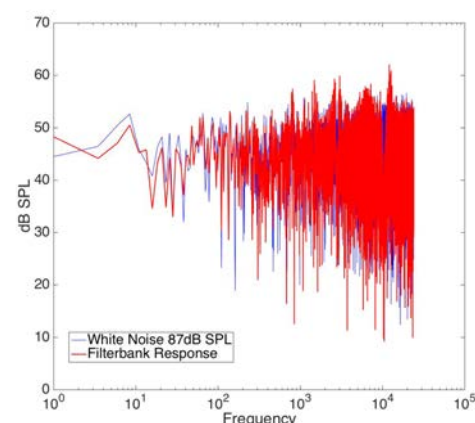
(γ') Τόνος 4 kHz



(δ') Τόνος 8 kHz

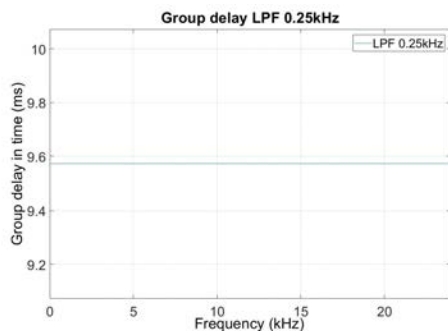


(ε') Δοκιμαστικός ήχος Roboscream

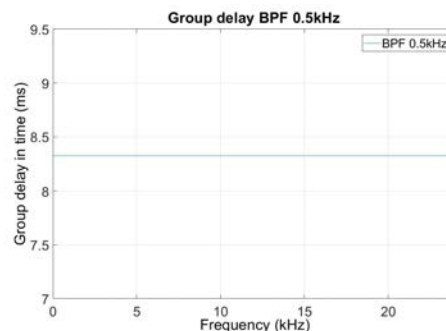


(Ϝ') Λευκός Θόρυβος

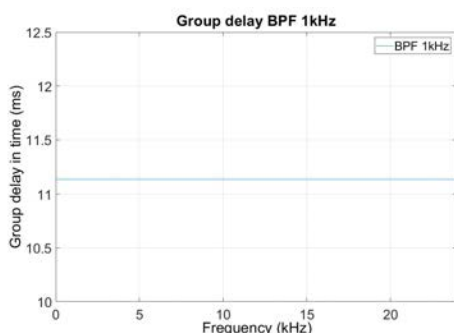
Σχήμα 5.12: Διαγράμματα υπολογισμού φασματικής απόκρισης της Τράπεζας Φίλτρων για: α) Τόνο 500 Hz, β) Τόνο 1 kHz, γ) Τόνο 4 kHz, δ) Τόνο 8 kHz, ε) Ηχητικό δείγμα «Roboscream», ζ) Ηχητικό δείγμα Λευκού θορύβου



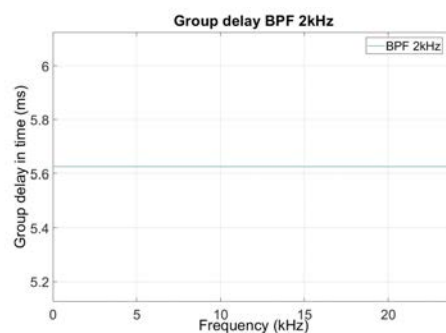
(α') Χαμηλοδιαβατό φίλτρο 250 Hz



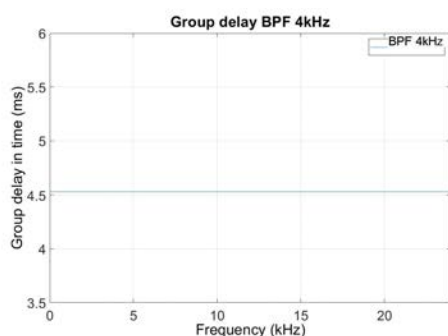
(β') Ζωνοπερατό φίλτρο 500 Hz



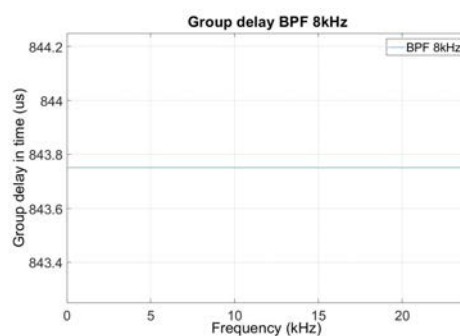
(γ') Ζωνοπερατό φίλτρο 1 kHz



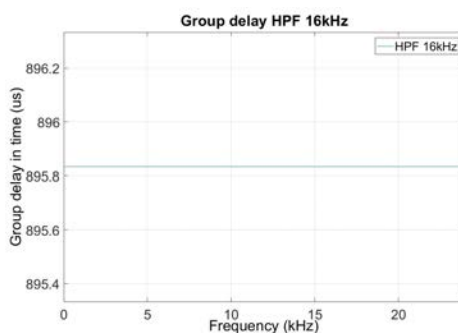
(δ') Ζωνοπερατό φίλτρο 2 kHz



(ε') Ζωνοπερατό φίλτρο 4 kHz



(ς') Ζωνοπερατό φίλτρο 8 kHz



(ζ') Υψηλοπερατό φίλτρο 16 kHz

Σχήμα 5.13: Διαγράμματα απόκρισης φάσης για καθυστέρηση ομάδας Τράπεζας Φίλτρων: α)Χαμηλοδιαβατό φίλτρο 250 Hz, β)Ζωνοπερατό φίλτρο 500 Hz, γ)Ζωνοπερατό φίλτρο 1 kHz, δ)Ζωνοπερατό φίλτρο 2 kHz, ε)Ζωνοπερατό φίλτρο 4 kHz, στ) Ζωνοπερατό φίλτρο 8 kHz, ζ)Υψηλοπερατό φίλτρο 16 kHz

Λόγω της γραμμικής απόκρισης φάσης, όπως φαίνεται και στο Σχήμα 5.13, η καθυστέρηση της ομάδας σε όλα τα φίλτρα, είναι σταθερή με μέγιστη στο φίλτρο του  $1\text{ kHz}$  και ελάχιστη στο φίλτρο των  $8\text{ kHz}$  και  $16\text{ kHz}$ . Σύμφωνα με τον Blauert [12], η μη ακουστή καθυστέρηση κυμαίνεται από  $1\text{ ms}$  στα  $2\text{ kHz}$ ,  $1.5\text{ ms}$  στα  $4\text{ kHz}$ ,  $2\text{ ms}$  στο  $1\text{ kHz}$  και στα  $8\text{ kHz}$  και  $3.2\text{ ms}$  στα  $500\text{ Hz}$ . Ο συγχρονισμός του πλάτους του σήματος, πραγματοποιείται με την εφαρμογή καθυστέρησης στο κάθε φίλτρο λαμβάνοντας ως αρχική την μέγιστη η οποία είναι τα  $11.2\text{ ms}$  στο  $1\text{ kHz}$ .

Συνεπώς, ως προς την φάση των σημάτων, υπάρχει η καθυστέρηση των  $11.2\text{ ms}$  η οποία είναι ανεκτή για την προτεινόμενη μεθοδολογία και τις δοκιμές που γίνονται σχετικά με την αναβάθμιση του μοντέλου επαυξημένης μείξης ήχου. Επίσης, σε ήχους ευρέος φάσματος η απόκριση της τράπεζας φίλτρων, φαίνεται να λειτουργεί ομαλότερα σε σχέση με ήχους πολύ στενού φάσματος. Επιπροσθέτως, στις δοκιμές που έγιναν, το σήμα παράγεται χωρίς ακουστές παραμορφώσεις, οι οποίες θα καθιστούσαν προβληματική τη χρήση της συγκεκριμένης τράπεζας φίλτρων από το μοντέλο.

### 5.2.5 Υλοποίηση Πίνακα μετατροπής Αμφιωτικής Ακουστότητας

Σε αυτή την ενότητα θα αναλυθεί η υλοποίηση του πίνακα μετατροπής αμφιωτικής ακουστότητας σύμφωνα με τη μεθοδολογία που περιγράφηκε στην Ενότητα 4.4.2. Ο πίνακας λειτουργεί ως κανόνας για την εκτίμηση της αμφιωτικής ακουστότητας στις προκαθορισμένες ζώνες συχνοτήτων η οποία αποτελεί παράμετρο για τον αλγόριθμο προσαρμοστικής μείξης.

Η υλοποίηση του πίνακα διαιρέθηκε σε τρία στάδια. Αρχικά, το πρώτο στάδιο ήταν η υλοποίηση των απαραίτητων ηχητικών δειγμάτων προς ανάλυση. Όπως έχει προαναφερθεί, τα δείγματα αποτελούνται από καθαρούς ημιτονοειδείς τόνους σε προκαθορισμένες ηχητικές στάθμες. Τα δείγματα αυτά απαρτίζουν το υλικό για το οποίο θα εκτιμηθεί η αμφιωτική ακουστότητα σύμφωνα με το μοντέλο του Glasberg-Moore. Για την υλοποίηση των δειγμάτων χρησιμοποιήθηκε το προγραμματιστικό περιβάλλον της Matlab. Ο κώδικας που υλοποιήθηκε παρουσιάζεται στον Πηγαίο Κώδικα B.3 στο Παράρτημα Β'. Οι τόνοι που επιλέχθηκαν για ανάλυση είναι οι ακόλουθοι:  $250\text{ Hz}$ ,  $500\text{ Hz}$ ,  $1\text{ kHz}$ ,  $2\text{ kHz}$ ,  $4\text{ kHz}$ ,  $8\text{ kHz}$ ,  $16\text{ kHz}$ . Η επιλογή τους έγινε λαμβάνοντας υπόψιν το επίπεδο ανάλυσης του εικονικού περιβάλλοντος που προορίζεται για επεξεργασία. Σύμφωνα με αυτό, η ανάλυσή του πραγματοποιείται από την τράπεζα φίλτρων σε επτά συχνοτικές ζώνες με κεντρική συχνότητα του παραπάνω τόνους.

Ο κώδικας τρέχει μέσα σε ένα προκαθορισμένο αριθμό επαναλήψεων για να βαθμονομηθεί το κάθε ημίτονο στην αντίστοιχη στάθμη ακουστικής πίεσης. Πιο αναλυτικά, επιλέγεται ο αριθμός των δοκιμαστικών επαναλήψεων σύνθεσης μέσα στις οποίες πρέπει να έχουν δημιουργηθεί τα βαθμονομημένα ημίτονα του συγκεκριμένου τόνου. Σε αυτή την περίπτωση τέθηκε το όριο των 10000 επαναλήψεων οι οποίες επαρκούν σύμφωνα με τις δοκιμές για όλες τις περιπτώσεις των προκαθορισμένων τόνων. Κατόπιν ορίζεται η συχνότητα για την οποία θα παραχθούν τα βαθμονομημένα ημιτονοειδή σήματα. Για τη δημιουργία των ημιτονοειδών σημάτων διαφορετικού πλάτους, προγραμματίστηκε ο βρόχος με τις παρακάτω διεργασίες: ορισμός της διάρκειας των παραγόμενων σημάτων στα 5 δευτερόλεπτα, ορισμός του βήματος

με το οποίο αυξάνεται το πλάτος στην τιμή του  $10e - 4$ . Κατόπιν υπολογίζεται η *RMS* τιμή του παραγόμενου ημίτονου, μετατρέπεται σε κλίμακα *dB* και βαθμονομείται σύμφωνα με τη μέγιστη στάθμη για 1 *Pascal* όπου ορίζονται τα 94 *dB SPL*. Κατόπιν υπολογίζεται η τιμή της στάθμης ακουστικής πίεσης και συνεχίζεται ο βρόχος μέχρι να εκπληρωθεί το όριο των ορισθέντων επαναλήψεων. Για κάθε υπολογισμένη στάθμη ακουστικής πίεσης, γίνεται στρογγυλοποίηση και αποθηκεύεται στον επιλεγμένο φάκελο για τη συγκεκριμένη συχνότητα. Ο τελικός φάκελος της συγκεκριμένης συχνότητας περιέχει ηχητικά δείγματα με διαφορά 1 *dB SPL* για ολόκληρο το εύρος ηχητικών στάθμεων από 31 *dB SPL* μέχρι 91 *dB SPL*.

Με την αποθήκευση των δειγμάτων, ολοκληρώνεται το πρώτο στάδιο για την παραγωγή του πίνακα μετατροπής αμφιωτικής ακουστότητας. Στο επόμενο στάδιο, γίνεται η ανάλυση των βαθμονομημένων τόνων με τη χρήση μοντέλου αμφιωτικής ακουστότητας.

Το στάδιο της ανάλυσης και υπολογισμού της αμφιωτικής ακουστότητας των αποθηκευμένων τόνων, γίνεται στο περιβάλλον της Matlab. Για τον υπολογισμό της αμφιωτικής ακουστότητας του κάθε ηχητικού δείγματος, χρησιμοποιήθηκε το μοντέλο του Glasberg-Moore [42] μέσα από το πακέτο εργαλείων Loudness Toolbox 1.2 [38].

Ο υπολογισμός της αμφιωτικής ακουστότητας, έγινε σε κάθε ηχητικό δείγμα από το προηγούμενο επίπεδο με διάρκεια μεγαλύτερη από 100 *ms*. Επιλέχθηκε αυτή η διάρκεια, στην οποία η αμφιωτική ακουστότητα είναι ανεξάρτητη της χρονικής ενσωμάτωσής της [38]. Με αυτό τον τρόπο, οι υπολογισμένες τιμές στον πίνακα του Παραρτήματος Α', δείχνουν τη σχέση των συχνοτήτων και της μέγιστης αμφιωτικής ακουστότητας που μπορούν να πάρουν για τη συγκεκριμένη ακουστική στάθμη. Με τον παραπάνω τρόπο, υπολογίζονται οι μέγιστες τιμές αμφιωτικής ακουστότητας που μπορεί να έχει ο κάθε καθαρός τόνος στις προδιαγεγραμμένες στάθμες ακουστικής πίεσης.

Το τρίτο στάδιο αποτελεί την ενσωμάτωση των αποτελεσμάτων αμφιωτικής ακουστότητας σε έναν ενιαίο πίνακα μετατροπής στάθμης ακουστικής πίεσης σε τιμές αμφιωτικής ακουστότητας για τις προκαθορισμένες συχνότητες. Ο τελικός πίνακας παρουσιάζεται στο Παράρτημα Α'. Ο πίνακας εισάγεται στον αλγόριθμο εκτίμησης αμφιωτικής ακουστότητας από τον πίνακα ακουστικής διέγερσης.

Ο υπολογισμός του πίνακα μετατροπής αμφιωτικής ακουστότητας, λειτουργεί άτυπα ως ένα σύνολο από κανόνες οι οποίοι διέπουν οριζοντίως τις ακουστικές σχέσεις μεταξύ των συχνοτήτων ως προς την αμφιωτική ακουστότητα, αλλά και καθέτως καθορίζουν τη σχέση της κάθε συχνότητας ως προς την στάθμη ακουστικής πίεσης και αμφιωτικής ακουστότητας.

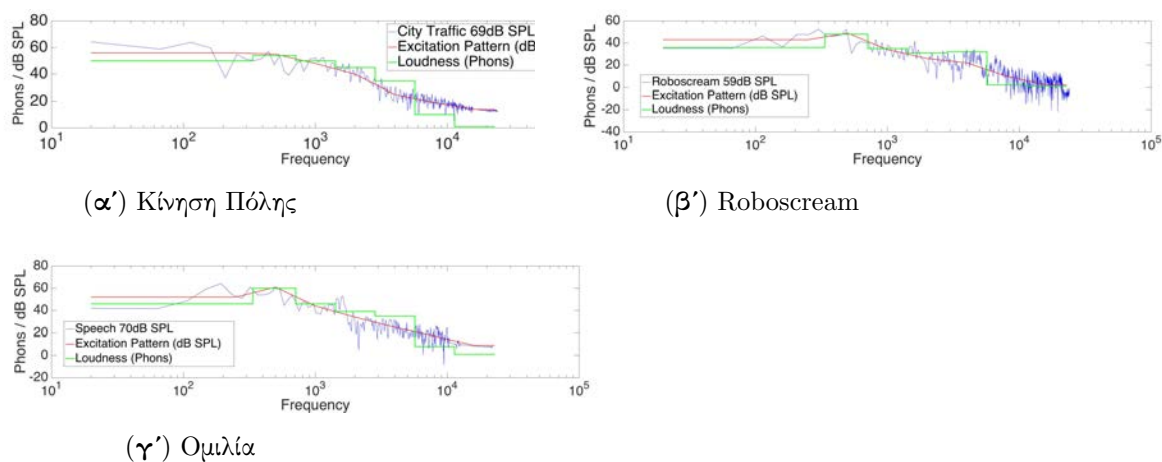
### 5.2.6 Εκτίμηση Αμφιωτικής Ακουστότητας Σημάτων Εισόδου

Στα σήματα εισόδου του μοντέλου προσαρμοστικής μείξης, εφαρμόζεται κοινή μεθοδολογία υπολογισμού της ενέργειας της ακουστικής διέγερσης, όπως περιγράφεται στην Ενότητα 5.2.3. Οι πίνακες ακουστικής διέγερσης του κάθε σήματος εισάγονται στον αλγόριθμο εκτίμησης της αμφιωτικής ακουστότητας για κάθε μπάντα συχνοτήτων.

Η υλοποίηση του αλγόριθμου υπολογισμού της αμφιωτικής ακουστότητας, γίνεται στο προγραμματιστικό περιβάλλον της Matlab. Η εκτίμηση της αμφιωτικής ακουστότητας από

την καμπύλη διέγερσης, αποτελεί το τμήμα του μοντέλου στο οποίο γίνεται η μετατροπή της ακουστικής διέγερσης σε μονάδα αμφιωτικής ακουστότητας για κάθε ζώνη συχνοτήτων. Ακολουθώντας την μεθοδολογία, στο πρώτο επίπεδο εισάγεται ο πίνακας μετατροπής αμφιωτικής ακουστότητας του Παραρτήματος Α' στη συνάρτηση υπολογισμού αμφιωτικής ακουστότητας (βλπ Πηγ. Κώδ. Β.4, Παρ.Β'). Ο πίνακας συμπυκνώνει ένα σύνολο σχέσεων μεταξύ ηχητικής στάθμης, αμφιωτικής ακουστότητας και συχνότητας και λειτουργεί ως κανόνας για τον υπολογισμό των τελικών τιμών αμφιωτικής ακουστότητας για τα σήματα εισόδου.

Ακολουθώντας, στη συνάρτηση εισάγονται οι πίνακες ακουστικής διέγερσης του ψευδοακουστικού περιβάλλοντος και του εικονικού αντίστοιχα. Στο υλοποιημένο μοντέλο, ο πίνακας διέγερσης αποτελείται από επτά τιμές αντίστοιχες των συχνοτικών ζωνών. Οι τιμές του πίνακα διέγερσης αποτυπώνονται σε μονάδα μέτρησης dB SPL. Στο επόμενο επίπεδο, υλοποιείται ο βρόχος μέσω του οποίου γίνεται η αναζήτηση στον πίνακα μετατροπής αμφιωτικής ακουστότητας, για την κοντινότερη τιμή αμφιωτικής ακουστότητας (Phons) της εκάστοτε ζώνης συχνοτήτων σύμφωνα με τις τιμές του πίνακα της ακουστικής διέγερσης (dB SPL) για κάθε ηχητικό περιβάλλον της επαυξημένης πραγματικότητας. Το τελικό αποτέλεσμα που εξάγεται από την παρούσα συνάρτηση, αποτελείται από τον πίνακα αμφιωτικής ακουστότητας εκφρασμένο σε Phons για το εικονικό περιβάλλον και ομοίως και για το ψευδοακουστικό. Παρακάτω, παρουσιάζονται παραδείγματα υπολογισμού της αμφιωτικής ακουστότητας με τη χρήση της υφιστάμενης συνάρτησης και του πίνακα μετατροπής ακουστότητας (βλπ Σχ. 5.14).



Σχήμα 5.14: Διαγράμματα εκτίμησης Αμφιωτικής Ακουστότητας μέσα από την Καμπύλη Διέγερσης για μοντέλο AR-ION: α)Κίνηση Πόλης (69 dB SPL), β)Roboscream (59 dB SPL), γ)Ομιλία (70 dB SPL)

Αρχικά, για τα παραδείγματα, επιλέχθηκαν ήχοι που θα μπορούσαν να συνθέτουν σενάριο σε ένα περιβάλλον επαυξημένης πραγματικότητας ήχου. Σε αυτά είναι ευδιάκριτο πως η καμπύλη διέγερσης ακολουθείται από την καμπύλη αμφιωτικής ακουστότητας, ωστόσο για να εξαχθούν συμπεράσματα ως προς την ποιότητα απόδοσης της αμφιωτικής ακουστότητας, θα πρέπει να πραγματοποιηθεί σύγκριση με υφιστάμενο μοντέλο. Για αυτό τον λόγο, σχεδιάστηκε μια μεθοδολογία σύγκρισης μεταξύ του προσαρμοστικού μοντέλου που υλοποιήθηκε για την

παρούσα εργασία με το όνομα AR-ION (Augmented Reality of IONian University και του μοντέλου του Glasberg - Moore (GM)).

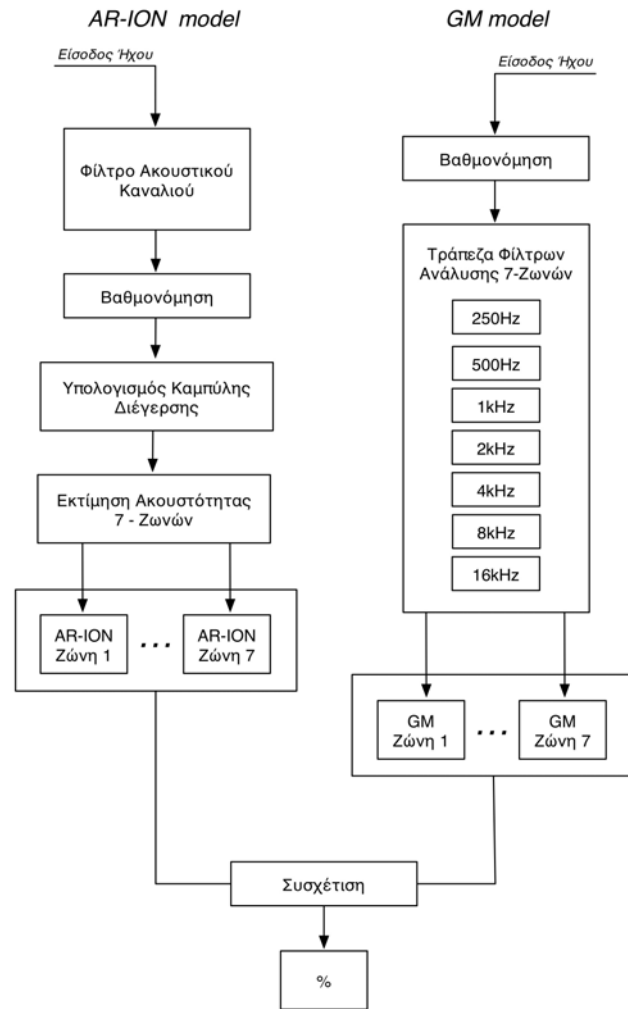
Η μεθοδολογία, παρουσιάζεται στο Σχήμα 5.15. Στόχος είναι η σύγκριση των τιμών αμφιωτικής ακουστότητας σε επίπεδο συχνοτικών ζωνών μεταξύ του υφιστάμενου μοντέλου και του προτεινόμενου και ο έλεγχος της ομοιότητάς τους αν υπάρχει. Για το προσαρμοστικό μοντέλο, η μεθοδολογία εκτίμησης της αμφιωτικής ακουστότητας ακολουθεί τα παρακάτω βήματα. Αρχικά ο εισερχόμενος ήχος φιλτράρεται σύμφωνα με τα υλοποιημένα φίλτρα αντιστάθμισης του ακουστικού πόρου. Ακολούθως γίνεται βαθμονόμηση του σήματος. Η βαθμονόμηση αποτελεί κοινό στοιχείο και για την ανάλυση του άλλου συγκρινόμενου μοντέλου, καθώς θα πρέπει να έχουν το ίδιο κέρδος. Στη συνέχεια, για το μοντέλο AR-ION πραγματοποιείται ο υπολογισμός της καμπύλης διέγερσης μέσω της οποίας γίνεται η εκτίμηση της αμφιωτικής ακουστότητας για τις προεπιλεγμένες επτά ζώνες συχνοτήτων.

Για το μοντέλο GM, σχεδιάστηκε και υλοποιήθηκε η παρακάτω μεθοδολογία. Το σήμα εισόδου εισάγεται για βαθμονόμηση όμοια με το παραπάνω μοντέλο χωρίς τη χρήση φίλτρου αντιστάθμισης ακουστικού πόρου, καθώς το μοντέλο που χρησιμοποιήθηκε το περιλαμβάνει στην ανάλυση που θα ακολουθήσει. Το βαθμονομημένο σήμα εισόδου αναλύεται σε επτά ζώνες συχνοτήτων με τη χρήση τράπεζας φίλτρων. Ακολούθως, στο κάθε φιλτραρισμένο σήμα, εφαρμόζεται το GM μοντέλο και υπολογίζεται η αμφιωτική ακουστότητα για κάθε ζώνη συχνοτήτων. Επιπροσθέτως, πρέπει να σημειωθεί πως η ανάλυση και στα δύο μοντέλα έγινε για αριθμό δειγμάτων ίσο με αυτόν που χρησιμοποιείται στο προσαρμοστικό μοντέλο δηλαδή 1024 δείγματα ώστε να υπάρχει αντιστοιχία.

Τα τελικά αποτελέσματα από τα δύο μοντέλα μπορούν να συγκριθούν και να συσχετισθούν ως προς την ομοιότητά τους και την τάση που ακολουθούν. Σύμφωνα με την παραπάνω μεθοδολογία, πραγματοποιήθηκαν μετρήσεις και τα αποτελέσματα καταγράφονται στον Πίνακα 5.2. Πιο αναλυτικά, επιλέχθηκαν οι εξής ήχοι: α)Κίνηση πόλης, β)Ο τεχνητός ήχος Roboscream, γ)Ομιλία ανδρική, δ)Ήχος από βιομηχανία.

Τα αποτελέσματα των μετρήσεων παρουσιάζονται στον Πίνακα 5.2. Πιο ειδικά για κάθε ηχητικό δείγμα, καταγράφηκε για το AR-ION μοντέλο η ακουστική διέγερση, και η αμφιωτική ακουστότητα για κάθε συχνοτική ζώνη και για το μοντέλο GM, η αμφιωτική ακουστότητά του. Τα αποτελέσματα αποτυπώνονται γραφικά και στα διαγράμματα του Σχήματος 5.15. Σε αυτά τα διαγράμματα έχει ενδιαφέρον να γίνουν κάποιες παρατηρήσεις. Αρχικά, φαίνεται ότι σε όλες τις περιπτώσεις οι εκτιμώμενες τιμές αμφιωτικής ακουστότητας του GM μοντέλου, είναι υψηλότερες από το μοντέλο AR-ION. Αυτό δείχνει μια ομοιομορφία στη λειτουργία του AR-ION μοντέλου στοιχείο το οποίο κρίνεται θετικά για την σύγκριση που επιδιώκεται στην προσαρμοστική μείξη καθώς προσδίδεται κοινή αντιμετώπιση της κάθε περίπτωσης. Για την καλύτερη σύγκριση της απόδοσης των δύο μοντέλων, έγινε υπολογισμός της συσχέτισής (Correlation) μεταξύ των παραγόμενων τιμών αμφιωτικής ακουστότητας για κάθε ζώνη τους με βάση τη στατιστική ανάλυση (Microsoft Excel) του Πίνακα 5.2.

Η συσχέτιση μεταξύ των δύο μοντέλων στα παραδείγματα που παρουσιάζονται είναι πάνω από 85%. Πιο αναλυτικά, στο παράδειγμα της κίνησης πόλης, παρουσιάζεται συσχέτιση 87%. Στο διάγραμμα σύγκρισης των καμπυλών αμφιωτικής ακουστότητας (βλπ Σχ. 5.16α'), είναι

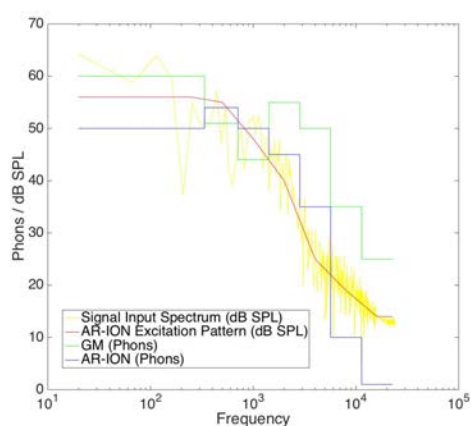


Σχήμα 5.15: Διάγραμμα Μεθοδολογίας Σύγκρισης Μοντέλων AR-ION με GM

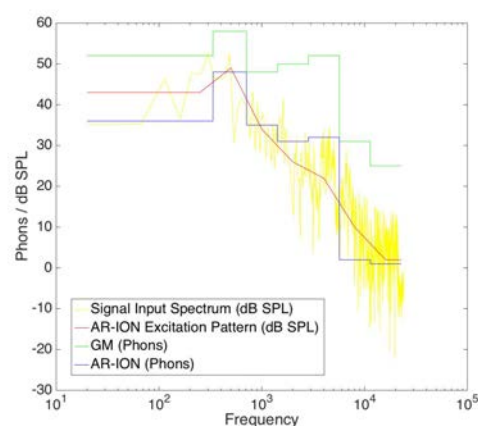
ευδιάκριτη η τάση της αμφιωτικής ακουστότητας των δύο μοντέλων στην οποία παρουσιάζεται μια διαφοροποίηση στη ζώνη των  $500\text{ Hz}$  και του  $1\text{ kHz}$ . Στο παράδειγμα του τεχνητού ήχου Roboscream (βλπ Σχ. 5.16β'), παρουσιάζεται μια διαφοροποίηση στην τάση των συγκρινόμενων μοντέλων αμφιωτικής ακουστότητας στα  $2\text{ kHz}$  με αντίστοιχη συσχέτιση όπως φαίνεται στον Πίνακα 5.16β', στα 97%. Ο επόμενος δοκιμαστικός ήχος είναι ανδρική ομιλία σε στάθμη  $70\text{ dB SPL}$ . Η συσχέτιση των δύο μοντέλων, παρουσιάζεται σε ποσοστό 91%. Στο διάγραμμα του Σχήματος 5.16γ', παρουσιάζεται η τάση των δύο μοντέλων στις καμπύλες αμφιωτικής ακουστότητας. Η διαφοροποίηση στην τάση, φαίνεται να υπάρχει στα  $500\text{ Hz}$  και στα  $2\text{ kHz}$ . Στο τελευταίο παράδειγμα, ο ήχος δοκιμής είναι ήχος βιομηχανίας. Ένα περιβάλλον το οποίο θα μπορούσε να είναι εν δυνάμει ήχος του ψευδοακουστικού περιβάλλοντος. Η συσχέτιση που παρουσιάζει ο συγκεκριμένος ήχος μεταξύ των δύο μοντέλων, είναι 97%. Μέσα σε αυτό το ποσοστό, σύμφωνα με το Σχήμα 5.16δ', παρουσιάζεται μια διαφοροποίηση στην τάση της ζώνης του  $1\text{ kHz}$ .

|           | Κίνηση Πόλης<br>69 dB SPL |        |       | Roboscream<br>59 dB SPL |        |       | Ομιλία<br>70 dB SPL |        |       | Βιομηχανία<br>70 dB SPL |        |       |
|-----------|---------------------------|--------|-------|-------------------------|--------|-------|---------------------|--------|-------|-------------------------|--------|-------|
|           | Διέγερση AR-ION           | AR-ION | GM    | Διέγερση AR-ION         | AR-ION | GM    | Διέγερση AR-ION     | AR-ION | GM    | Διέγερση AR-ION         | AR-ION | GM    |
|           | dB SPL                    | Phons  | Phons | dB SPL                  | Phons  | Phons | dB SPL              | Phons  | Phons | dB SPL                  | Phons  | Phons |
| 250 Hz    | 56                        | 50     | 60    | 43                      | 36     | 52    | 52                  | 46     | 66    | 36                      | 29     | 60    |
| 500 Hz    | 55                        | 54     | 51    | 49                      | 48     | 58    | 61                  | 60     | 63    | 48                      | 47     | 76    |
| 1 kHz     | 48                        | 50     | 44    | 34                      | 35     | 48    | 44                  | 46     | 49    | 35                      | 36     | 69    |
| 2 kHz     | 40                        | 45     | 55    | 26                      | 31     | 50    | 34                  | 39     | 57    | 33                      | 38     | 67    |
| 4 kHz     | 25                        | 35     | 50    | 22                      | 32     | 52    | 25                  | 35     | 53    | 21                      | 31     | 65    |
| 8 kHz     | 19                        | 10     | 35    | 10                      | 2      | 31    | 17                  | 8      | 35    | 18                      | 9      | 50    |
| 16 kHz    | 14                        | 1      | 25    | 2                       | 1      | 25    | 19                  | 1      | 18    | 27                      | 1      | 32    |
| Συσχέτιση |                           | 87%    |       |                         | 97%    |       |                     | 91%    |       |                         | 97%    |       |

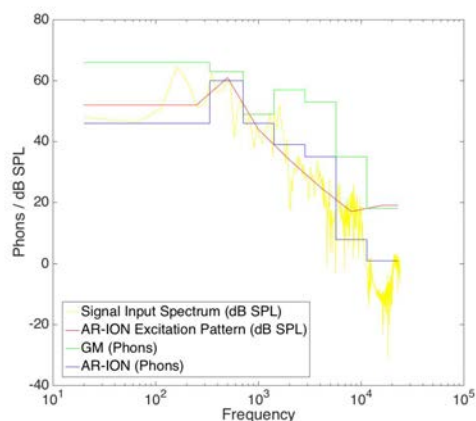
Πίνακας 5.2: Συγκριτικός Πίνακας Συσχέτισης μοντέλων AR-ION και GM



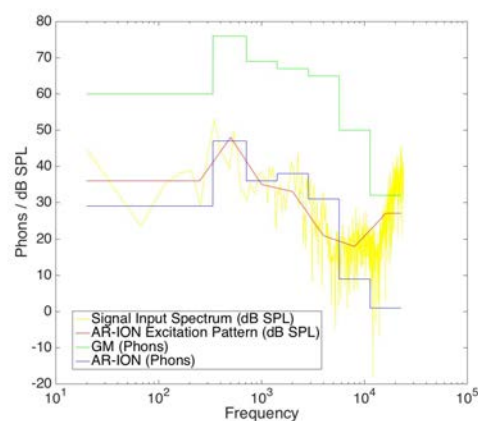
(α') Κίνηση Πόλης 69dB



(β') Roboscream 59dB



(γ') Ομιλία 70dB



(δ') Βιομηχανία 70dB

Σχήμα 5.16: Συγκριτικά Διαγράμματα Συσχέτισης μοντέλων AR-ION και GM: α) Κίνηση Πόλης 69 dB SPL, β) Roboscream 59 dB SPL, γ) Ομιλία 70 dB SPL, δ) Βιομηχανία 70 dB SPL

Κατά συνέπεια, σύμφωνα με την παραπάνω σύγκριση των δύο μοντέλων, το προτεινόμενο μοντέλο AR-ION συσχετίζεται σε μεγάλο βαθμό με το υφιστάμενο μοντέλο GM με διαφορο-

ποιήσεις σε επίπεδο στάθμης αμφιωτικής ακουστότητας αλλά ακολουθώντας μια κοινή τάση στα αποτελέσματα. Η παραγόμενη καμπύλη αμφιωτικής ακουστότητας, στο προσαρμοστικό μοντέλο μείξης υπολογίζεται για το ψευδοακουστικό και το εικονικό περιβάλλον και εισάγεται στον αλγόριθμο προσαρμογής ο οποίος αναλύεται ακολούθως.

### 5.2.7 Υλοποίηση Αλγόριθμου προσαρμογής

Σε αυτή την ενότητα της διατριβής, θα αναλυθεί η υλοποίηση του αλγόριθμου προσαρμογής σύμφωνα με τον οποίο υλοποιήθηκε το μοντέλο προσαρμοστικής επαυξημένης μείξης. Ο αλγόριθμος, περιγράφεται στην Ενότητα 4.7.1. Η υλοποίηση πραγματοποιήθηκε έγινε στο προγραμματιστικό περιβάλλον της Matlab και δοκιμάστηκε για χρήση σε παράθυρο των 1024 δειγμάτων. Παρακάτω, θα αναλυθεί ο αλγόριθμος προσαρμογής μέσα από μεθοδολογία μετρήσεων της απόδοσής του με ηχητικά παραδείγματα.

Σύμφωνα με τις γενικές προδιαγραφές του μοντέλου, ο αλγόριθμος προσαρμογής, υλοποιείται για επεξεργασία καμπύλης αμφιωτικής ακουστότητας επτά ζωνών. Ειδικότερα, λαμβάνει ως είσοδο τον πίνακα αμφιωτικής ακουστότητας του ψευδοακουστικού περιβάλλοντος και του εικονικού αντίστοιχα. Ο αλγόριθμος προσαρμογής αναλαμβάνει να υπολογίσει την διαφορά της αμφιωτικής ακουστότητας μεταξύ των δύο περιβαλλόντων και ακολούθως να την εκφράσει ως συντελεστή ρύθμισης της ενέργειας για κάθε συχνοτική ζώνη του εικονικού σήματος. Πρέπει να υπενθυμιστεί πως το εικονικό σήμα αναλύεται σε ζώνες συχνοτήτων όπως περιγράφεται στην Ενότητα 5.2.4 με τη χρήση τράπεζας φίλτρων και στο κάθε φιλτραρισμένο σήμα εφαρμόζεται ο συντελεστής που υπολογίζεται από τον αλγόριθμο προσαρμογής.

Στον Πηγαίο Κώδικα B.5 του Παραρτήματος Β', παρουσιάζεται η υλοποίηση του αλγόριθμου προσαρμογής. Οι μεταβλητές εισόδου της συνάρτησης προσαρμογής παρουσιάζονται παρακάτω. Η πρώτη μεταβλητή, αποτελείται από τις φιλτραρισμένες ροές ήχου του εικονικού ηχητικού περιβάλλοντος. Ειδικότερα, πρόκειται για τις αριθμητικές ακολουθίες που εξάγονται από την τράπεζα φίλτρων ανάλυσης επτά ζωνών που έχει υλοποιηθεί για το μοντέλο και συνθέτουν τον πίνακα ανάλυσης του εικονικού περιβάλλοντος στις προκαθορισμένες ζώνες. Οι δύο επόμενες μεταβλητές είναι ο πίνακας αμφιωτικής ακουστότητας του ψευδοακουστικού περιβάλλοντος και ο πίνακας αμφιωτικής ακουστότητας του εικονικού περιβάλλοντος. Οι δυο πίνακες εισάγουν την εκτίμηση αμφιωτικής ακουστότητας του μοντέλου για τα ακουστικά περιβάλλοντα. Επίσης, για λόγους μετρήσεων, έχει προστεθεί ένας δείκτης απενεργοποίησης του λόγου που περιγράφεται από την Εξίσωση 4.22 για λειτουργία χωρίς τον συντελεστή αντιστάθμισης της ακουστικής διαφοράς κάθε ζώνης. Επιπροσθέτως, έχει εισαχθεί ένας συντελεστής διαβάθμισης της μείξης ο οποίος αναλύεται παρακάτω (βλπ Πηγ. Κώδ. B.5, Παρ.Β' η μεταβλητή DifferCoef). Η έξοδος του αλγόριθμου προσαρμογής αποτελείται από έναν πίνακα επτά ακολουθιών. Η κάθε ακολουθία αποτελεί το ηχητικό σήμα του εικονικού περιβάλλοντος της κάθε συχνοτικής ζώνης. Οι ακολουθίες, στη συνέχεια εισάγονται στην τράπεζα φίλτρων σύνθεσης για το φιλτράρισμα τυχόν παραμορφώσεων στα όρια της εκάστοτε ζώνης και κατόπιν γίνεται η πρόσθεση των σημάτων για τη δημιουργία του εικονικού προσαρμοσμένου ηχητικού σήματος.

Ο αλγόριθμος προσαρμογής λειτουργεί αυτόνομα όπως και όλο το μοντέλο επαυξημένης μείξης ήχου. Ωστόσο, η διαβάθμιση της μείξης εικονικού περιβάλλοντος με πραγματικό αποτελεί μια προσωπική επιλογή του θεατή [57]. Κατά συνέπεια, έγινε η δοκιμαστική εισαγωγή ενός συντελεστή διαβάθμισης της μείξης για την μελέτη των ορίων στην ισορροπία της μείξης μεταξύ του ψευδοακουστικού και του εικονικού περιβάλλοντος ήχου. Ο συντελεστής διαβάθμισης εισάγεται σε κάθε συχνοτική ζώνη ενισχύοντας την ακουστική διαφορά μεταξύ των περιβαλλόντων. Η αντικειμενική αξιολόγηση της λειτουργίας του αλγόριθμου προσαρμογής και κατ'επέκταση και του συνολικού μοντέλου παρουσιάζεται στην Ενότητα 6.2.1.

### 5.3 Σύνοψη Κεφαλαίου

Σε αυτό το κεφάλαιο παρουσιάστηκαν τα στάδια υλοποίησης του μοντέλου προσαρμοστικής μείξης για περιβάλλοντα επαυξημένης πραγματικότητας ήχου. Στο πρώτο στάδιο, δόθηκε η περιγραφή της υλοποίησης της αμφιωτικής αναπαραγωγής του πραγματικού περιβάλλοντος με την χρήση τεχνολογίας ψηφιακών φίλτρων για την αντιστάθμιση της επίδρασης του κλειστού ακουστικού καναλιού, καθώς και την χρήση ειδικού ζεύγους αμφιωτικών μικροφώνων-ακουστικών τύπου In-Ear για την ρεαλιστικότερη αναπαραγωγή του ψευδοακουστικού περιβάλλοντος, διατηρώντας στο μέγιστο βαθμό τις ανακλάσεις που προκαλεί στον ήχο το εξωτερικό αυτί και το σώμα του ακροατή. Στο επόμενο στάδιο, περιγράφηκε η υλοποίηση βαθμονόμησης των ηχητικών σημάτων που εισάγονται στο μοντέλο. Η βαθμονόμηση, καθορίζεται από το περιβάλλον υλοποίησης του μοντέλου. Σε αυτή την περίπτωση, η υλοποίηση του μοντέλου, επιτελέστηκε στο προγραμματιστικό περιβάλλον της Simulink. Αρχικά, μετατράπηκε η στάθμη του εισερχόμενου σήματος σε κλίμακα Pascal με βάση τις στάθμες που δίνονται από το σύστημα και κατόπιν βαθμονομείται με την εύρεση του συντελεστή βαθμονόμησης σύμφωνα με την πραγματική στάθμη που έχει επιλεχθεί να έχει στον πραγματικό κόσμο. Κατόπιν περιγράφηκε η υλοποίηση του αλγόριθμου μέτρησης της ενέργειας διέγερσης των βαθμονομημένων σημάτων εισόδου με την παρουσίαση μετρήσεων. Στη συνέχεια, δόθηκε η υλοποίηση της τράπεζας φίλτρων που επιτρέπει την ανάλυση του εικονικού περιβάλλοντος για την επεξεργασία του στις προκαθορισμένες συχνοτικές ζώνες, καθώς και η τράπεζα σύνθεσης του τελικού εικονικού σήματος. Η υλοποίηση του πίνακα αμφιωτικής ακουστότητας περιγράφηκε στην Ενότητα 5.2.5 και το αποτέλεσμά της παρουσιάζεται στο Παράρτημα Α'. Η μετατροπή του πίνακα διέγερσης σε αμφιωτική ακουστότητα λαμβάνοντας υπόψιν του κανόνες με βάση τον πίνακα αμφιωτικής ακουστότητας, δίνεται στην Ενότητα 5.2.6, όπου γίνεται επιπλέον μια σύγκριση με το υφιστάμενο μοντέλο αμφιωτικής ακουστότητας. Τέλος, παρουσιάζεται η υλοποίηση του αλγόριθμου προσαρμογής μέσα στο μοντέλο επαυξημένης μείξης. Η συνολική αξιολόγηση του μοντέλου, παρουσιάζεται στο επόμενο κεφάλαιο μέσα από την αντικειμενική και υποκειμενική αξιολόγησή του.



## Κεφάλαιο 6

# Αξιολόγηση μοντέλου Προσαρμοστικής ARA Μείξης

### 6.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο, παρουσιάστηκε η υλοποίηση του προσαρμοσμένου μοντέλου μείξης επαυξημένης πραγματικότητας ήχου. Σε αυτό το κεφάλαιο θα παρουσιαστεί η αξιολόγηση του συνολικού μοντέλου υπό το πρίσμα τριών τύπων αξιολόγησης. Η πρώτη μεθοδολογία αξιολόγησης που παρουσιάζεται είναι αντικειμενική. Βασίζεται σε δεδομένα μετρήσεων της απόδοσης του μοντέλου, σύγκρισης και ανάλυσης των αποτελεσμάτων. Η δεύτερη μεθοδολογία, η οποία αναφέρεται ως υποκειμενική, ακολουθεί την μεθοδολογική προσέγγιση των υποκειμενικών πειραμάτων. Μέσα από την αξιολόγηση του μοντέλου από υποκείμενα και την στατιστική ανάλυση των αποτελεσμάτων, δίνονται τα υποκειμενικά στοιχεία που συνθέτουν την συνολική εικόνα του μοντέλου μείξης. Τέλος, δίνεται η μεθοδολογική προσέγγιση αξιολόγησης ενός ηχητικού παιχνιδιού μέσα στο οποίο εφαρμόζεται το μοντέλο μείξης και η αξιολόγηση της απόδοσης των υποκειμένων κατά την παιγνιακή εμπειρία.

### 6.2 Μεθοδολογία Αξιολόγησης

Η αξιολόγηση του προσαρμοστικού μοντέλου μείξης, στηρίζεται σε δύο πυλώνες. Ο πρώτος πυλώνας είναι η μέτρηση της απόδοσης και ο δεύτερος η υποκειμενική αξιολόγηση λαμβάνοντας υπόψιν υποκειμενικά κριτήρια. Σχετικά με τον πρώτο πυλώνα σχεδιάστηκε μια μεθοδολογία σύγκρισης αποτελεσμάτων, χρησιμοποιώντας εργαλεία ανάλυσης, και συσχέτισης ακουστικών αποτελεσμάτων τα οποία χρησιμοποιούνται στον ευρύτερο τομέα της στατιστικής ανάλυσης αλλά και κάποια τα οποία χρησιμοποιούνται στον τομέα του ήχου. Αναλυτικά οι διεργασίες αντικειμενικής αξιολόγησης παρουσιάζονται στην Ενότητα 6.2.1.

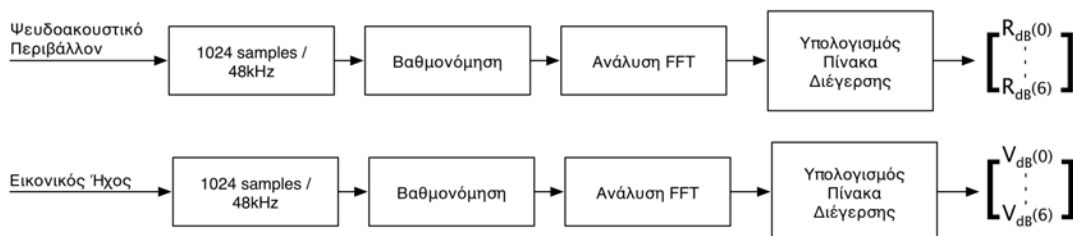
Για τον δεύτερο πυλώνα, αναπτύχθηκε μια μεθοδολογία αξιολόγησης η οποία στηρίζεται σε σύγκριση υποκειμενικών κριτηρίων ήχου από υποκείμενα με υπάρχουσες μεθόδους κοινά αποδεκτές σε διεθνή επίπεδο. Επιπλέον, δίνεται η αξιολόγηση του προσαρμοστικού μοντέλου μέσα από την χρήση του σε ένα ηχητικό παιχνίδι επαυξημένης πραγματικότητας με στοιχεία

διάδρασης, παρουσιάζοντας μια ενδεικτική εφαρμοσμένη χρήση του.

### 6.2.1 Αντικειμενική Αξιολόγηση

Η μέτρηση της απόδοσης του μοντέλου, αποτελεί την πρώτη προσέγγιση αξιολόγησης που υλοποιήθηκε με σκοπό την αντικειμενική αποτύπωση της απόκρισης του μοντέλου μέσα στο σύστημα επαυξημένης μείξης ήχου. Η απόκριση του μοντέλου αποτελεί το αποτέλεσμα της πολυεπίπεδης διαδικασίας που λαμβάνει η εικονική πηγή ήχου και κατ' επέκταση το εικονικό περιβάλλον ήχου, μέσα στο ευρύτερο πλαίσιο της επαυξημένης πραγματικότητας ήχου. Η μέτρηση της απόδοσης, αποτελεί μια διαδικασία με την οποία θα παρουσιαστεί το ηχητικό αποτέλεσμα ποσοτικοποιημένο με σκοπό την αξιολόγησή του. Μέσα από την ποσοτικοποίηση, γίνεται εφικτό να πραγματοποιηθούν συγκρίσεις και να βγουν συμπεράσματα ως προς την λειτουργία του μοντέλου προσαρμοστικής μείξης. Παρακάτω παρουσιάζονται τα στάδια που ακολουθούνται για την αξιολόγηση, καθώς και οι τεχνικές που χρησιμοποιούνται για την εξαγωγή αντικειμενικών αποτελεσμάτων. Η παρουσίαση των σταδίων, δομείται ακολουθώντας την πορεία του εικονικού περιβάλλοντος ήχου από την εισαγωγή του στο μοντέλο μέχρι την εξαγωγή του από αυτό. Η συνολική διαδικασία της αξιολόγησης, διαιρέθηκε σε τμήματα, ανάλογα με την εκάστοτε διεργασία, με σκοπό να υπάρχει ο απόλυτος έλεγχος της διαδικασίας και των μετρήσεων.

Αρχικά, επιλέχθηκαν τα ηχητικά δείγματα τα οποία θα χρησιμοποιηθούν για την αξιολόγηση του μοντέλου. Η επιλογή έγινε έχοντας ως γνώμονα τη χρήση δειγμάτων που ανταποκρίνονται σε μια υποθετική πραγματική χρήση του μοντέλου. Έτσι, ως πραγματικό περιβάλλον, επιλέχθηκαν ηχοτοπία τα οποία συναντάμε στην καθημερινή ζωή. Πιο συγκεκριμένα, επιλέχθηκε το ηχοτοπίο από μια καφετέρια (CoffeeShop) καθώς και το ηχοτοπίο από ένα βιομηχανικό περιβάλλον (Industry). Το εικονικό περιβάλλον ήχου, παρουσιάζεται μέσα από την επιλογή τεχνητών ήχων η οποίοι αποτελούν ηχητικά συμβάντα που προδιαγράφονται κατά την ηχοποίηση (Sonification) σεναρίων επαυξημένης πραγματικότητας ήχου και ειδικότερα ηχητικών παιχνιδιών. Ο πρώτος εικονικός ήχος μπορεί να περιγραφεί ως κραυγή από ένα ρομπότ (κωδικό όνομα: Roboscream) και ο δεύτερος ως ήχος νίκης (κωδικό όνομα: ScoreTrim).

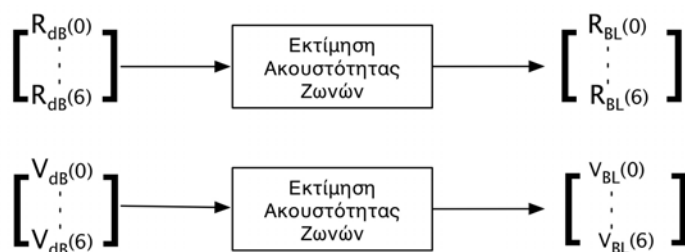


Σχήμα 6.1: Διάγραμμα Μεθόδου δημιουργίας δειγμάτων και υπολογισμός πίνακα διέγερσης περιβαλλόντων για υλοποίηση μετρήσεων

Για όλα τα ηχητικά δείγματα ακολουθήθηκε η διαδικασία επιλογής του αριθμού των δειγ-

μάτων για τα οποία θα δοκιμαστεί το μοντέλο. Ο αριθμός των δειγμάτων είναι σύμφωνος με τον μέγιστο χρόνο ενσωμάτωσης της ακουστότητας του οποίου η διάρκεια είναι 200 ms. Επιλέχθηκε αυτή η παράμετρος κατά την αντικειμενική αξιολόγηση για να υπάρχει συνάφεια στην χρήση του μοντέλου σε πραγματικό χρόνο και στην διεργασία της αξιολόγησης. Το παράθυρο των 9600 δειγμάτων ή αλλιώς 200 ms αποτελεί κοινή συνισταμένη για όλα τα ηχητικά δείγματα, καθώς και ο ρυθμός δειγματοληψίας, ο οποίος καθορίζεται με τακτική όμοια με εκείνη του αριθμού των δειγμάτων, στα 48 kHz.

Ακολουθώντας, σε κάθε περίπτωση, διενεργείται βαθμονόμηση των ηχητικών δειγμάτων. Οι τιμές βαθμονόμησης για κάθε ηχητικό δείγμα τέθηκαν στο επίπεδο των 82 dB SPL. Τα ηχητικά δείγματα αναλύονται με την χρήση του γρήγορου μετασχηματισμού Fourier. Στη συνέχεια, υπολογίζεται ο πίνακας διέγερσης ακολουθώντας την υλοποίηση που περιγράφηκε στην Ενότητα 5.2.3. Τέλος, σε αυτό το στάδιο εξάγονται οι τιμές διέγερσης για κάθε ζώνη συχνοτήτων για κάθε περίπτωση βαθμονομημένου ηχητικού δείγματος.

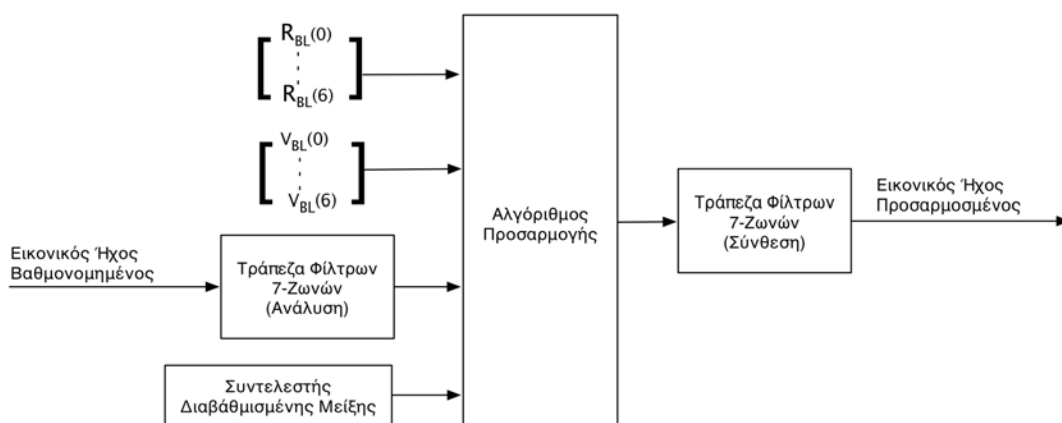


Σχήμα 6.2: Διάγραμμα Μεθόδου Εκτίμησης Ακουστότητας από πίνακα Διέγερσης

Έχοντας ως δεδομένο τον πίνακα διέγερσης του εκάστοτε περιβάλλοντος, γίνεται η εκτίμηση της ακουστότητας για κάθε ζώνη χρησιμοποιώντας την τεχνική που περιγράφεται στην Ενότητα 5.2.6. Οι τιμές ακουστότητας για κάθε ζώνη του εκάστοτε ηχητικού δείγματος αποθηκεύονται για να χρησιμοποιηθούν στο επόμενο στάδιο.

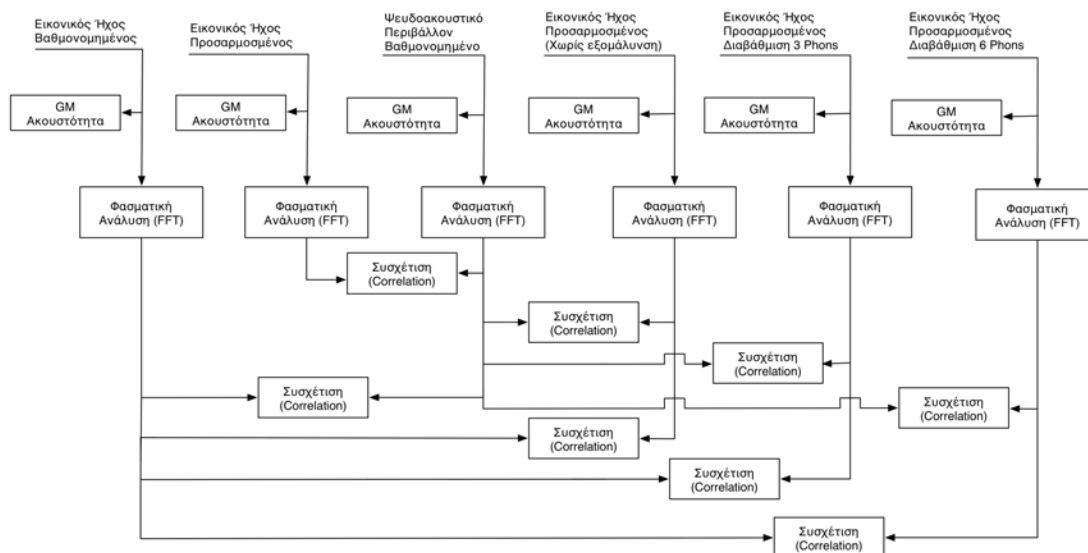
Αφού γίνει η εκτίμηση της ακουστότητας, ακολουθεί η λειτουργία του αλγόριθμου προσαρμογής. Ο αλγόριθμος προσαρμογής, λειτουργεί σύμφωνα με την Ενότητα 5.2.7. Ως είσοδο στην κάθε περίπτωση προσαρμογής εικονικού με ψευδοακουστικό περιβάλλον, χρησιμοποιείται ο πίνακας εκτιμώμενης ακουστότητας των επτά συχνοτικών ζωνών. Στην περίπτωση της αντικειμενικής αξιολόγησης, προκαθορίστηκαν ζευγάρια ήχων τα οποία παρουσιάζονται στην Ενότητα 7. Για το κάθε ζευγάρι ήχου, εισάγεται ο εικονικός ήχος στον αλγόριθμο προσαρμογής, αφού προηγουμένως έχει αναλυθεί σε ζώνες από την τράπεζα φίλτρων. Στην περίπτωση της αξιολόγησης, θα εξεταστεί η διαβάθμιση της μείζης μέσα από την επίδραση του συντελεστή διαβαθμισμένης μείζης.

Με αυτό τον τρόπο θα μετρηθεί η απόδοση του μοντέλου σε μια τέτοια αλλαγή αντικειμενικά. Ο συντελεστής εισάγεται σύμφωνα με την Εξίσωση 4.28 και μεταβάλλει τη διαφορά μεταξύ των συχνοτικών ζωνών του πραγματικού περιβάλλοντος με το εικονικό. Οι επιλογές της στα-



Σχήμα 6.3: Διάγραμμα Μεθόδου Μετρήσεων Αλγόριθμου Προσαρμογής

θεράς διαβάθμισης που ορίστηκαν για δοκιμή είναι η «Διαβάθμιση 0» όπου ως συντελεστής ορίζεται η τιμή 0 και το αποτέλεσμα είναι η αύξηση της διαφοράς μεταξύ του εικονικού με το πραγματικό περιβάλλον μέχρι 3 *Phons*, η «Διαβάθμιση 3» στην οποία ο συντελεστής έχει τιμή 3 και η διαφορά που δίνεται επιπλέον μεταξύ πραγματικού με εικονικό περιβάλλον φτάνει μέχρι 6 *Phons* και η «Διαβάθμιση 6» κατά την οποία ο συντελεστής παίρνει τιμή 6 και ως αποτέλεσμα η διαφορά μεταξύ εικονικού και πραγματικού περιβάλλοντος φτάνει τα 9 *Phons*.



Σχήμα 6.4: Διάγραμμα Μεθόδου Ανάλυσης μετρήσεων μοντέλου προσαρμογής

Για την ανάλυση των ηχητικών αποτελεσμάτων του μοντέλου, ακολουθείται η μεθοδολογία που απεικονίζεται στο Σχήμα 6.4. Σύμφωνα με αυτό, το κάθε ζευγάρι εικονικού ήχου - ψευδοακουστικού περιβάλλοντος, αξιολογείται βασιζόμενο σε 3 τύπους ανάλυσης: α) Ανάλυση

ακουστότητας σύμφωνα με το μοντέλο GM, β) Φασματική Ανάλυση (FFT), γ) Συσχέτιση φασματικής ανάλυσης μεταξύ των ηχητικών δειγμάτων πριν την επεξεργασία τους από τον αλγόριθμο προσαρμογής και κατόπιν της προσαρμογής τους.

Αναλυτικότερα, θα υπολογιστεί η συσχέτιση του εικονικού περιβάλλοντος με το πραγματικό πριν και μετά την εφαρμογή της προσαρμοστικής μείξης. Για να επιτευχθεί αυτό, αρχικά θα γίνει η φασματική ανάλυση τύπου FFT και κατόπιν για τα αποτελέσματα της ανάλυσης θα γίνει μέτρηση της συσχέτισης μεταξύ τους. Επίσης, θα συγκριθούν οι διαφορετικές λειτουργίες του αλγόριθμου προσαρμογής σε σχέση με τη διαβαθμισμένη λειτουργία της μείξης σε φασματικό επίπεδο καθώς και σε επίπεδο ακουστότητας για την καλύτερη κατανόηση της λειτουργίας του μοντέλου. Ο υπολογισμός της ακουστότητας κατά το μοντέλο GM θα πραγματοποιηθεί σε όλα τα ηχητικά δείγματα για επιπλέον συγκρίσεις. Τα αποτελέσματα της αξιολόγησης, παρουσιάζονται στην Ενότητα 7.2 μαζί με την ανάλυσή τους.

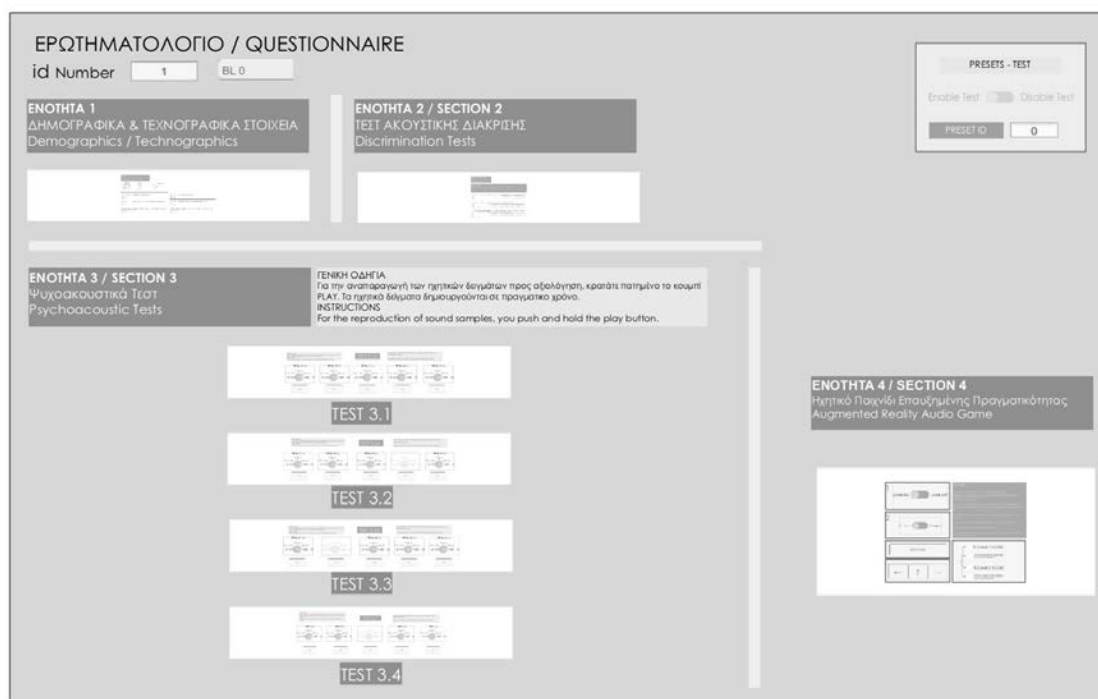
### 6.2.2 Υποκειμενική Αξιολόγηση

Η αξιολόγηση του μοντέλου μείξης από υποκείμενα αποτελεί το ζητούμενο σε αυτήν την ενότητα, καθώς πρόκειται για ένα αντιληπτικό μοντέλο μείξης, το οποίο είναι αδύνατο η αξιολόγησή του να βασιστεί μόνο σε αντικειμενικές μετρήσεις απόδοσης. Γι'αυτό το λόγο, οργανώθηκε μια μεθοδολογία υποκειμενικής αξιολόγησης του μοντέλου από υποκείμενα μέσα από μια αλληλουχία δοκιμών.

Η αξιολογητική διαδικασία σχεδιάστηκε μέσα σε εργαστηριακό περιβάλλον. Με αυτό τον τρόπο είναι εφικτό να υπάρχει πλήρης έλεγχος των διάφορων παραμέτρων που συντελούν στη λειτουργία του μοντέλου προσαρμοστικής μείξης. Επιπλέον, τέθηκε ως στόχος η αξιολόγηση του μοντέλου μέσα από τη λειτουργία του σε πραγματικό χρόνο, ώστε να υπάρχει ο βέλτιστος συγχρονισμός ψευδοακουστικού περιβάλλοντος με το προσαρμοσμένο από το μοντέλο μείξης. Σύμφωνα με αυτό τον στόχο, σχεδιάστηκε η μορφή της αξιολόγησης, τα ερωτήματα που τίθενται άμεσα και έμμεσα, οι τύποι αξιολόγησης που χρησιμοποιούνται και όλα αυτά υπό την μορφή ενός ηλεκτρονικού ερωτηματολογίου με έμφαση στην ευχρηστία και στην συγκέντρωση του υποκειμένου στον τελικό στόχο της αξιολόγησης.

Η διαίρεση του ηλεκτρονικού ερωτηματολογίου γίνεται σε τέσσερα μέρη (βλπ Σχ. 6.5). Αρχικά, το πρώτο μέρος περιλαμβάνει τη συμπλήρωση δημογραφικών και τεχνογραφικών στοιχείων, το δεύτερο μέρος περιέχει ψυχοακουστικά τεστ για την διάκριση του υποκειμένου σε κατηγορία σχετικά με την εμπειρία του με τον ήχο, το τρίτο μέρος περιέχει την αξιολόγηση της προσαρμοστικής μείξης μέσα από μια ακολουθία τεστ και στο τελικό μέρος την διεξαγωγή ενός ηχητικού παιχνιδιού επαυξημένης πραγματικότητας ήχου το οποίο αναλύεται στην επόμενη Ενότητα (βλπ Εν. 6.2.3).

Το υποκείμενο στην πρώτη ενότητα όπως παρουσιάζεται στο Σχήμα 6.6, καλείται να δώσει στοιχεία σχετικά με το φύλο του, την ηλικία του και την ασχολία του. Κατόπιν στα τεχνογραφικά στοιχεία, απαντάει σε ερωτήσεις με μονολεκτική απάντηση τύπου Ναι/Όχι σχετικά με τη χρήση νέων τεχνολογιών, την ενασχόλησή του με τη μουσική, την εμπειρία του με ηλεκτρονικές ηχητικές εφαρμογές, την σχέση του με την ψηφιακή επεξεργασία ήχου, τη χρήση



Σχήμα 6.5: Κύρια οθόνη ηλεκτρονικού ερωτηματολογίου

εφαρμογών Επαυξημένης πραγματικότητας Ήχου και τη χρήση συστημάτων προσαρμοστικής ηχητικής μείξης.

Στο δεύτερο μέρος όπως προαναφέρθηκε, το υποκείμενο υπόκειται σε τρία ψυχοακουστικά τέστ σύμφωνα με τα οποία το υποκείμενο δοκιμάζεται ως προς την ικανότητά του να κατανοεί διαφορές ακουστότητας [8], να διαχωρίζει ήχους διαφορετικής τονικότητας, και να αντιλαμβάνεται τον τρισδιάστατο ηχητικό χώρο. Σύμφωνα με το αποτέλεσμα του τεστ διάκρισης, γίνεται η κατάταξή του, σε κατηγορία ακουστικής εμπειρίας (βλπ Σχ. 6.7). Στο πρώτο ψυχοακουστικό τέστ, το υποκείμενο συγκρίνει δύο ήχους ως προς την ακουστότητά τους. Ο πρώτος ήχος είναι ένας τόνος  $1\text{ kHz}$  με ψηφιακό πλάτος  $0\text{ dBFS}$  και ο δεύτερος ένας τόνος  $3\text{ kHz}$  με ψηφιακό πλάτος  $0\text{ dBFS}$ . Στο δεύτερο ψυχοακουστικό τεστ, το υποκείμενο πρέπει να διακρίνει ανάμεσα σε δύο ήχους ποιος περιλαμβάνεται σε ένα τρίτο ηχητικό δείγμα. Ο πρώτος ήχος (A) είναι ένας τόνος  $1\text{ kHz}$  με ψηφιακό πλάτος  $0\text{ dBFS}$ , ο δεύτερος ήχος είναι ένας τόνος  $500\text{ Hz}$  με ψηφιακό πλάτος  $0\text{ dBFS}$  και η τελική ηχητική σύνθεση αποτελείται από τον τόνο των  $500\text{ Hz}$  και έναν τόνο  $3\text{ kHz}$  με ψηφιακό πλάτος  $0\text{ dBFS}$ . Στο τρίτο ψυχοακουστικό τεστ, το υποκείμενο καλείται να βρει ποιος από τους ήχους τύπου «A» και «B» βρίσκεται στο τρισδιάστατο εικονικό ηχητικό χώρο του «X».

**ΕΝΟΤΗΤΑ 1**  
ΔΗΜΟΓΡΑΦΙΚΑ & ΤΕΧΝΟΓΡΑΦΙΚΑ ΣΤΟΙΧΕΙΑ  
Demographics / Technographics

1.1 GENDER/ΦΥΛΟ  
☐ MALE/ΑΝΔΡΑΣ  
☒ FEMALE/ΓΥΝΑΙΚΑ  
☐ OTHER/ΆΛΛΟ

1.2 AGE/ΗΛΙΚΙΑ  
☒ <18  
☐ 18-25  
☐ 25-35  
☐ >35

1.3 OCCUPATION / ΑΙΣΧΟΛΙΑ  
☒ PUPIL/MΑΘΗΤΗΣ  
☐ STUDENT /ΣΠΟΥΔΑΣΤΗΣ  
☐ OTHER/ΆΛΛΟ

---

1.4 Μου αρέσει η χρήση νέων τεχνολογιών / I like using new technologies  
☒ ΝΑΙ / YES  
☐ ΟΧΙ / NO

1.5 Είμαι σχετικός με μουσική / I have a musical background  
☒ ΝΑΙ / YES  
☐ ΟΧΙ / NO

1.6 Είμαι σχετικός με ηλεκτρονικές ηχητικές εφαρμογές / I am experienced with electronic music applications  
☒ ΝΑΙ / YES  
☐ ΟΧΙ / NO

1.7 Είμαι σχετικός με ψηφιακή επεξεργασία ήχου / I am experienced with digital signal processing  
☒ ΝΑΙ / YES  
☐ ΟΧΙ / NO

1.8 Έχω χρησιμοποιήσει στο παρελθόν εφαρμογή Επαυξημένης Ηχητικής Προσαρμοστικότητας / In the past I have used Augmented Reality Audio application  
☐ ΝΑΙ / YES  
☒ ΟΧΙ / NO

1.9 Έχω χρησιμοποιήσει στο παρελθόν σύστημα προσαρμοσμένης ηχητικής μίξης / In the past I have used a system of audio adaptive mix  
☐ ΝΑΙ / YES  
☒ ΟΧΙ / NO

Σχήμα 6.6: Οθόνη Τεχνογραφικών και Δημογραφικών στοιχείων ηλεκτρονικού ερωτηματολογίου

**ΕΝΟΤΗΤΑ 2 / SECTION 2**  
ΤΕΣΤ ΑΚΟΥΣΤΙΚΗΣ ΔΙΑΚΡΙΣΗΣ  
Discrimination Test

**ΓΕΝΙΚΗ ΟΔΗΓΙΑ**

Για την αναπαραγωγή των ηχητικών δειγμάτων προς αξιολόγηση, κρατάτε πατημένο το κουμπί PLAY. Τα ηχητικά δείγματα δημιουργούνται σε πραγματικό χρόνο.

2.1 Ποιον ήχο θεωρείτε πιο δυνατά ακουστό; Which sound you think is louder?  
☒ A  
☐ B

2.2 Ποιος ήχος από τους A & B περιλαμβάνεται στον X; / Which of A & B sounds is included in X?  
☒ A  
☐ B  
☐ Κανένας / None

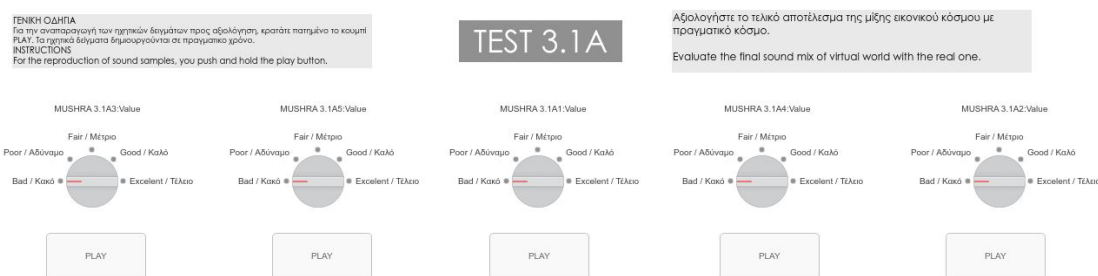
2.3 Ποιος ήχος από τους A & B βρίσκεται αριστερά στον τρισδιάστατο εικονικό ηχητικό χώρο του X; Which of A & B sounds is located on the left of 3d virtual audio space?  
☒ A  
☐ B  
☐ Κανένας/None

Σχήμα 6.7: Οθόνη Ψυχοακουστικών τεστ Ακουστικής Διάκρισης

Ο αριθμός των σωστών απαντήσεων, κατηγοριοποιεί το υποκείμενο σε βαθμίδα ακουστικής εμπειρίας [8]. Έχοντας απαντηθεί σωστά μια ερώτηση, τότε η βαθμίδα ακουστικής εμπειρίας αντιστοιχεί στον «Αρχάριο» (Beginner), στην περίπτωση των δύο σωστών απαντήσεων, η

βαθμίδα ακουστικής εμπειρίας αντιστοιχεί στον «Έμπειρο» (Experienced), στην περίπτωση όλων των σωστών απαντήσεων, το υποκείμενο κατηγοριοποιείται ως «Ειδικός» (Expert) και στην περίπτωση μη σωστής απάντησης, θεωρείται ως «Αφελής» (Naive).

Στην τρίτη ενότητα του ηλεκτρονικού ερωτηματολογίου, το υποκείμενο, υπόκειται σε υποκειμενικά τέστ σύγκρισης του μοντέλου ηχητικής μείξης ακολουθώντας την μεθοδολογία των τέστ τύπου MUSHRA (Multi Stimulus Hidden Reference) [108]. Μέσα από αυτό τον τύπο τεστ (βλπ Σχ. 6.8), ο ακροατής έχει την δυνατότητα αξιολόγησης πέντε διαφορετικών εκδοχών επαυξημένης μείξης ήχου με συγκεκριμένες προδιαγραφές, χωρίς να γνωρίζει το δείγμα αναφοράς. Η πρώτη μείξη, αποτελείται από τον ήχο του εικονικού περιβάλλοντος ο οποίος διέρχεται από το σύστημα βαθμονόμησης, καθώς και από την τράπεζα φίλτρων ανάλυσης και σύνθεσης χωρίς να γίνεται καμιά επιπλέον επεξεργασία σε αυτόν. Η δεύτερη εκδοχή αποτελεί την προσαρμοστική μείξη ήχου επαυξημένης πραγματικότητας, στην οποία η εικονική πηγή ήχου προσαρμόζεται στο πραγματικό περιβάλλον σύμφωνα με το προτεινόμενο μοντέλο. Οι υπόλοιπες τρεις εκδοχές αποτελούν παραλλαγές της πρώτης εκδοχής με μεταβαλλόμενα τα συχνοτικά εύρη τα οποία κινούνται αντίστοιχα στα 3.5 kHz, 7 kHz και 10 kHz.



Σχήμα 6.8: Οθόνη Υποκειμενικών Τεστ τύπου MUSHRA

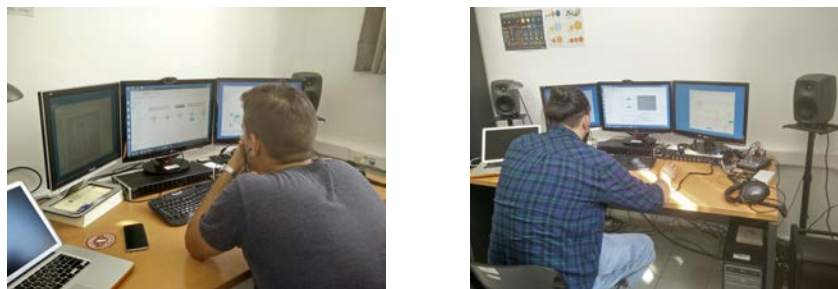
Ο ακροατής, πρέπει να αξιολογήσει το τελικό αποτέλεσμα κάθε εκδοχής της μείξης του εικονικού κόσμου με τον πραγματικό. Η κλίμακα αξιολόγησης του αποτελέσματος της κάθε μείξης, τέθηκε μέσα από μια κλίμακα πέντε επιπέδων: Κακό, Αδύναμο, Μέτριο, Καλό, Τέλειο.

Η αναπαραγωγή του πραγματικού περιβάλλοντος, γίνεται μέσα από την χρήση ηχείων (Genelec 8040B) τα οποία είναι τοποθετημένα αριστερά και δεξιά από τον ακροατή σε απόσταση 1 μέτρου από τον ίδιο, έχοντας 2 μέτρα απόσταση μεταξύ τους. Η βαθμονόμηση του πραγματικού περιβάλλοντος έγινε με την χρήση ηχόμετρου στην θέση του κεφαλιού του υποκειμένου σύμφωνα με το οποίο καθορίστηκαν οι στάθμες του πραγματικού περιβάλλοντος για κάθε MUSHRA τεστ.



Σχήμα 6.9: Βαθμονόμηση πραγματικού περιβάλλοντος αξιολόγησης

Η διαδικασία της αξιολόγησης είχε ατομικό χαρακτήρα. Για την διεξαγωγή της, παραχωρήθηκε έγγραφη άδεια από την επιτροπή δεοντολογίας του Ιονίου Πανεπιστημίου (βλπ Παράρτημα Δ' Σχ.Δ'.1). Το κάθε υποκείμενο εισήλθε ατομικά στο δωμάτιο δοκιμής. Για την συμμετοχή του, συμπλήρωσε τη δήλωση συγκατάθεσης (βλπ Παράρτημα Δ' Σχ.Δ'.2) η οποία ήταν διαθέσιμη και σε αγγλική γλώσσα (βλπ Παράρτημα Δ' Σχ.Δ'.3), και κάθισε στη θέση ακρόασης. Πρέπει να σημειωθεί πως δεν υπήρξε διαδικασία προετοιμασίας και εκπαίδευσης του υποκειμένου σε καμία ενότητα του ερωτηματολογίου ώστε το δείγμα να είναι αμερόληπτο (no bias). Οι μόνες οδηγίες που δόθηκαν ήταν καθαρά και μόνο για την περιήγηση στο ερωτηματολόγιο και για τον τρόπο χρήσης του. Δεν τέθηκε χρονικός περιορισμός για την διεξαγωγή της αξιολόγησης ούτε περιορισμός στον τρόπο συμπλήρωσής του. Επιπροσθέτως, το ερωτηματολόγιο, ήταν ανώνυμο έχοντας μόνο έναν αύξοντα αριθμό. Επίσης, για την αξιολόγηση του συντελεστή διαβαθμισμένης μείξης, πριν την έναρξη της αξιολόγησης, ρυθμιζόταν σε μια σταθερή τιμή για κάθε υποκείμενο. Οι προεπιλεγμένες τιμές διαβάθμισης που επιλέχθηκαν ήταν οι ακόλουθες: «Διαβάθμιση 0», «Διαβάθμιση 3» και «Διαβάθμιση 6». Συνεπώς εκτός από τις κατηγορίες σύμφωνα με την εμπειρία του ακροατή, υπάρχουν και οι κατηγορίες σύμφωνα με την διαβάθμιση της προσαρμοστικής μείξης.



Σχήμα 6.10: Διαδικασία Υποκειμενικής Αξιολόγησης

Τα επαυξημένα περιβάλλοντα που υλοποιούνται μέσα από τα τέστ τύπου MUSHRA, περιγράφονται παρακάτω:

Τέστ MUSHRA με κωδικό 3.1: Το πραγματικό περιβάλλον παρουσιάζεται ως ήχος βιομηχανίας βαθμονομημένος στα  $82\text{ dB SPL}$  και το εικονικό περιβάλλον παρουσιάζεται ως εικονικός ήχος τύπου Roboscream.

Τέστ MUSHRA με κωδικό 3.2: Το πραγματικό περιβάλλον παρουσιάζεται ως ήχος βιομηχανίας με βαθμονόμηση στα  $82\text{ dB SPL}$  και το εικονικό περιβάλλον ως δοκιμαστικός ήχος με την επωνυμία Miss (Ψηφιακός ήχος αποτυχίας ηχητικού παιχνιδιού).

Τέστ MUSHRA με κωδικό 3.3: Το πραγματικό περιβάλλον αναπαράγεται ως περιβάλλον δρόμου πόλης με κίνηση βαθμονομημένο στα  $82\text{ dB SPL}$  και το εικονικό περιβάλλον αποτελεί τον ήχο με το όνομα Score (Ψηφιακός ήχος νίκης ηχητικού παιχνιδιού).

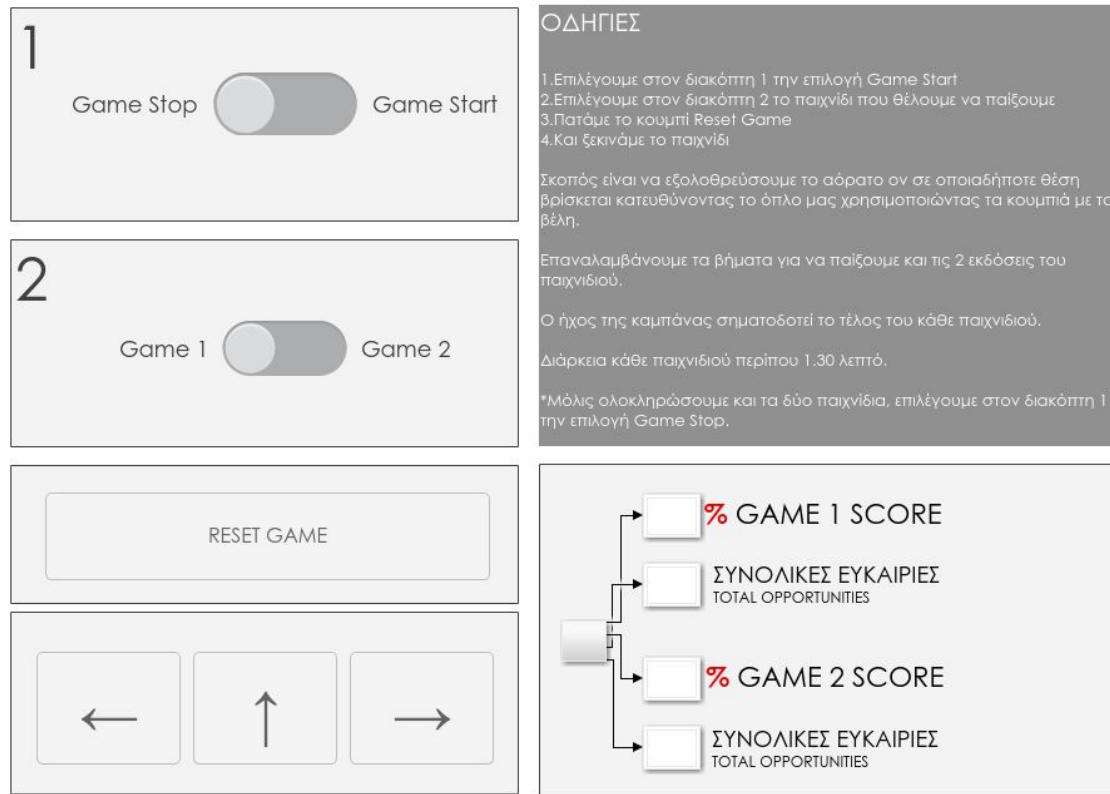
Τέστ MUSHRA με κωδικό 3.4: Το πραγματικό περιβάλλον παρουσιάζεται ως καφετέρια με όχληση βαθμονομημένη στα  $82\text{ dB SPL}$  και το εικονικό περιβάλλον αποτελείται από τον ήχο με κωδικό όνομα Score.

Για τις παραπάνω περιπτώσεις, τίθεται ως αρχική υπόθεση (Null Hypothesis) ότι η προσαρμοστική μείξη δεν είναι αρεστή στο ανθρώπινο αυτί. Ως εναλλακτική υπόθεση, τίθεται πως η προσαρμοστική μείξη ήχου βελτιώνει την ποιότητα της επαυξημένης μείξης ήχου. Τα αποτελέσματα της παραπάνω αξιολόγησης, παρουσιάζονται στην Ενότητα 7.3.

### 6.2.3 Περίπτωση Ηχητικού Παιχνιδιού τύπου ARA

Στην τελευταία ενότητα της αξιολόγησης του μοντέλου, παρουσιάζεται η περίπτωση ενός ηχητικού παιχνιδιού επαυξημένης πραγματικότητας. Η ιδιαιτερότητα του συγκεκριμένου παιχνιδιού είναι η χρήση του μοντέλου προσαρμοστικής μείξης ήχου σε όλη την διάρκεια της παιγνιακής διαδικασίας. Ο ακροατής λαμβάνει μέρος σε μια παιγνιακή εμπειρία σε εργαστηριακό επίπεδο μέσα σε ένα περιβάλλον ελεγχόμενο. Το σενάριο του ηχητικού παιχνιδιού αποτελεί την παραλλαγή του ηχητικού παιχνιδιού που υλοποιήθηκε στα πλαίσια του ερευνητικού έργου 'Ηχητικά Παιχνίδια Επαυξημένης Πραγματικότητας' [111, 96]. Πιο συγκεκριμένα, το σενάριο του ηχητικού παιχνιδιού που υλοποιήθηκε, είναι το θαύμα του Αγίου με το οποίο ο Άγιος Σπυρίδων της Κέρκυρας έδιωξε την πανούκλα από το νησί. Στην παραλλαγή που υλοποιήθηκε για τον σκοπό της αξιολόγησης, ο παίκτης έχει προκαθορισμένη θέση και η διάδραση επιτυγχάνεται μέσα από τον χειρισμό μιας εικονικής σκανδάλης η οποία στοχεύει προς τρεις

κατευθύνσεις. Ο εικονικός ήχος της 'Πανούκλας' ο οποίος αποτελείται από έναν συνθετικό ήχο μεσαίου συχνοτικού φάσματος υπό τη μορφή ψιθύρου, ακολουθεί προκαθορισμένη τροχιά στον τρισδιάστατο ηχητικό χώρο. Στο Σχήμα 6.11, παρουσιάζεται η οθόνη στην οποία ο θεατής χειρίζεται την κατεύθυνση της εικονικής σκανδάλης με το πάτημα των κουμπιών με το εικονίδιο του βέλους.



Σχήμα 6.11: Οθόνη Ηχητικού παιχνιδιού Επαυξημένης Πραγματικότητας Ήχου

Σκοπός του παιχνιδιού, είναι η επιτυχής στόχευση και ενεργοποίηση της εικονικής σκανδάλης τη στιγμή που βρίσκεται ο εικονικός ήχος της «Πανούκλας» στην αντίστοιχη θέση μπροστά, αριστερά ή δεξιά με ένα εύρος 10% γύρω από τον στόχο. Η τροχιά που διαγράφει ο εικονικός ήχος είναι ίδια για όλους τους παίκτες για να υπάρχει κοινό μέτρο σύγκρισης καθώς και ο χρόνος διεξαγωγής του ηχητικού παιχνιδιού είναι τα 100 δευτερόλεπτα. Ο παίκτης, παίζει δύο εκδοχές παιχνιδιού διαφορετικής ηχητικής μείξης. Στην μία εκδοχή δεν υπάρχει η χρήση του προσαρμοστικού μοντέλου μείξης, ενώ στην δεύτερη, χρησιμοποιείται. Η σειρά με την οποία ο παίκτης παίζει τις δύο ηχητικές εκδοχές, είναι τυχαία και επίσης δεν προηγείται καμιά μορφή εκπαίδευσης του παίκτη.

Το πραγματικό περιβάλλον το οποίο επιλέχθηκε για την διεξαγωγή του ηχητικού παιχνιδιού επαυξημένης πραγματικότητας, είναι το ηχητικό περιβάλλον της πόλης σε ώρα κίνησης αναπαραγόμενο με την ίδια μεθοδολογία που ακολουθείται κατά την διάρκεια της υποκειμενικής αξιολόγησης του μοντέλου με βαθμονόμηση στα 85 dB SPL. Η επιλογή του συγκεκριμένου τύπου ήχου, επιλέχθηκε λόγω των έντονων διακυμάνσεων που περιέχονται μέσα σε αυτόν,

καθώς στόχος του προσαρμοστικού μοντέλου μείξης, είναι η βέλτιστη επαυξημένη μείξη ειδικά σε τέτοια πραγματικά περιβάλλοντα ήχου με έντονες διακυμάνσεις.

Για την αξιολόγηση του ηχητικού παιχνιδιού, δημιουργήθηκαν μετρητές καταγραφής συγχεκριμένων συμβάντων. Αρχικά, δημιουργήθηκε ο μετρητής του χρόνου μέσα στον οποίο ο παίκτης καλείται να πετύχει το αόρατο ον της 'Πανούκλας'. Ο δεύτερος μετρητής, καταγράφει τις χαμένες ευκαιρίες που είχε ο θεατής για να πετύχει το εχθρικό ον καθώς πέρασε από κοντά του αλλά δεν το πέτυχε. Ο τρίτος μετρητής, καταγράφει τις επιτυχημένες βολές του παίκτη και τέλος υπολογίζεται το ποσοστό επιτυχίας, λαμβάνοντας υπόψιν το λόγο των επιτυχημένων βολών ως προς τις συνολικές ευκαιρίες που είχε ο παίκτης κατά τη διάρκεια του παιχνιδιού [93].

Το ηχητικό παιχνίδι επαυξημένης πραγματικότητας, αποτελεί το τελευταίο κομμάτι του ηλεκτρονικού ερωτηματολογίου για την αξιολόγηση του προτεινόμενου μοντέλου προσαρμοστικής μείξης σε περιβάλλοντα επαυξημένης πραγματικότητας ήχου. Τα αποτελέσματα της αξιολόγησης καταγράφονται στην Ενότητα 7.4.

### 6.3 Σύνοψη Κεφαλαίου

Σε αυτό το κεφάλαιο, παρουσιάστηκε η μεθοδολογία που σχεδιάστηκε και υλοποιήθηκε για την αξιολόγηση του μοντέλου προσαρμοστικής μείξης. Αρχικά εισάγεται η μεθοδολογία της αντικειμενικής αξιολόγησης του μοντέλου. Για την υλοποίησή της, το προσαρμοστικό μοντέλο, διαίρεται σε τμήματα για την καταγραφή της απόκρισής του σε κάθε τμήμα του. Βασικά εργαλεία που επιλέχθηκαν για την αντικειμενική αξιολόγηση των τελικών αποτελεσμάτων, είναι η μέτρηση της ακουστότητας σύμφωνα με το μοντέλο GM, καθώς και ο υπολογισμός της συσχέτισης μεταξύ των μειξαρισμένων περιβαλλόντων σε συχνотικό επίπεδο. Εκτός από την αντικειμενική αξιολόγηση, σχεδιάστηκε και υλοποιήθηκε υποκειμενική αξιολόγηση με τη χρήση υποκειμένων. Για την υποκειμενική αξιολόγηση, δημιουργήθηκε ένα ηλεκτρονικό ερωτηματολόγιο το οποίο αποτελείται από τέσσερα μέρη. Βασικό στοιχείο του ερωτηματολογίου είναι η διάκριση του υποκειμένου ανάλογα με την εμπειρία του μέσα από ψυχοακουστικά τεστ, καθώς και η διεξαγωγή ακουστικών τεστ τύπου MUSHRA για την αξιολόγηση του προσαρμοστικού μοντέλου επαυξημένης μείξης ήχου. Το τελευταίο μέρος του ερωτηματολογίου, αποτελεί την εφαρμοσμένη χρήση του προσαρμοστικού μοντέλου μείξης σε ένα ηχητικό παιχνίδι επαυξημένης πραγματικότητας σε εργαστηριακό περιβάλλον. Τα αποτελέσματα από την συνολική αξιολογητική διαδικασία, παρουσιάζονται στο επόμενο κεφάλαιο.

## Κεφάλαιο 7

# Αποτελέσματα

### 7.1 Εισαγωγή

Το πρώτο μέρος της αντικειμενικής και υποκειμενικής αξιολόγησης το οποίο περιλαμβάνει τη μεθοδολογία και την υλοποίηση, ολοκληρώθηκε στο προηγούμενο κεφάλαιο. Σε αυτό το κεφάλαιο, παρουσιάζονται τα αποτελέσματα των αντικειμενικών καθώς και των υποκειμενικών μετρήσεων με τις αντίστοιχες αναλύσεις τους. Η ανάλυση των αποτελεσμάτων, στοχεύει στην ορθή κατανόηση της απόδοσης του μοντέλου με την εξαγωγή αντικειμενικών δεικτών στην πρώτη φάση της αντικειμενικής αξιολόγησης και την αποτύπωση της επίδρασης στον ακροατή μέσα από τα αποτελέσματα της υποκειμενικής αξιολόγησης.

### 7.2 Αποτελέσματα αντικειμενικών μετρήσεων

Η μεθοδολογία υλοποίησης των αντικειμενικών μετρήσεων, παρουσιάστηκε στην Ενότητα 6.2.1. Σε αυτή την ενότητα, θα παρουσιαστούν τα αποτελέσματα των μετρήσεων επιλεγμένων εκδοχών σεναρίου ηχητικής μείξης επαυξημένης πραγματικότητας συμπεριλαμβανομένου και του προσαρμοστικού μοντέλου μείξης.

Οι περιπτώσεις παραδειγμάτων εφαρμογής του μοντέλου προσαρμοστικής επαυξημένης μείξης για τις οποίες έγιναν μετρήσεις παρουσιάζονται παρακάτω με τον αντίστοιχο κωδικό σεναρίου. Οι περιπτώσεις είναι οι ίδιες για τις οποίες γίνεται η υποκειμενική αξιολόγηση τύπου MUSHRA (βλπ Εν. 6.2.2). Η επιλογή των ήχων έγινε λαμβάνοντας υπόψιν τις ανάγκες της κατηγορίας των ηχητικών παιχνιδιών επαυξημένης πραγματικότητας τα οποία αποτελούν την εφαρμογή στόχο της παρούσας διατριβής. Τα περιβάλλοντα που επιλέχθηκαν, περιέχουν έντονες διακυμάνσεις με διαφορετικό συχνотικό περιεχόμενο. Ως προς το εικονικό περιβάλλον, επιλέχθηκαν συνθετικοί ήχοι, ένα χαρακτηριστικό που συναντάται στα ηχητικά παιχνίδια τύπου ARA. Στο Παράρτημα Γ' παρουσιάζονται για όλα τα ηχητικά δείγματα που χρησιμοποιήθηκαν στην αξιολόγηση, η απεικόνιση της κυματομορφής τους και η φασματική ανάλυση τους.

Σενάριο 3.1: Επαυξημένη μείξη στην οποία ως πραγματικό περιβάλλον παρουσιάζεται ο ήχος βιομηχανίας και ως εικονικό περιβάλλον, ο ήχος με όνομα Roboscream. Ο ήχος βιομη-

χανίας σύμφωνα με την κυματομορφή του (βλπ Παράρτημα Γ', Σχ. Γ'.5), φαίνεται να παρουσιάζει μια επαναληπτικότητα λόγω της λειτουργίας μηχανημάτων. Το συχνοτικό περιεχόμενο του ήχου βιομηχανίας σύμφωνα με τη φασματική ανάλυση (βλπ Παράρτημα Γ', Σχ. Γ'.6), παρουσιάζει υψηλή συγκέντρωση ενέργειας για συχνότητες μικρότερες από  $1\text{ kHz}$  ενώ για συχνότητες πάνω από  $1\text{ kHz}$ , υπάρχει φθίνουσα πορεία της κατανεμημένης ενέργειας. Ο ήχος Roboscream αποτελεί ένα συνθετικό ήχο διάρκειας 2 δευτερολέπτων σύμφωνα με την κυματομορφή του (βλπ Παράρτημα Γ', Σχ.Γ'.9) και αποτελεί έναν σύντομο ήχο προειδοποίησης σαν μια ψηφιακή κραυγή της οποίας το συχνοτικό περιεχόμενο συγκεντρώνεται κυρίως μεταξύ των συχνοτήτων των  $300\text{ Hz}$  και  $2\text{ kHz}$  (βλπ Παράρτημα Γ', Σχ. Γ'.10). Ο ήχος της βιομηχανίας φαίνεται να καλύπτει το συχνοτικό περιεχόμενο του ήχου Roboscream γεγονός που δημιουργεί δυσκολία την κατανόησή του.

Σενάριο 3.2: Ως πραγματικό περιβάλλον παρουσιάζεται ο ήχος βιομηχανίας ομοίως με το σενάριο 3.1 και ως εικονικός, ο ήχος με το όνομα Miss. Ο ήχος Miss, ανήκει στην κατηγορία των συνθετικών ήχων. Η διάρκεια του ήχου σύμφωνα με την κυματομορφή του (βλπ Παράρτημα Γ', Σχ.Γ'.7), είναι 3 δευτερόλεπτα και το περιεχόμενό του είναι ένας συνθετικός ήχος γέλιου. Το συχνοτικό περιεχόμενο του ήχου Miss (βλπ Παράρτημα Γ', Σχ.Γ'.8), συγκεντρώνεται κυρίως μεταξύ των  $200\text{ Hz}$  με  $2\text{ kHz}$  με μεγαλύτερη συγκέντρωση ενέργειας στα  $700\text{ Hz}$ . Ομοίως με το Σενάριο 3.1, ο ήχος της βιομηχανίας φαίνεται να υπερτερεί ως προς συχνοτικό περιεχόμενο του ήχου Miss.

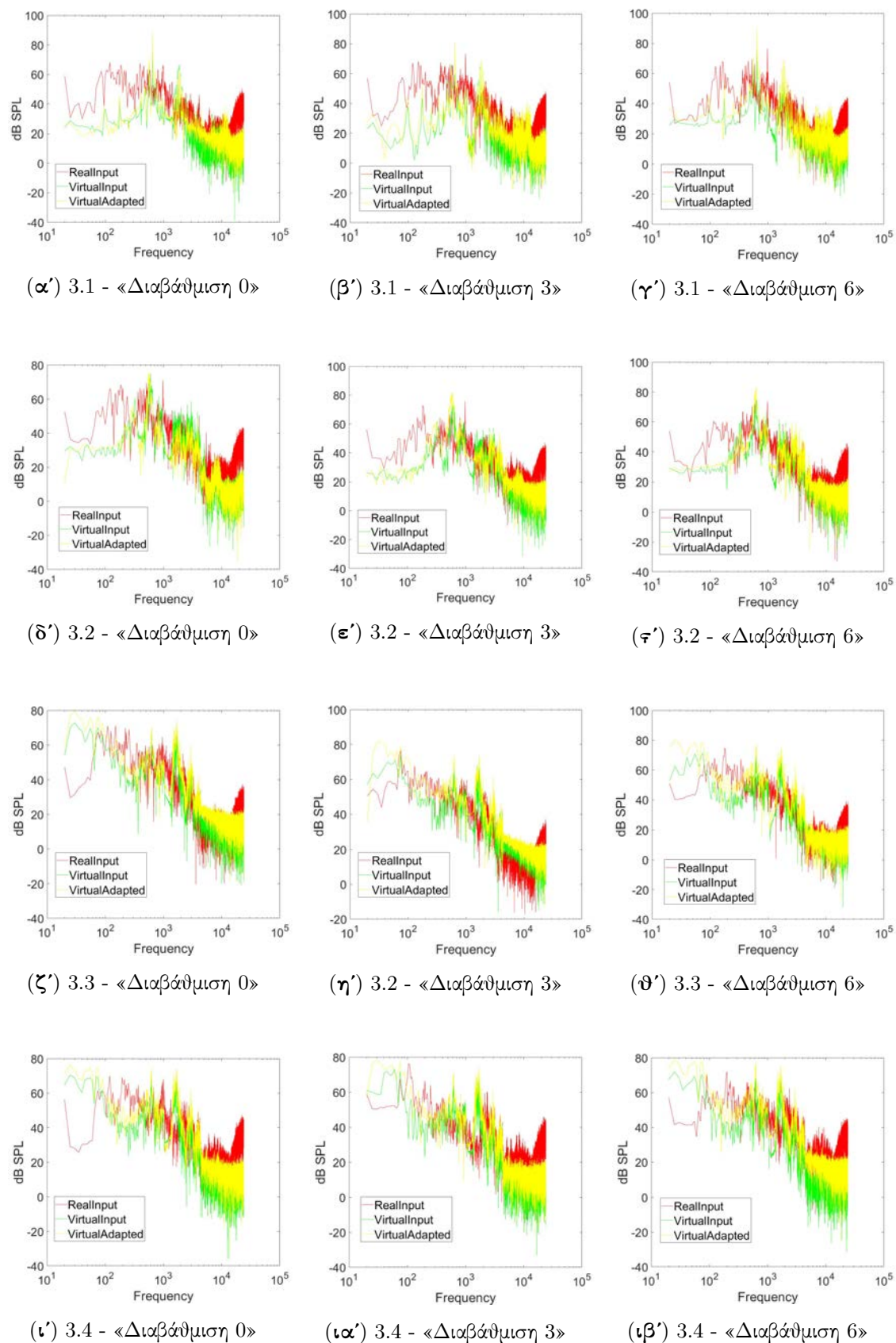
Σενάριο 3.3: Ως πραγματικό περιβάλλον παρουσιάζεται ο ήχος από έναν κεντρικό δρόμο πόλης σε ώρα αιχμής. Την εικονική πηγή ήχου αποτελεί ο τεχνητός ήχος με όνομα Score. Το συχνοτικό περιεχόμενο του ήχου πόλης σε ώρα αιχμής, παρουσιάζει έντονες διακυμάνσεις ηχητικής στάθμης δημιουργώντας ένα δύσκολο περιβάλλον για την επαυξημένη μείξη ήχου και κατά συνέπεια για την εφαρμογή στόχου όπως έχει οριστεί στην παρούσα διατριβή η κατηγορία των παιχνιδιών επαυξημένης πραγματικότητας ήχου. Η συχνοτική απόκριση του ηχητικού δείγματος της πόλης (βλπ Παράρτημα Γ', Σχ.Γ'.2) περιέχει μεγάλη συγκέντρωση ενέργειας για συχνότητες μικρότερες του  $1\text{ kHz}$ , ενώ για συχνότητες μεγαλύτερες, η ενέργεια, παρουσιάζει υποχώρηση. Ο ήχος Score, αποτελεί ένα συνθετικό ήχο κραυγής πόνου, ο οποίος αντικατοπτρίζει την επιτυχή βολή προς τον αντίπαλο από το ηχητικό παιχνίδι εξόντωσης της πανούκλας [111]. Η διάρκεια του ήχου Score, είναι 3 δευτερόλεπτα σύμφωνα με την κυματομορφή του (βλπ Παράρτημα Γ', Σχ.Γ'.11). Το συχνοτικό περιεχόμενο του ήχου σύμφωνα με τη φασματική ανάλυση (βλπ Παράρτημα Γ', Σχ.Γ'.12), περιέχει κυρίως το συχνοτικό φάσμα μέχρι τα  $2\text{ kHz}$ .

Σενάριο 3.4: Το πραγματικό περιβάλλον αποτελείται από τον ήχο ενός καφέ με κόσμο και αντίστοιχα η εικονική πηγή ήχου παρουσιάζεται ο τεχνητός ήχος με όνομα Score. Ο ήχος Score, είναι ο ίδιος που περιγράφηκε και στο προηγούμενο σενάριο, οπότε στη συνέχεια θα αναλυθεί ο ήχος του καφέ. Ο ήχος περιβάλλοντος του καφέ, αποτελεί έναν ήχο, αντίστοιχα με τον ήχο πόλης, έντονων διακυμάνσεων. Το συχνοτικό περιεχόμενό του παρουσιάζει μεγάλο εύρος. Στη φασματική ανάλυση (βλπ Παράρτημα Γ', Σχ.Γ'.4), φαίνεται να υπάρχει υψηλή συγκέντρωση ενέργειας μέχρι τα  $2\text{ kHz}$  με μια μικρή ελάττωση μέχρι τα  $12\text{ kHz}$ . Σε σύγκριση με τα υπόλοιπα ηχητικά περιβάλλοντα, ο ήχος του καφέ αποτελεί το πιο θορυβώδες περιβάλλον

μέσα στο οποίο η προσαρμοστική μείξη θα πρέπει να αναμείξει το εικονικό στοιχείο.

Αρχικά, έχει πραγματοποιηθεί βαθμονόμηση στο ακουστικό πεδίο για τα πραγματικά περιβάλλοντα με χρήση ηχόμετρου στα 82 dB SPL, καθώς και στο πρόγραμμα εξομίωσης για όλους τους ήχους συμπεριλαμβανομένου και των εικονικών, στη στάθμη των 82 dB SPL. Συνεπώς, η επιθυμητή ενέργεια στα σήματα εισόδου, έχει κοινή στάθμη κατά τη βαθμονόμηση, για να μην υπάρχει απόκλιση στην ηχητική στάθμη στις περιπτώσεις που δεν χρησιμοποιείται η προσαρμοστική μείξη, ώστε να παρουσιαστεί η διαφοροποίηση της προσαρμοστικής μείξης σε σχέση με τη περίπτωση βέλτιστης ρύθμισης της ηχητικής στάθμης μεταξύ πραγματικού και εικονικού περιβάλλοντος. Κατόπιν, όλα τα σήματα εισόδου και εξόδου ανά σενάριο αναλύθηκαν στο πεδίο της συχνότητας. Οι αναλύσεις περιλαμβάνουν τη φασματική ανάλυση του πραγματικού περιβάλλοντος, του εικονικού περιβάλλοντος πριν την προσαρμοστική μείξη, και του εικονικού περιβάλλοντος με τη χρήση προσαρμοστικής μείξης. Οι παραπάνω αναλύσεις, περιλαμβάνουν τις περιπτώσεις των τριών διαφορετικών διαβαθμίσεων μείξης, δηλαδή τη «Διαβάθμιση 0», τη «Διαβάθμιση 3», και τη «Διαβάθμιση 6». Το αποτέλεσμα των φασματικών αναλύσεων, παρουσιάζονται στα διαγράμματα του Σχήματος 7.1. Με μια πρώτη ματιά, διακρίνεται πως το προσαρμοσμένο εικονικό περιβάλλον έχει την τάση να ακολουθήσει το συχνοτικό φάσμα του πραγματικού περιβάλλοντος σε σύγκριση με το εικονικό περιβάλλον χωρίς την χρήση του προσαρμοστικού μοντέλου. Αυτό φαίνεται και στα τρία επίπεδα διαβάθμισης της προσαρμοστικής μείξης. Σχετικά με το συντελεστή διαβάθμισης, διακρίνεται μια αύξηση της ενέργειας ανάμεσα στα δείγματα επαυξημένης μείξης της κάθε περίπτωσης, ωστόσο χρειάζεται επιπλέον ανάλυση για να υπολογιστεί αντικειμενικά η επίδραση στην ενέργεια και στην ακουστότητα του προσαρμοσμένου εικονικού περιβάλλοντος.

Για την αντικειμενική αποτύπωση της φασματικής σχέσης στο πεδίο της συχνότητας μεταξύ των σημάτων εισόδου και εξόδου του προσαρμοστικού μοντέλου, χρησιμοποιείται η συνάρτηση συσχέτισης με την οποία υπολογίζεται το ποσοστό συσχέτισης του φάσματος μεταξύ των συγκρινόμενων περιβαλλόντων. Στον Πίνακα 7.1, παρουσιάζονται τα αποτελέσματα της συσχέτισης μεταξύ των ηχητικών συνιστωσών του επαυξημένου περιβάλλοντος. Τα ζευγάρια σημάτων που έχουν επιλεχθεί για να συγκριθούν τα φάσματά τους είναι τα ακόλουθα: Το πραγματικό περιβάλλον με το αρχικό σήμα της εικονικής πηγής ήχου, το δεύτερο ζευγάρι είναι το προσαρμοσμένο εικονικό σήμα με το πραγματικό ηχητικό περιβάλλον και το τελευταίο ζευγάρι αποτελεί το αρχικό εικονικό σήμα με το προσαρμοσμένο εικονικό σήμα. Στην πλειονότητα των μετρήσεων, η συσχέτιση του προσαρμοσμένου με το πραγματικό σε σύγκριση με το πραγματικό - εικονικό, δείχνει μια μείωση του ποσοστού συσχέτισης για το προσαρμοσμένο. Αυτό δεν παρατηρείται στην περίπτωση 3.3 για «Διαβάθμιση 0» και «Διαβάθμιση 6». Κατά συνέπεια, το προσαρμοσμένο μοντέλο μείξης, δημιουργεί μια τάση μείωσης της συσχέτισης σε ποσοστό από 1% – 10% στο 83% των περιπτώσεων. Ακολούθως για το ζευγάρι σύγκρισης αρχικού εικονικού σήματος και του αντίστοιχου προσαρμοσμένου, φαίνεται σε όλες τις περιπτώσεις να δημιουργείται μια μείωση της συσχέτισης που προκαλείται από την προσαρμογή και κυμαίνεται από 14%-32%. Η παραπάνω μείωση δεν φαίνεται να ακολουθεί κάποια συγκεκριμένη τάση σε σχέση και με την διαβάθμιση της προσαρμοστικής μείξης και επηρεάζεται από τα φασματικά χαρακτηριστικά του κάθε περιβάλλοντος.



Σχήμα 7.1: Διαγράμματα Φασματικής Ανάλυσης περιβαλλόντων παραδειγμάτων εφαρμογής αλγόριθμου προσαρμογής για τις περιπτώσεις 3.1, 3.2, 3.3, 3.4 με διαβάθμιση μείξης 0, 3, 6

Εκτός από τον υπολογισμό της συσχέτισης, για τις ίδιες περιπτώσεις σεναρίων, πραγματοποιείται μέτρηση της ενέργειας, καθώς και της ακουστότητας, με την χρήση του μοντέλου GM. Στον Πίνακα 7.2 παρουσιάζονται τα αποτελέσματα των μετρήσεων ενέργειας σε κλίμακα *dB SPL*, καθώς και η αμφιωτική ακουστότητα του πραγματικού περιβάλλοντος, του αρχικού εικονικού σήματος και του προσαρμοσμένου εικονικού σήματος.

Σε γενική εικόνα των μετρήσεων ακουστότητας, η προσαρμοστική μείξη, φαίνεται στις περισσότερες περιπτώσεις να επιτυγχάνει την ελάττωση της διαφοράς μεταξύ της ακουστότητας του πραγματικού με το προσαρμοσμένο εικονικό περιβάλλον σε σύγκριση με το αρχικό εικονικό σήμα. Στην περίπτωση 3.1 για «Διαβάθμιση 0», η ακουστότητα του πραγματικού περιβάλλοντος έχει μια διαφορά 13% με το εικονικό, την οποία η προσαρμοστική μείξη την ελαττώνει στο 5%. Αυτή η μείωση δεν σημαίνει βέβαια και αύξηση στην συσχέτιση του φάσματος προσαρμοσμένου με πραγματικό σύμφωνα και με τις προηγούμενες μετρήσεις. Κατά συνέπεια, μπορεί η ακουστότητα να τείνει να ισοσκελιστεί μεταξύ του εικονικού και του πραγματικού περιβάλλοντος μέσα από μια προσαρμοστική μείξη, ωστόσο δεν σημαίνει ότι φασματικά τα δύο περιβάλλοντα θα γίνουν όμοια.

Σύμφωνα με τις μετρήσεις συσχέτισης προσαρμοσμένου περιβάλλοντος με πραγματικό, στην περίπτωση 3.1 και στην περίπτωση 3.3 η αύξηση του συντελεστή διαβάθμισης, είναι ανάλογη με την συσχέτιση, καθώς παρατηρείται ομοίως βελτίωση της συσχέτισης. Στα σενάρια 3.2 και 3.4 παρατηρείται πως την μεγαλύτερη συσχέτιση εμφανίζει η διαβάθμιση 6 και υπάρχει μια σύγκριση μεταξύ της διαβάθμισης 0 και 3.

| Σενάριο | Συντελεστής Διαβάθμισης Μείξης | Πραγματικό         | Εικονικό | Προσαρμοσμένο      | Πραγματικό | Εικονικό           | Προσαρμοσμένο |
|---------|--------------------------------|--------------------|----------|--------------------|------------|--------------------|---------------|
|         |                                | Συσχέτιση Φάσματος |          | Συσχέτιση Φάσματος |            | Συσχέτιση Φάσματος |               |
| 3.1     | 0                              | 38%                |          | 29%                |            | 68%                |               |
|         | 3                              | 33%                |          | 32%                |            | 85%                |               |
|         | 6                              | 40%                |          | 38%                |            | 72%                |               |
| 3.2     | 0                              | 47%                |          | 43%                |            | 83%                |               |
|         | 3                              | 43%                |          | 37%                |            | 80%                |               |
|         | 6                              | 46%                |          | 45%                |            | 84%                |               |
| 3.3     | 0                              | 56%                |          | 54%                |            | 86%                |               |
|         | 3                              | 57%                |          | 58%                |            | 88%                |               |
|         | 6                              | 56%                |          | 61%                |            | 84%                |               |
| 3.4     | 0                              | 51%                |          | 45%                |            | 85%                |               |
|         | 3                              | 46%                |          | 42%                |            | 83%                |               |
|         | 6                              | 55%                |          | 51%                |            | 85%                |               |

Πίνακας 7.1: Πίνακας συσχέτισης φάσματος επαυξημένων περιβαλλόντων για παραδείγματα 3.1, 3.2, 3.3, 3.4 με διαβάθμιση μείξης 0, 3, 6

Η πρόβλεψη της ακουστότητας από το μοντέλο, αποτελεί μια απαιτητική διεργασία για το μοντέλο προσαρμοστικής μείξης. Παρόλα αυτά, σύμφωνα με τις μετρήσεις, φαίνεται να υπάρχει μια ομοιόμορφη λειτουργία του και σε συνδυασμό με τον συντελεστή διαβάθμισης, δίνεται η ευκαιρία της υποκειμενικής αξιολόγησης της επιθυμητής διαφοροποίησης μεταξύ των περιβαλλόντων που αναμειγνύονται. Στην επόμενη ενότητα παρουσιάζονται τα αποτελέσματα της υποκειμενικής αξιολόγησης που διενεργήθηκε για το μοντέλο προσαρμοστικής μείξης

επαυξημένου ήχου.

| Σενάριο | Συντελεστής<br>Διαβάθμισης Μείξης | Πραγματικό<br>dB SPL | Εικονικό<br>dB SPL | Προσαρμοσμένο<br>dB SPL | Πραγματικό<br>GM (Phons) | Εικονικό<br>GM (Phons) | Προσαρμοσμένο<br>GM (Phons) |
|---------|-----------------------------------|----------------------|--------------------|-------------------------|--------------------------|------------------------|-----------------------------|
| 3.1     | 0                                 | 82                   | 82                 | 89                      | 93                       | 80                     | 88                          |
|         | 3                                 | 82                   | 82                 | 82                      | 95                       | 91                     | 92                          |
|         | 6                                 | 82                   | 82                 | 90                      | 94                       | 90                     | 95                          |
| 3.2     | 0                                 | 82                   | 82                 | 83                      | 94                       | 94                     | 92                          |
|         | 3                                 | 82                   | 82                 | 90                      | 94                       | 94                     | 96                          |
|         | 6                                 | 82                   | 82                 | 90                      | 95                       | 95                     | 97                          |
| 3.3     | 0                                 | 82                   | 82                 | 88                      | 93                       | 91                     | 95                          |
|         | 3                                 | 82                   | 82                 | 90                      | 92                       | 92                     | 97                          |
|         | 6                                 | 82                   | 82                 | 90                      | 93                       | 93                     | 98                          |
| 3.4     | 0                                 | 82                   | 82                 | 88                      | 96                       | 92                     | 97                          |
|         | 3                                 | 82                   | 82                 | 88                      | 92                       | 92                     | 96                          |
|         | 6                                 | 82                   | 82                 | 90                      | 96                       | 92                     | 98                          |

Πίνακας 7.2: Πίνακας υπολογισμών ενέργειας και Ακουστότητας (Μοντέλο GM) για περιπτώσεις παραδειγμάτων 3.1, 3.2, 3.3, 3.4 με διαβάθμιση μείξης 0, 3, 6

### 7.3 Αποτελέσματα υποκειμενικής αξιολόγησης ARA μείξης

Η διαδικασία της αξιολόγησης, εγκρίθηκε από την επιτροπή δεοντολογίας του Ιονίου Πανεπιστημίου στις 13/12/2019 με την απόφαση της με αριθμό πρωτοκόλλου 4542 (βλπ Παράρτημα Δ'). Η διενέργειά της τελέστηκε στο «Εργαστήριο Επεξεργασίας Οπτικοακουστικών Σημάτων - ΕΠ.ΟΑ.ΣΗ», του τμήματος Τεχνών Ήχου και Εικόνας του Ιονίου Πανεπιστημίου. Στη διαδικασία πήραν μέρος συνολικά 78 υποκείμενα τα οποία συμπλήρωσαν την αίτηση συγκατάθεσης συμμετοχής στην πειραματική διαδικασία (βλπ Παράρτημα Δ'). Στην διαδικασία τηρήθηκαν όλα τα υγειονομικά μέτρα προστασίας για τους συμμετέχοντες, με τον καθαρισμό του ακουστικού εξοπλισμού πριν την ατομική χρήση του.

Για την ανάλυση των αποτελεσμάτων, δημιουργήθηκαν διακριτές κατηγορίες υποκειμένων με βάση συγκεκριμένα κριτήρια. Η πρώτη κατηγορία, αποτελεί το σύνολο των υποκειμένων όπου συμμετείχαν στην πειραματική διαδικασία χωρίς κάποια εξαίρεση. Ο συνολικός αριθμός σε αυτή την κατηγορία είναι τα 78 υποκείμενα. Η δεύτερη κατηγορία, δημιουργείται έχοντας κριτήριο το τεστ ακουστικής διάκρισης του υποκειμένου και τις σωστές απαντήσεις που έδωσε (βλπ Σχ.6.7). Σε αυτή την κατηγορία, σύμφωνα με τα αποτελέσματα, 3 υποκείμενα δεν απάντησαν σωστά σε καμία ερώτηση και κατατάσσονται στην κατηγορία των «Αφελών» (Naïve). Από το συνολικό αριθμό των υποκειμένων, 27 άτομα απάντησαν σωστά σε ένα τεστ διάκρισης και κατηγοριοποιούνται ως «Αρχάριοι». Στην κατηγορία «Έμπειροι», εξακριβώθηκαν συνολικά 26 υποκείμενα τα οποία είχαν δύο σωστές απαντήσεις στο τεστ ακουστικής διάκρισης. Τέλος στην κατηγορία «Ειδικοί», κατηγοριοποιήθηκαν 22 άτομα με σωστές το σύνολο των απαντήσεων του τεστ διάκρισης.

Η τρίτη κατηγορία του συνόλου των υποκειμένων, δημιουργήθηκε με βάση τον συντελεστή διαβάθμισης της μείξης που ρυθμίστηκε πριν την διεξαγωγή της κάθε πειραματικής διαδικα-

σίας. Η ύπαρξή του δεν γίνεται γνωστή στο υποκείμενο. Σε αυτή την κατηγορία δημιουργήθηκαν τρία σύνολα αντίστοιχα με τους τρεις συντελεστές διαβάθμισης που επιλέχθηκαν. Η κατηγορία με συντελεστή διαβάθμισης 0 περιέχει 26 υποκείμενα όπως και η κατηγορία με συντελεστή 3 και 6 αντίστοιχα. Παρακάτω, θα παρουσιαστούν τα αποτελέσματα των τεσσάρων τεστ τύπου MUSHRA. Τα ηχητικά σενάρια που ακολουθήθηκαν για την επιλογή του πραγματικού ηχητικού περιβάλλοντος και του εικονικού είναι τα ίδια με τα τέσσερα σενάρια που χρησιμοποιήθηκαν κατά την αντικειμενική αξιολόγηση του προσαρμοστικού μοντέλου μείζης. Τα αποτελέσματα αναλύονται σύμφωνα με την μέση τιμή της αξιολόγησης που δόθηκε από όλα τα υποκείμενα καθώς και από τη διενέργεια t-Test για τον έλεγχο της στατιστικής σημαντικότητας μεταξύ της προσαρμοστικής εικονικής πηγής ήχου ως ζευγάρι με την κάθε μία από τις υπόλοιπες εκδοχές μείζης. Για  $p - Value < 0.05$  θεωρείται ως στατιστικά σημαντικό το αποτέλεσμα και καταρρίπτεται η αρχική υπόθεση ότι η προσαρμοστική μείξη δεν είναι αρεστή στον ακροατή (στους πίνακες των αποτελεσμάτων συμβολίζεται με το λεκτικό Signif.). Σε αντίθετη περίπτωση το αποτέλεσμα δεν είναι σημαντικό και ισχύει η αρχική υπόθεση (στους πίνακες των αποτελεσμάτων συμβολίζεται με το λεκτικό Null Acc.) Η ερώτηση πάνω στην οποία καλείται το υποκείμενο να βαθμολογήσει τις εκδοχές μείζης είναι το πόσο του αρέσει το ηχητικό αποτέλεσμα της μείξης μεταξύ του πραγματικού κόσμου και της εικονικής πηγής ήχου.

| MUSHRA Τέστ 3.1                                                      |                                |                 |          |            |           |            |           |            |          |              |          |              |          |              |           |
|----------------------------------------------------------------------|--------------------------------|-----------------|----------|------------|-----------|------------|-----------|------------|----------|--------------|----------|--------------|----------|--------------|-----------|
| Κλίμακα: Κακό = 1 / Αδύναμο = 2 / Μέτριο = 3 / Καλό = 4 / Τέλειο = 5 |                                |                 |          |            |           |            |           |            |          |              |          |              |          |              |           |
| Πραγματικό Περιβάλλον                                                | Εικονικό Περιβάλλον            | Κατηγοριοποίηση |          |            |           |            |           |            |          |              |          |              |          |              |           |
|                                                                      |                                | Συνολικό        |          | Αρχάριοι   |           | Έμπειροι   |           | Ειδικοί    |          | Διαβάθμιση 0 |          | Διαβάθμιση 3 |          | Διαβάθμιση 6 |           |
|                                                                      |                                | Μέσος Όρος      | p < 0.05 | Μέσος Όρος | p < 0.05  | Μέσος Όρος | p < 0.05  | Μέσος Όρος | p < 0.05 | Μέσος Όρος   | p < 0.05 | Μέσος Όρος   | p < 0.05 | Μέσος Όρος   | p < 0.05  |
| Βιομηχανία                                                           | Roboscream                     | 3.03            | Signif.  | 2.89       | Signif.   | 3.20       | Signif.   | 3.05       | Signif.  | 2.92         | Signif.  | 3.12         | Signif.  | 3.04         | Signif.   |
| Βιομηχανία                                                           | Roboscream Προσαρμοσμένο       | 3.90            | -        | 3.92       | -         | 3.96       | -         | 4.15       | -        | 3.96         | -        | 4.13         | -        | 3.75         | -         |
| Βιομηχανία                                                           | Roboscream 3.5 kHz Εύρος Ζώνης | 3.00            | Signif.  | 3.16       | Signif.   | 3.17       | Signif.   | 2.95       | Signif.  | 2.83         | Signif.  | 3.33         | Signif.  | 2.88         | Signif.   |
| Βιομηχανία                                                           | Roboscream 7 kHz Εύρος Ζώνης   | 3.10            | Signif.  | 3.32       | Null Acc. | 3.00       | Signif.   | 3.15       | Signif.  | 2.96         | Signif.  | 3.08         | Signif.  | 3.17         | Null Acc. |
| Βιομηχανία                                                           | Roboscream 10 kHz Εύρος Ζώνης  | 3.06            | Signif.  | 2.84       | Signif.   | 3.30       | Null Acc. | 3.20       | Signif.  | 3.25         | Signif.  | 2.79         | Signif.  | 3.21         | Null Acc. |

Πίνακας 7.3: MUSHRA Τέστ 3.1

Στο πρώτο MUSHRA τεστ του Πίνακα 7.3, το υποκείμενο ακούει ως πραγματικό περιβάλλον τον ήχο από μια «Βιομηχανία» και ως εικονικό περιβάλλον τον τεχνητό ήχο τύπου «Roboscream». Η μέση τιμή αξιολόγησης της προσαρμοστικής μείξης σε όλες τις κατηγορίες φαίνεται να είναι σε υψηλότερα επίπεδα σε σύγκριση με τις υπόλοιπες εκδοχές. Η μέση τιμή των κατηγοριών για την προσαρμοστική μείξη κυμαίνεται μεταξύ του 3,75 – 4,14, η οποία, σε μια αντιστοίχιση ως προς την κλίμακα μείζης, είναι στα επίπεδα του καλού. Ως προς την στατιστική σημαντικότητα των αποτελεσμάτων, στα t-Test που διενεργήθηκαν για την προσαρμοστική μείξη σε αντιπαράθεση με τις υπόλοιπες εκδοχές, παρουσιάζεται στην πλειονότητα των περιπτώσεων το αποτέλεσμα να είναι στατιστικά σημαντικό, εκτός από την κάποιες μεμονωμένες περιπτώσεις σε αρχάριους και έμπειρους. Στην περίπτωση της «Διαβάθμισης 6», φαίνεται ότι στις περιπτώσεις της μείξης με εύρος ζώνης εικονικού περιβάλλοντος 7 kHz και 10 kHz, να μην αποδεικνύεται σημαντική η διαφορά της μέσης τιμής.

Στο δεύτερο MUSHRA τεστ του οποίου τα αποτελέσματα παρουσιάζονται στον Πίνακα

7.4, το πραγματικό περιβάλλον αναπαράγεται ως ήχος «Βιομηχανίας», ενώ το εικονικό αντιπροσωπεύει ο τεχνητός ήχος με το όνομα «Miss». Ο υπολογισμός της μέσης τιμής των αποτελεσμάτων για την μείξη, φαίνεται πως και σε αυτή την περίπτωση είναι θετικός προς την προσαρμοστική μείξη. Ωστόσο στην ανάλυση της σχέσης των τιμών της προσαρμοστικής μείξης με τις υπόλοιπες μέσω της χρήσης t-Test, φαίνεται πως ενώ στο γενικό σύνολο η διαφορά είναι στατιστικά σημαντική εκτός από την περίπτωση της εικονικής πηγής εύρους 10 kHz, στις υπόλοιπες περιπτώσεις των διάφορων κατηγοριών παρουσιάζονται διαφοροποιήσεις. Πιο αναλυτικά, στην πλειονότητα των t-Test, για τις υπόλοιπες κατηγορίες, φαίνεται πως δεν είναι στατιστικά σημαντική αυτή η διαφορά. Συνοψίζοντας τα στατιστικά στοιχεία για το δεύτερο MUSHRA τεστ, φαίνεται πως η προσαρμοστική μείξη στο γενικό σύνολο, φαίνεται να έχει σημαντική διαφορά ωστόσο δεν μπορεί να βγει κάποιο ασφαλές συμπέρασμα στις υπόλοιπες κατηγορίες.

| MUSHRA Τέστ 3.2                                                      |                          |                 |                 |            |                 |            |                 |            |                 |              |                 |              |                 |              |                 |
|----------------------------------------------------------------------|--------------------------|-----------------|-----------------|------------|-----------------|------------|-----------------|------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|
| Κλίμακα: Κακό = 1 / Αδύναμο = 2 / Μέτριο = 3 / Καλό = 4 / Τέλειο = 5 |                          |                 |                 |            |                 |            |                 |            |                 |              |                 |              |                 |              |                 |
| Πραγματικό Περιβάλλον                                                | Εικονικό Περιβάλλον      | Κατηγοριοποίηση |                 |            |                 |            |                 |            |                 |              |                 |              |                 |              |                 |
|                                                                      |                          | Συνολικό        |                 | Αρχάριοι   |                 | Έμπειροι   |                 | Είδοκοι    |                 | Διαβάθμιση 0 |                 | Διαβάθμιση 3 |                 | Διαβάθμιση 6 |                 |
|                                                                      |                          | Μέσος Όρος      | <i>p</i> < 0.05 | Μέσος Όρος | <i>p</i> < 0.05 | Μέσος Όρος | <i>p</i> < 0.05 | Μέσος Όρος | <i>p</i> < 0.05 | Μέσος Όρος   | <i>p</i> < 0.05 | Μέσος Όρος   | <i>p</i> < 0.05 | Μέσος Όρος   | <i>p</i> < 0.05 |
| Βιομηχανία                                                           | Miss                     | 3.40            | Signif.         | 3.52       | Null Acc.       | 3.40       | Null Acc.       | 3.45       | Null Acc.       | 3.58         | Null Acc.       | 3.50         | Null Acc.       | 3.12         | Null Acc.       |
| Βιομηχανία                                                           | Miss Προσαρμοσμένο       | 3.76            | -               | 3.76       | -               | 3.96       | -               | 3.65       | -               | 3.88         | -               | 3.71         | -               | 3.79         | -               |
| Βιομηχανία                                                           | Miss 3.5 kHz Εύρος Ζώνης | 3.10            | Signif.         | 3.04       | Signif.         | 3.30       | Signif.         | 3.20       | Null Acc.       | 2.96         | Signif.         | 3.17         | Null Acc.       | 3.33         | Null Acc.       |
| Βιομηχανία                                                           | Miss 7 kHz Εύρος Ζώνης   | 3.24            | Signif.         | 3.36       | Null Acc.       | 3.39       | Null Acc.       | 3.10       | Null Acc.       | 3.38         | Null Acc.       | 3.42         | Null Acc.       | 3.13         | Signif.         |
| Βιομηχανία                                                           | Miss 10 kHz Εύρος Ζώνης  | 3.51            | Null Acc.       | 3.44       | Null Acc.       | 3.48       | Null Acc.       | 3.80       | Null Acc.       | 3.71         | Null Acc.       | 3.54         | Null Acc.       | 3.33         | Null Acc.       |

Πίνακας 7.4: MUSHRA Τέστ 3.2

| MUSHRA Τέστ 3.3                                                      |                           |                 |           |            |           |            |           |            |          |              |           |              |           |              |           |
|----------------------------------------------------------------------|---------------------------|-----------------|-----------|------------|-----------|------------|-----------|------------|----------|--------------|-----------|--------------|-----------|--------------|-----------|
| Κλίμακα: Κακό = 1 / Αδύναμο = 2 / Μέτριο = 3 / Καλό = 4 / Τέλειο = 5 |                           |                 |           |            |           |            |           |            |          |              |           |              |           |              |           |
| Πραγματικό Περιβάλλον                                                | Εικονικό Περιβάλλον       | Κατηγοριοποίηση |           |            |           |            |           |            |          |              |           |              |           |              |           |
|                                                                      |                           | Συνολικό        |           | Αρχάριοι   |           | Έμπειροι   |           | Είδοι      |          | Διαβάθμιση 0 |           | Διαβάθμιση 3 |           | Διαβάθμιση 6 |           |
|                                                                      |                           | Μέσος Όρος      | p < 0.05  | Μέσος Όρος | p < 0.05  | Μέσος Όρος | p < 0.05  | Μέσος Όρος | p < 0.05 | Μέσος Όρος   | p < 0.05  | Μέσος Όρος   | p < 0.05  | Μέσος Όρος   | p < 0.05  |
| Πόλη                                                                 | Score                     | 3.03            | Null Acc. | 3.04       | Null Acc. | 3.08       | Null Acc. | 2.91       | Signif.  | 2.88         | Null Acc. | 3.42         | Null Acc. | 2.77         | Null Acc. |
| Πόλη                                                                 | Score Προσαρμοσμένο       | 3.59            | -         | 3.48       | -         | 3.35       | -         | 3.85       | -        | 3.25         | -         | 3.79         | -         | 3.58         | -         |
| Πόλη                                                                 | Score 3.5 kHz Εύρος Ζώνης | 2.99            | Signif.   | 2.80       | Signif.   | 3.00       | Null Acc. | 3.15       | Signif.  | 2.71         | Signif.   | 3.42         | Null Acc. | 3.00         | Signif.   |
| Πόλη                                                                 | Score 7 kHz Εύρος Ζώνης   | 3.12            | Signif.   | 3.24       | Null Acc. | 2.96       | Null Acc. | 3.20       | Signif.  | 2.88         | Null Acc. | 3.13         | Null Acc. | 3.46         | Null Acc. |
| Πόλη                                                                 | Score 10 kHz Εύρος Ζώνης  | 3.38            | Null Acc. | 3.40       | Null Acc. | 3.39       | Null Acc. | 3.40       | Signif.  | 3.29         | Null Acc. | 3.67         | Null Acc. | 3.29         | Null Acc. |

Πίνακας 7.5: MUSHRA Τέστ 3.3

Στο MUSHRA τεστ 3.3 του οποίου τα αποτελέσματα παρουσιάζονται στον Πίνακα 7.5, το πραγματικό περιβάλλον μεταφέρεται στο υποκείμενο ως «Ήχος πόλης σε ώρα αιχμής», ενώ το εικονικό περιβάλλον ως τεχνητός ήχος τύπου «Score». Στην περίπτωση υπολογισμού της μέσης τιμής, φαίνεται το προσαρμοστικό μοντέλο καταφέρνει και έχει την υψηλότερη σε σύγκριση με τις υπόλοιπες περιπτώσεις και κυμαίνεται το αποτέλεσμα στην κλίμακα από μέτριο μέχρι καλό. Ως προς την στατιστική ανάλυση του αποτελέσματος, σύμφωνα με τις αναλύσεις τύπου T-test, στο γενικό σύνολο, το αποτέλεσμα της προσαρμοστικής μείξης έναντι του εικονικού ήχου εύρους ζώνης 3.5kHz και 7kHz, παρουσιάζει στατιστική σημαντικότητα σε αντίθεση με το εικονικό περιβάλλον χωρίς προσαρμογή και εκείνο με εύρος 10kHz. Στην

κατηγορία των ειδικών παρόλα αυτά, παρουσιάζεται πως η προσαρμοστική μείζη, εμφανίζει σημαντική διαφοροποίηση σε σύγκριση με τις υπόλοιπες εκδοχές. Σε αυτό το παράδειγμα φαίνεται πως η εμπειρία και η ακουστική διαφοροποίηση των υποσυνόλων, επιδρά στην αντίληψη της προσαρμοστικής μείζης.

Στο τελευταίο MUSHRA τεστ (βλπ Πίν.7.6), ως πραγματικό περιβάλλον αναπαράγεται ο ήχος από ένα «Καφέ» και ως εικονικό περιβάλλον, ο τεχνητός ήχος τύπου «Score». Στα αποτελέσματα που εξήχθησαν, η μέση τιμή του προσαρμοσμένου εικονικού περιβάλλοντος, φαίνεται να είναι σε πιο υψηλά επίπεδα σε σύγκριση με τις υπόλοιπες εκδοχές μείζης. Η μέση τιμή για το περιβάλλον προσαρμοστικής μείζης κυμαίνεται από τιμές 3,54 – 3,79 οι οποίες στην κλίμακα κατάταξης της μείζης τείνουν προς το επίπεδο του καλού αποτελέσματος. Πιο αναλυτικά, στο γενικό σύνολο, το αποτέλεσμα παρουσιάζει σημαντική διαφορά στις τρεις από τις τέσσερις εκδοχές μείζης. Ειδικότερα, πρέπει να σημειωθεί πως, σε όλες τις κατηγορίες, η προσαρμοστική μείζης σε σύγκριση με το αρχικό εικονικό περιβάλλον, παρουσιάζει σημαντική διαφοροποίηση. Επίσης, στην κατηγορία των ειδικών, φαίνεται πως υπάρχει σημαντική διαφοροποίηση για την προσαρμοστική μείζη. Επιπροσθέτως όμως, παρουσιάζεται σημαντική διαφορά στην «Διαβάθμιση 0» και στην «Διαβάθμιση 3» σε τρεις από τις τέσσερις περιπτώσεις σε αντίθεση με τη «Διαβάθμιση 6» όπου παρουσιάζεται σημαντική διαφοροποίηση μόνο σε σχέση με τη μη επεξεργασμένη εικονική πηγή ήχου.

| MUSHRA Τέστ 3.4                                                      |                          |                 |           |            |           |            |           |            |           |              |           |              |           |              |           |
|----------------------------------------------------------------------|--------------------------|-----------------|-----------|------------|-----------|------------|-----------|------------|-----------|--------------|-----------|--------------|-----------|--------------|-----------|
| Κλίμακα: Κακό = 1 / Αδύναμο = 2 / Μέτριο = 3 / Καλό = 4 / Τέλειο = 5 |                          |                 |           |            |           |            |           |            |           |              |           |              |           |              |           |
| Πραγματικό Περιβάλλον                                                | Εικονικό Περιβάλλον      | Κατηγοριοποίηση |           |            |           |            |           |            |           |              |           |              |           |              |           |
|                                                                      |                          | Συνολικό        |           | Αρχάριοι   |           | Έμπειροι   |           | Ειδικοί    |           | Διαβάθμιση 0 |           | Διαβάθμιση 3 |           | Διαβάθμιση 6 |           |
|                                                                      |                          | Μέσος Όρος      | p < 0.05  | Μέσος Όρος | p < 0.05  | Μέσος Όρος | p < 0.05  | Μέσος Όρος | p < 0.05  | Μέσος Όρος   | p < 0.05  | Μέσος Όρος   | p < 0.05  | Μέσος Όρος   | p < 0.05  |
| Καφέ                                                                 | Score                    | 2,76            | Signif.   | 2,67       | Signif.   | 3,04       | Signif.   | 2,64       | Signif.   | 2,58         | Signif.   | 2,88         | Signif.   | 2,81         | Signif.   |
| Καφέ                                                                 | Score Προσαρμοσμένο      | 3,69            | -         | 3,56       | -         | 3,78       | -         | 3,70       | -         | 3,75         | -         | 3,79         | -         | 3,54         | -         |
| Καφέ                                                                 | Score 3.5kHz Εύρος Ζώνης | 3,13            | Signif.   | 2,96       | Null Acc. | 3,22       | Null Acc. | 3,15       | Signif.   | 3,00         | Signif.   | 3,13         | Signif.   | 3,25         | Null Acc. |
| Καφέ                                                                 | Score 7kHz Εύρος Ζώνης   | 3,17            | Signif.   | 3,16       | Null Acc. | 3,22       | Null Acc. | 3,10       | Signif.   | 3,08         | Signif.   | 3,17         | Signif.   | 3,38         | Null Acc. |
| Καφέ                                                                 | Score 10kHz Εύρος Ζώνης  | 3,37            | Null Acc. | 3,60       | Null Acc. | 3,22       | Null Acc. | 3,35       | Null Acc. | 3,67         | Null Acc. | 3,17         | Null Acc. | 3,21         | Null Acc. |

Πίνακας 7.6: MUSHRA Τέστ 3.4

Σύμφωνα με τα παραπάνω, στο γενικό σύνολο όλων των τεστ τύπου MUSHRA, εμφανίζεται στατιστική σημαντικότητα το αποτέλεσμα της μέσης τιμής της προσαρμοστικής μείζης. Εκτός από το γενικό σύνολο, στις κατηγορίες σύμφωνα με την ακουστική εμπειρία, στα δύο από τα τέσσερα τεστ (MUSHRA 3.1 και MUSHRA 3.4, εμφανίζεται στατιστική σημαντικότητα για τους ειδικούς οι οποίοι εκλαμβάνουν ως καλύτερη την προσαρμοστική μείζη σε σύγκριση με τις υπόλοιπες εκδοχές. Επίσης, στα τεστ MUSHRA 3.1 και MUSHRA 3.4, οι κατηγορίες διαβάθμισης 0 και 3, παρουσιάζουν σημαντική στατιστική διαφορά για την προσαρμοστική μείζη. Σίγουρα πρόκειται για μια υποκειμενική αντιμετώπιση της επαυξημένης μείζης ήχου έχοντας το κάθε υποκείμενο δική του ακουστική αντίληψη. Στο σύνολο όμως της αξιολόγησης σύμφωνα με τα αποτελέσματα της μέσης τιμής, φαίνεται πως η προσαρμοστική μείζη μπορεί να αναβαθμίσει την αισθητική ποιότητα της επαυξημένης μείζης, κάνοντας αυτήν πιο ελκυστική. Ωστόσο, λόγω του ότι συχνά στο πεδίο της επαυξημένης πραγματικότητας ήχου υπάρχει συνδυασμός διάδρασης για την επίτευξη κάποιου στόχου, όπως έχει ήδη αναφερθεί,

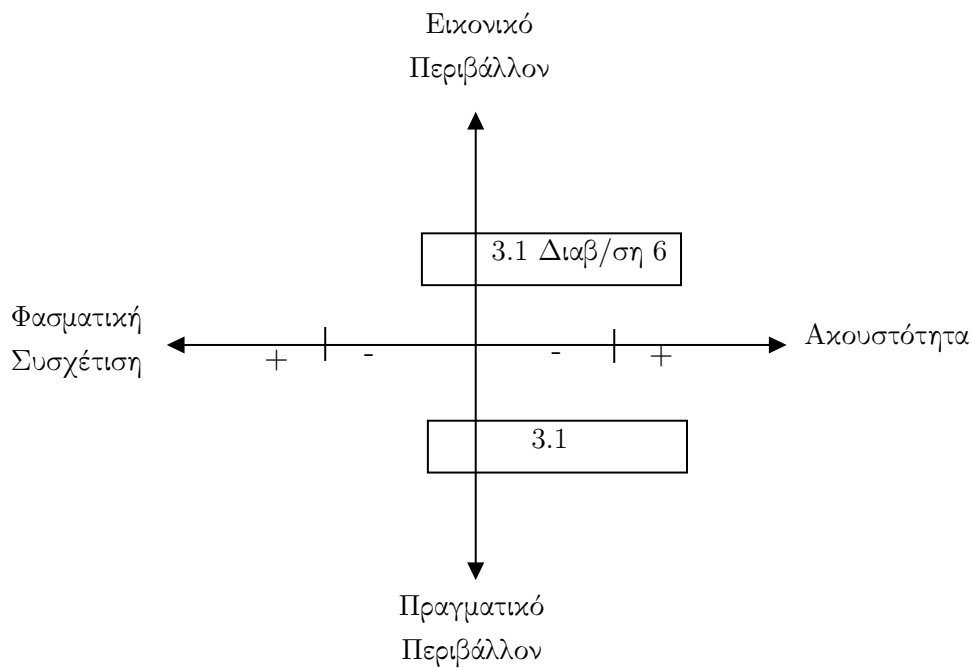
υλοποιήθηκε η υποκειμενική αξιολόγηση ενός ηχητικού παιχνιδιού τύπου ARA η οποία αποτελεί την εφαρμογή στόχο για την πειραματική αξιολόγηση του προσαρμοστικού μοντέλου μείξης, της οποίας τα αποτελέσματα παρουσιάζονται στην επόμενη ενότητα.

Για την συσχέτιση των αποτελεσμάτων των αντικειμενικών μετρήσεων της Ενότητας 7.2 και της στατιστικής σημαντικότητας των αποτελεσμάτων της υποκειμενικής αξιολόγησης, δημιουργήθηκε ένα πεδίο απεικόνισης της σύγκρισης της προσαρμοστικής μείξης με την απλή μείξη ανάμεσα σε δύο άξονες. Ο πρώτος άξονας έχει στα άκρα του το εικονικό περιβάλλον και το πραγματικό περιβάλλον. Με αυτό τον άξονα απεικονίζεται η ισορροπία του πραγματικού περιβάλλοντος ως προς το εικονικό για την προσαρμοστική μείξη με βάση την ακουστότητα. Στον οριζόντιο άξονα απεικονίζεται στην μια πλευρά του η φασματική συσχέτιση μεταξύ των συνισταμένων περιβαλλόντων με διάκριση στην τάση συσχέτισης αν είναι αυξητική σε σύγκριση με την μείξη χωρίς προσαρμογή και, αντίστοιχα, στην δεξιά πλευρά, έχει τοποθετηθεί η ακουστότητα με διάκριση της τάσης (αύξηση, μείωση), σε σχέση με την αρχική μείξη χωρίς προσαρμογή.

Στο Σχήμα 7.2 απεικονίζεται το πεδίο στο οποίο βρίσκεται η προσαρμοστική μείξη σε σχέση με την αρχική μείξη για το σενάριο 3.1. Καταγράφοντας την ισορροπία της μείξης, φαίνεται πως στο σύνολο των εκδοχών εκτός από τη «Διαβάθμιση 6», η ισορροπία της μείξης τείνει προς το πραγματικό περιβάλλον. Σχετικά με την ακουστότητα, παρουσιάζει αυξητική τάση σε αντίθεση με την φασματική συσχέτιση η οποία μειώνεται. Σε αυτή την περίπτωση τα αποτελέσματα του MUSHRA τέστ, εμφανίζουν στατιστική σημαντικότητα για το γενικό σύνολο καθώς και για όλες τις διαβαθμίσεις εκτός από την «Διαβάθμιση 6». Σύμφωνα με αυτό φαίνεται πως για τα υποκείμενα στην περίπτωση της ενσωμάτωσης του εικονικού περιβάλλοντος στο πραγματικό με αύξηση της ακουστότητας σε σχέση με τη μείξη χωρίς προσαρμογή, διατηρώντας την ισορροπία της μείξης προς το πραγματικό περιβάλλον, εκλαμβάνεται θετικά από τα υποκείμενα.

Για το σενάριο 3.2, έγινε η ίδια ανάλυση με παραπάνω για την τοποθέτηση της προσαρμοστικής μείξης στο πεδίο σύγκρισης (βλπ Σχ.7.3). Η προσαρμοστική μείξη του σεναρίου, παρουσιάζει για το γενικό σύνολο και τις διαβαθμίσεις με εξαίρεση την «Διαβάθμιση 0» ισορροπία μείξης προς το εικονικό περιβάλλον. Η σχέση φασματικής συσχέτισης και ακουστότητας για την προσαρμοστική μείξη τείνει για όλες τις περιπτώσεις προς την αύξηση της ακουστότητας. Σε αυτή την περίπτωση με τη μετατόπιση ισορροπίας προς το εικονικό περιβάλλον παρατηρείται στα υποκειμενικά αποτελέσματα για το γενικό σύνολο μη σημαντική διαφορά για την ποιότητα της μείξης σε σύγκριση με τις υπόλοιπες εκδοχές, δημιουργώντας μια σύγχυση στα υποκείμενα.

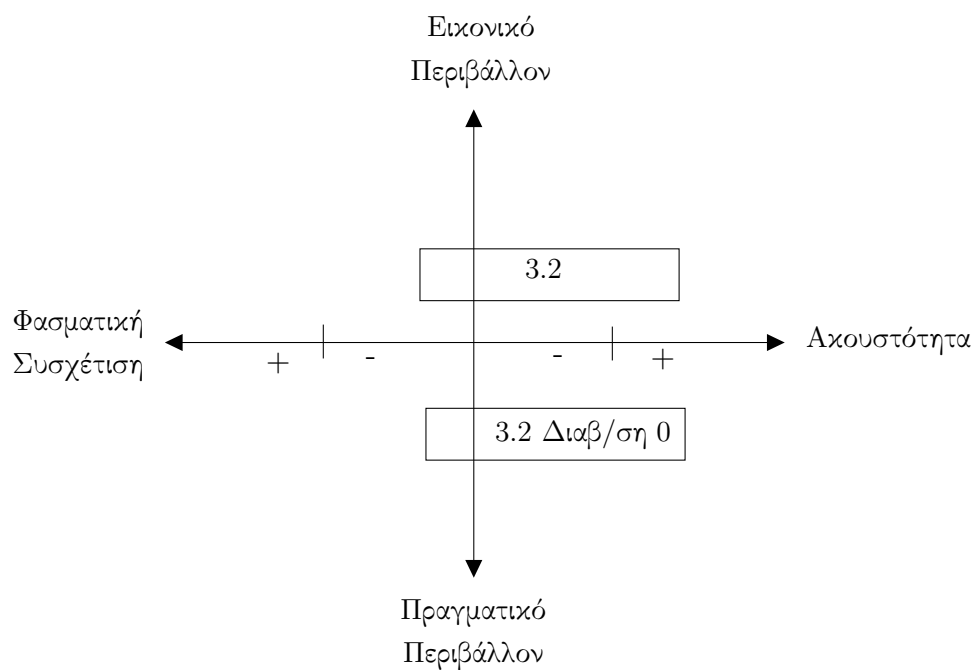
Για το σενάριο 3.3 (βλπ Σχ.7.5), η ισορροπία της προσαρμοστικής μείξης υπερτερεί ως προς το εικονικό περιβάλλον. Στο γενικό σύνολο, η προσαρμοστική μείξη κινείται με αυξητικό ρυθμό ως προς την ακουστότητα, αλλά και ως προς την φασματική συσχέτιση. Το αποτέλεσμα της στατιστικής ανάλυσης είναι η στατιστική σημαντικότητα των αποτελεσμάτων ως προς το ήμισυ των υπόλοιπων εκδοχών μείξης. Σχετικά με τις διαβαθμίσεις μείξης, παρατηρείται ότι η «Διαβάθμιση 0» και 6 παρουσιάζει στατιστικά μη σημαντικό αποτέλεσμα σε τρεις από τις τέσσερις εκδοχές μείξης ενώ στη «Διαβάθμιση 3» σε όλες τις εκδοχές μείξης. Σε σύγκριση με τις προηγούμενες εκδοχές, παρατηρείται ότι η αύξηση του ποσοστού συσχέτισης του



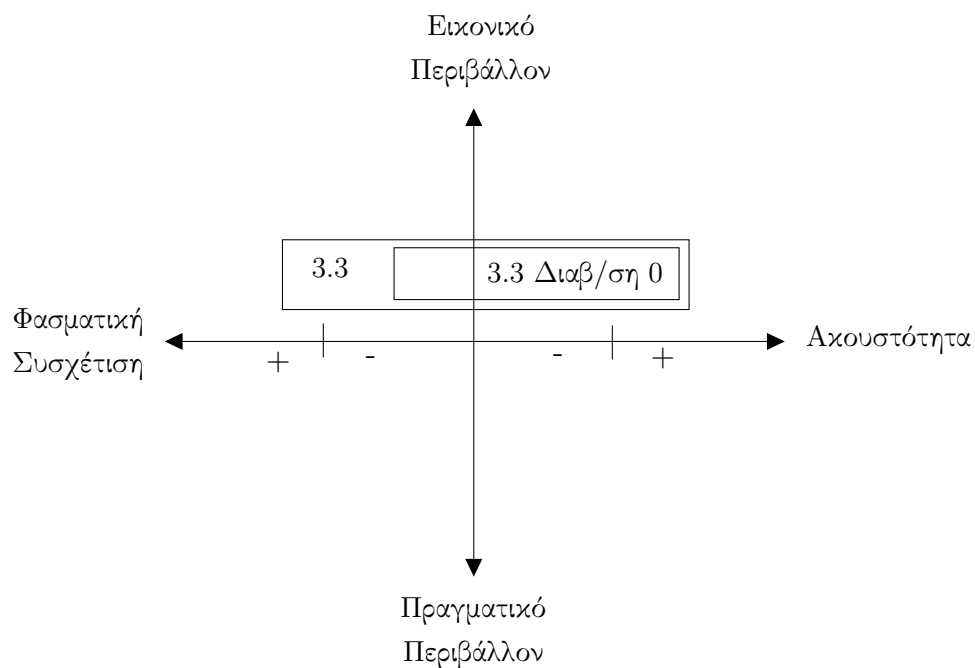
Σχήμα 7.2: Πεδίο ARA μείζης για σενάριο 3.1

φάσματος και της ακουστότητας ως προς το αισθητικό κομμάτι μείζης, δημιουργεί σύγχυση στα υποκείμενα. Σίγουρα η αισθητική αντίληψη της μείζης είναι υποκειμενική, ωστόσο, σε αυτή την περίπτωση, φαίνεται να μπερδεύεται για τον ακροατή ανάμεσα και στις υπόλοιπες εκδοχές μείζης.

Για το σενάριο 3.4 (βλπ Σχ.7.4), η ισορροπία μείζης μεταξύ του εικονικού και του πραγματικού περιβάλλοντος, τοποθετείται προς την πλευρά του εικονικού. Η προσαρμοστική μείξη για το γενικό σύνολο και για τις διαβαθμίσεις, τείνει αυξητικά ως προς την αμφιωτική ακουστότητα και με μείωση ως προς τη φασματική συσχέτιση. Στην στατιστική ανάλυση, διαπιστώνεται πως στο γενικό σύνολο παρουσιάζει όμοιο αποτέλεσμα με το σενάριο 3.2 με το οποίο έχουν κοινά στοιχεία σύμφωνα με την ανάλυση του Σχήματος 7.4 και 7.3 στα οποία οι περιπτώσεις εκτός της «Διαβάθμισης 6», συμπεριφέρονται όμοια. Και στις δύο περιπτώσεις το συνολικό αποτέλεσμα είναι στατιστικά σημαντικό σε όλες τις εκδοχές μείζης εκτός από εκείνη με εύρος 10 kHz (βλπ Πίν. 7.4 και Πίνακα 7.6).

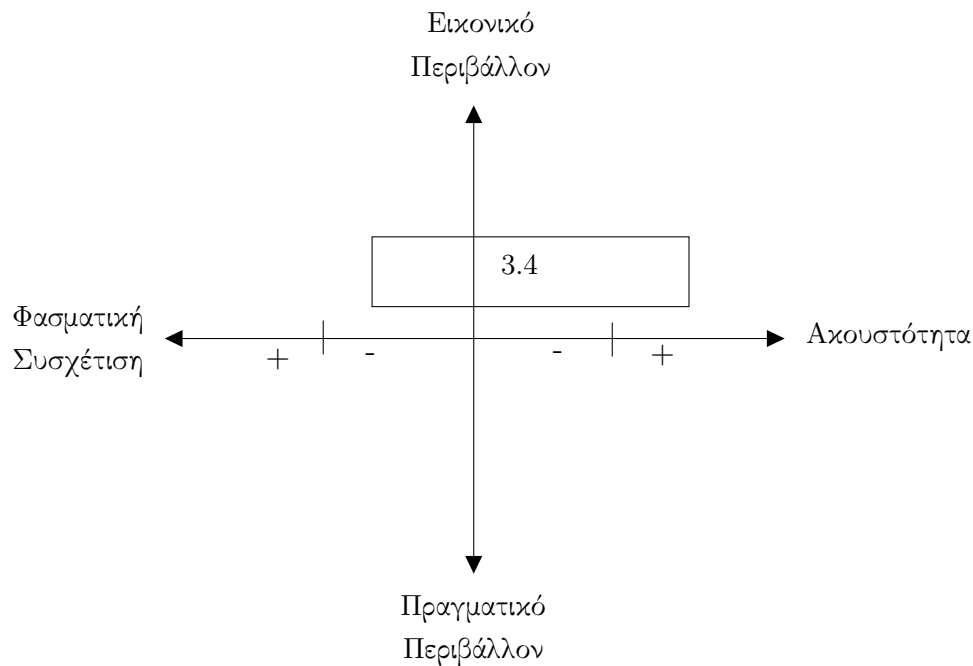


Σχήμα 7.3: Πεδίο ARA μείξης για σενάριο 3.2



Σχήμα 7.5: Πεδίο ARA μείξης για σενάριο 3.3

Συμπερασματικά, η βέλτιστη περίπτωση μείξης σύμφωνα με την στατιστική ανάλυση, αποτελεί το σενάριο 3.1 στο οποίο η ισορροπία της προσαρμοστικής μείξης τείνει προς το πραγματικό περιβάλλον και υπάρχει προσαρμογή της ακουστότητας αυξητικά. Οι επόμενες περιπτώσεις με σημαντικά αποτελέσματα, αποτελούν τα σενάρια 3.2 και 3.4, τα οποία, στο γενικό σύνολο, εμφανίζουν στο μεγαλύτερο ποσοστό τους (τρεις από τις 4 εκδοχές σύγκρι-



Σχήμα 7.4: Πεδίο ARA μείξης για σενάριο 3.4

σης μείξης με προσαρμοστική μείξης) στατιστικά σημαντικό αποτέλεσμα. Στην περίπτωση το σεναρίου 3.3, όπου παρατηρείται ταυτόχρονα αύξηση συσχέτισης και ακουστότητας με την ισορροπία μείξης να βρίσκεται προς το εικονικό περιβάλλον, παρατηρείται το λιγότερο σημαντικό αποτέλεσμα. Πιθανόν η αμφίδρομη αύξηση των δύο χαρακτηριστικών του εικονικού περιβάλλοντος κατά την προσαρμοστική μείξη, δημιουργεί μια σύγχυση στην αντίληψη του ακουστικού αποτελέσματος επηρεάζοντας την ποιότητα της προσαρμοστικής μείξης.

Η «Διαβάθμιση 6» σύμφωνα με την στατιστική ανάλυση των αποτελεσμάτων με την χρήση των t-Test σε όλα τα σενάρια, δεν αποτελεί αισθητικά την ιδανική μείξη, καθώς συγχέεται υποκειμενικά και με τις υπόλοιπες εκδοχές. Η «Διαβάθμιση 3» αποτελεί μια επιλογή μείξης με μέτρια απόδοση, καθώς παρατηρείται πως, για παράδειγμα, στα σενάρια 3.2 και 3.3 δεν είναι αποδεκτή, σε αντίθεση με το σενάριο 3.1, στο οποίο διαπιστώνεται σημαντική διαφοροποίηση και στο σενάριο 3.2, το οποίο διαφοροποιείται συγκριτικά με την στατιστική ανάλυση του 3.1 σε μια εκδοχή μείξης.

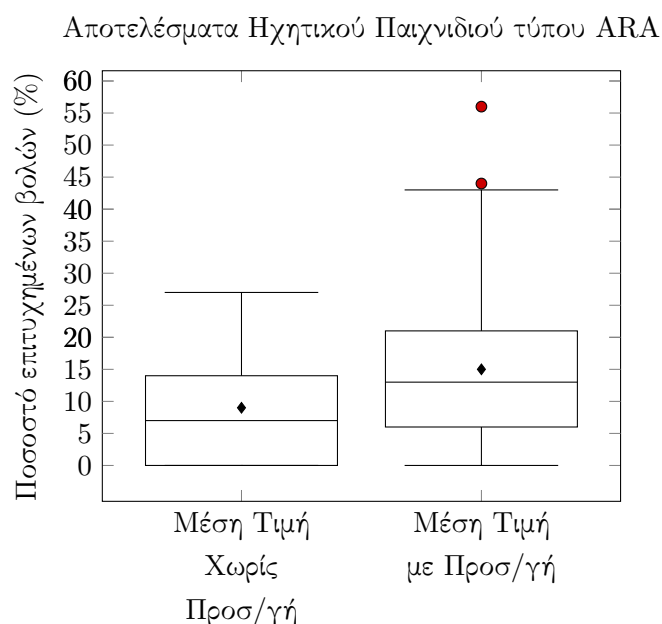
Σχετικά με τις κατηγορίες διάκρισης εμπειρίας, η προσαρμοστική μείξη, φαίνεται να ευνοείται από την εμπειρία του ακροατή καθώς παρατηρείται πως η κατηγορία των ειδικών σημειώνει σημαντική διαφορά υπέρ της προσαρμοστικής μείξης σε όλα τα σενάρια. Αυτό μπορεί να το εξηγήσει η ακουστική αντίληψη που έχει δημιουργήσει ένα πιο «έμπειρο» αυτί στις ακουστικές μεταβολές. Ωστόσο, στο σενάριο 3.1, η άποψη των ειδικών συμπλέει με την άποψη των έμπειρων και των αρχάριων σε πολύ μεγάλο βαθμό (βλπ Πίν. 7.3). Αυτή η παρατήρηση παρουσιάζει την δυνατότητα που μπορεί να δώσει η προσαρμοστική μείξη, ώστε να μην εξαρτάται η ακουστική αντίληψη από την εμπειρία, αλλά από την κατάλληλη προσαρμογή. Στην περίπτωση του σεναρίου 3.1 φαίνεται το χαρακτηριστικό της ισορροπίας της μείξης να τείνει προς το πραγματικό περιβάλλον (βλπ Σχ. 7.2) που δεν υπάρχει στα υπόλοιπα σενάρια και το οποίο φαίνεται

να ικανοποιεί το σύνολο των υποκειμένων ανεξαρτήτου εμπειρίας. Επιπροσθέτως, κατά την αξιολόγηση, υπήρξε η άποψη στο τέλος της διαδικασίας από μερικά υποκείμενα, πως η κρίση της μείξης ήταν με βάση το πόσο καλύπτει το ένα περιβάλλον το άλλο, και καταγράφηκε η άποψη ότι είναι αρεστή η μείξη όταν η ισορροπία της μείξης τείνει προς το πραγματικό.

#### 7.4 Αποτελέσματα υποκειμενικής αξιολόγησης Ηχητικού Παιχνιδιού τύπου ARA

Η υποκειμενική αξιολόγηση, αποτελεί το τελευταίο μέρος της υποκειμενικής αξιολόγησης του μοντέλου προσαρμοστικής μείξης σε επίπεδο πλέον εφαρμογής σε σενάριο ηχητικού παιχνιδιού επαυξημένης πραγματικότητας ήχου. Στη διαδικασία αξιολόγησης του παιχνιδιού, πήραν μέρος οι 78 συμμετέχοντες της παραπάνω διαδικασίας αξιολόγησης. Οι κατηγορίες σύμφωνα με τις οποίες έγινε η στατιστική ανάλυση ακολουθεί κοινή μεθοδολογία με τα τέστ της προηγούμενης ενότητας. Αναλυτικότερα, θα παρουσιαστούν οι κατανομές των αποτελεσμάτων για το γενικό σύνολο, καθώς και για τις κατηγορίες σύμφωνα με την διάκριση εμπειρίας και τη διαβάθμιση μείξης για την έκδοση του παιχνιδιού με προσαρμοστική μείξη και χωρίς. Ακολούθως, δίνονται τα ποσοστά επιτυχίας για την κάθε κατηγορία υποκειμένων για την εκάστοτε έκδοση του παιχνιδιού, καθώς και τα ποσοστά χαμένων ευκαιριών σύμφωνα με την μεθοδολογία που περιγράφηκε στην Ενότητα 6.2.3. Τέλος, παρουσιάζεται και η στατιστική ανάλυση των αποτελεσμάτων μέσα από την διεξαγωγή t-Test για την διερεύνηση η μη της στατιστικής σημαντικότητας των αποτελεσμάτων.

Η παιγνιακή δραστηριότητα, διήρκεσε για όλα τα υποκείμενα 100 δευτερόλεπτα. Δεν υπήρξε προηγούμενη εκπαίδευση παρά μόνο μια σύντομη γραπτή περιγραφή των οδηγιών. Η επιλογή της σειράς που το υποκείμενο θα έπαιζε τις δύο εκδόσεις του ηχητικού παιχνιδιού (με προσαρμοστική μείξη και χωρίς), ήταν τυχαία. Όλα τα υποκείμενα ολοκλήρωσαν την παιγνιακή δραστηριότητα. Στο γενικό σύνολο, η κατανομή των αποτελεσμάτων σύμφωνα με το ποσοστό των επιτυχημένων βολών, απεικονίζεται στο Σχήμα 7.6. Σύμφωνα με αυτό το σχήμα, για το γενικό σύνολο, το ποσοστό επιτυχημένων βολών για το μεγαλύτερο μέρος του συνόλου, κυμαίνεται για το παιχνίδι χωρίς προσαρμοστική μείξη μεταξύ του 0%-14% με μέση τιμή 9% και ενδιαμέση τιμή 7%. Αντίστοιχα για το ηχητικό παιχνίδι με ενεργοποιημένη την προσαρμοστική μείξη, το ποσοστό επιτυχημένων βολών για την πλειονότητα του γενικού συνόλου κυμαίνεται από 6%-21% με μέση τιμή 15% και ενδιαμέση στα 13%. Συγκρίνοντας τις δύο κατηγορίες μείξης στο γενικό σύνολο, διαπιστώνεται μια σημαντική αύξηση στο ποσοστό επιτυχημένων βολών για την προσαρμοστική μείξη με περιπτώσεις που έφτασαν στο 44% και 56%.

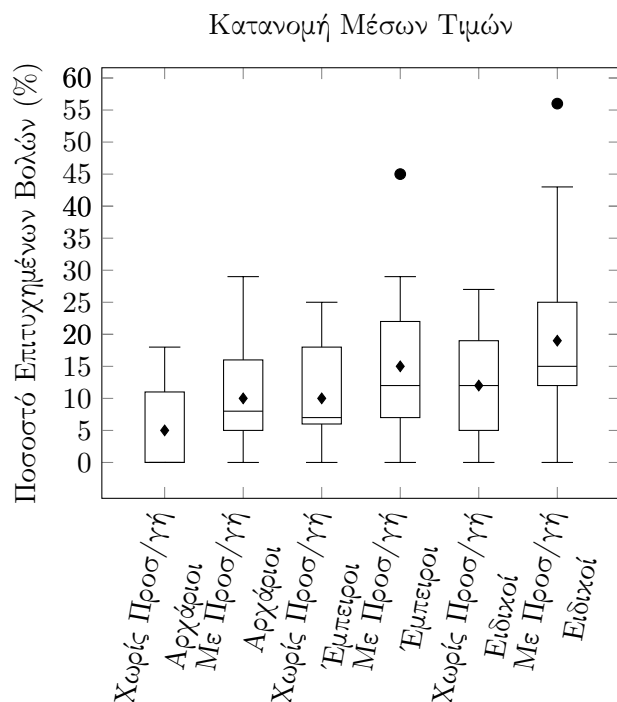


Σχήμα 7.6: Κατανομή Μέσης Τιμής Γενικού Συνόλου για κατηγορία προσαρμοστικής και μη προσαρμοστικής μείξης σε Ηχητικό Παιχνίδι τύπου ARA

Προχωρώντας στην παρουσίαση των αποτελεσμάτων, σειρά έχει η αποτύπωση της κατανομής των επιτυχημένων βολών μέσα από την κατηγοριοποίηση του γενικού συνόλου σύμφωνα με την εμπειρία των υποκειμένων (βλπ Σχ.7.7). Σε μια γενική επισκόπηση των αποτελεσμάτων, διακρίνεται η βελτίωση των επιτυχημένων βολών στην περίπτωση χρήσης της προσαρμοστικής μείξης. Ειδικότερα, τα σύνολα με προσαρμοστική μείξη και χωρίς για κάθε κατηγορία εμπειρίας εμφανίζουν τις ακόλουθες τιμές. Στην κατηγορία των «Αρχαρίων», τα υποκείμενα έχουν ποσοστό επιτυχίας από 0%-11% στη μη προσαρμοστική μείξη ενώ στην περίπτωση της προσαρμοστικής, το ποσοστό κυμαίνεται από 5%-16%. Σχετικά με τις τιμές του μέσου όρου και της ενδιάμεσης τιμής, παρατηρείται πως στη μη εφαρμογή της προσαρμοστικής μείξης, η μέση τιμή είναι 5% και η ενδιάμεση 0% σε αντιδιαστολή με την περίπτωση της προσαρμοστικής όπου η ενδιάμεση τιμή είναι 8% και η μέση τιμή 10%. Κατά συνέπεια στην περίπτωση των «Αρχαρίων», φαίνεται πως η προσαρμοστική μείξη δημιουργεί μια συνθήκη για καλύτερη απόδοση του υποκειμένου στο σενάριο του ηχητικού παιχνιδιού.

Στην κατηγορία «Έμπειροι», στην εκδοχή του παιχνιδιού χωρίς την προσαρμοστική μείξη, η μέση τιμή του ποσοστού επιτυχίας είναι 10% και η ενδιάμεση τιμή της κατανομής είναι 7% ενώ στην περίπτωση με την προσαρμοστική μείξη, η μέση τιμή αυξάνεται στο 15% και η ενδιάμεση τιμή στο 12% και επιπροσθέτως υπάρχει και μια περίπτωση με ποσοστό 45%. Σε σύγκριση με την κατηγορία των «Αρχαρίων», φαίνεται πως υπάρχει μια διαφοροποίηση ανοδική και για τους δύο τρόπου μείξης διακρίνοντας όμως αυξημένο ποσοστό επιτυχίας στην περίπτωση της προσαρμοστικής μείξης όπου φαίνεται πως έχει βοηθήσει στην βελτίωση της απόδοσης του υποκειμένου.

Στην κατηγορία «Ειδικοί», ο μεγαλύτερος όγκος των αποτελεσμάτων επιτυχίας χωρίς το μοντέλο προσαρμογής κυμαίνονται από 5%-19% με μέση και ενδιάμεση τιμή στο 12%. Στην

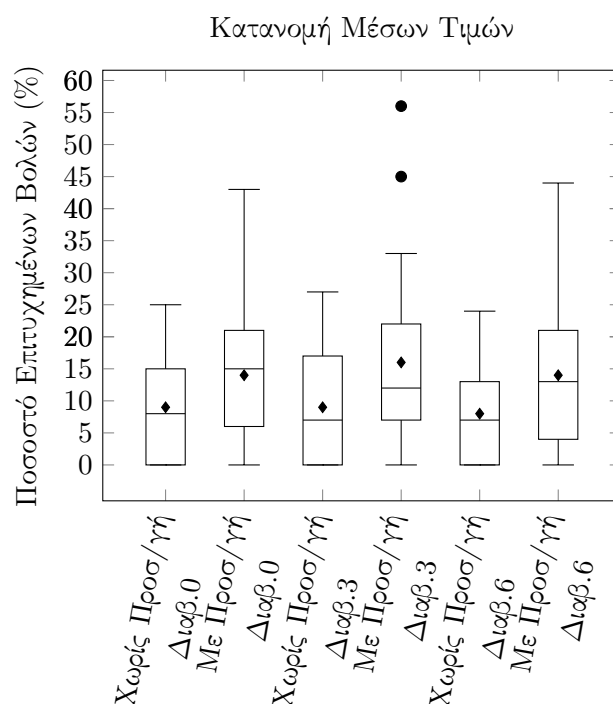


Σχήμα 7.7: Κατανομή Μέσων Τιμών Επιτυχίας Ηχητικού Παιχνιδιού τύπου ARA για κατηγορίες με Διάκριση Εμπειρίας

περίπτωση του μοντέλου προσαρμοστικής μείξης, η μέση τιμή αυξάνεται στο 19% και η ενδιάμεση στο 15% και κύριος όγκος των αποτελεσμάτων, κυμαίνεται από 12%-25%. Συγκριτικά με την προηγούμενη εκδοχή μείξης για αυτή την κατηγορία, διακρίνεται βελτίωση της κατανομής των αποτελεσμάτων όταν εφαρμόζεται η προσαρμοστική μείξη. Σχετικά με την σύγκριση μεταξύ των κατηγοριών, φαίνεται πως η εμπειρία του υποκειμένου επηρεάζει την παιγνιακή του απόδοση μέσα στο επαυξημένο περιβάλλον ήχου του ηχητικού παιχνιδιού αλλά και η εισαγωγή της προσαρμοστικής μείξης λειτουργεί ευεργετικά προς την βελτίωση του ποσοστού επιτυχίας. Σε όλες τις κατηγορίες υπάρχουν περιπτώσεις με πολύ χαμηλά ποσοστά επιτυχίας ωστόσο, λαμβάνοντας υπόψιν τον κύριο όγκο του δείγματος στην κάθε κατηγορία, το αποτέλεσμα τείνει θετικά προς την εκδοχή της προσαρμοστικής μείξης.

Το επόμενο μέρος των αποτελεσμάτων που παρουσιάζονται, κατηγοριοποιούνται σύμφωνα με τη διαβάθμιση μείξης που έχει δοθεί στην προσαρμοστική μείξη. Στην κάθε κατηγορία διαβαθμισμένης μείξης, παρουσιάζεται για το ίδιο σύνολο το ποσοστό επιτυχίας χωρίς την χρήση του μοντέλου προσαρμοστικής μείξης για να υπάρχει μια κοινή συνισταμένη σύγκρισης (βλπ Σχ.7.8). Στην πρώτη κατηγορία με «Διαβάθμιση 0», η κατανομή του συνόλου χωρίς την χρήση του προσαρμοστικού μοντέλου, παρουσιάζει μέση τιμή επιτυχημένων βολών 9% με ενδιάμεση τιμή κατανομής 8% με το σύνολο των αποτελεσμάτων να κυμαίνονται από 0%-15%. Στην περίπτωση της εισαγωγής του μοντέλου προσαρμογής, υπάρχει διεύρυνση του ποσοστού επιτυχίας με τον κύριο όγκο του να κατανέμεται μεταξύ των τιμών 6%-21% με μέση τιμή 14% και ενδιάμεση 15% παρουσιάζοντας αύξηση σε σύγκριση με την προηγούμενη εκδοχή. Η κα-

τηγορία «Διαβάθμισης 3», εμφανίζει για το σύνολο χωρίς την προσαρμοστική μείξη ποσοστά επιτυχίας από 0%-17% με ενδιάμεση τιμή 7% και μέση τιμή 9%. Η εισαγωγή της προσαρμοστικής μείξης με «Διαβάθμιση 3», αυξάνει τη μέση τιμή στο 16% και την ενδιάμεση στο 12% με το σύνολο να κυμαίνεται σε ποσοστά επιτυχίας μεταξύ 7%-22% με περιπτώσεις υψηλών ποσοστών επιτυχίας στα 45% και 56%. Στην τρίτη κατηγορία με «Διαβάθμιση 6», τα ποσοστά επιτυχίας χωρίς την προσαρμοστική μείξη, παρουσιάζουν μέση τιμή 8% και ενδιάμεση 7% και η κατανομή του κυρίως συνόλου κυμαίνεται από 0%-13%. Στην κατηγορία «Διαβάθμιση 6» με την εισαγωγή της προσαρμοστικής μείξης, η μέση τιμή υπολογίζεται στο 14% και η ενδιάμεση τιμή στα 13% με κύρια κατανομή μεταξύ του 4% και 21% εμφανίζοντας μια αυξητική τάση σε σύγκριση με την περίπτωση χωρίς την προσαρμοστική μείξη.



Σχήμα 7.8: Κατανομή Μέσων Τιμών Επιτυχίας Ηχητικού Παιχνιδιού τύπου ARA για κατηγορίες Διαβαθμισμένης Μείξης

Από τα αποτελέσματα εφαρμογής της διαβάθμισης στην προσαρμοστική μείξη, φαίνεται αρχικά πως σε σύγκριση με την περίπτωση χωρίς προσαρμοστική μείξη, υπάρχει μια διαφοροποίηση με θετικά αποτελέσματα για την προσαρμοστική μείξη. Μεταξύ των διαβαθμίσεων, η «Διαβάθμιση 3» εμφανίζει μεγαλύτερα ποσοστά για το 50% της κατανομής μεταξύ 7%-22% έναντι 6%-21% για την «Διαβάθμιση 0» και 4%-21% για την «Διαβάθμιση 6». Σχετικά με τα αποτελέσματα του μέσου όρου, την υψηλότερη τιμή συγκεντρώνει η «Διαβάθμιση 3» με 16% σε σύγκριση με την «Διαβάθμιση 0» και «Διαβάθμιση 6» στο 14%. Η ενδιάμεση τιμή των κατανομών φαίνεται να είναι υψηλότερη στην «Διαβάθμιση 0» στο 12% ενώ έπεται η «Διαβάθμιση 6» στο 13% και «Διαβάθμιση 3» στο 12%. Μεταξύ των τριών διαβαθμίσεων, και λαμβάνοντας υπόψιν σε μεγάλο βαθμό την κατανομή των αποτελεσμάτων, η «Διαβάθμιση 0» και η «Δια-

βάθμιση 3» φαίνεται να εμφανίζουν ένα μικρό προβάδισμα σε σύγκριση με την «Διαβάθμιση 6». Ωστόσο, πρέπει να επισημανθεί πως, λόγω του στενού χρονικού πλαισίου διενέργειας της παιγνιακής δραστηριότητας, ακόμα και οι μικρές ποσοστιαίες διαφορές αποτελούν σημαντικό αποτέλεσμα για την αξιολόγηση του μοντέλου.

Αναλύοντας επιπλέον το ποσοστό επιτυχίας, στους Πίνακες 7.7 και 7.8, παρατίθενται η μέση τιμή των ποσοστών επιτυχημένων βολών ως προς τις συνολικές ευκαιρίες για την κάθε κατηγορία εμπειρίας και για το γενικό σύνολο. Στο γενικό σύνολο, η διαφορά της μέσης τιμής των επιτυχημένων βολών είναι 1.12 και σε ποσοστό σύγκρισης, το μοντέλο προσαρμογής αυξάνει κατά 41% την επιτυχία των βολών μεταξύ του παιχνιδιού χωρίς προσαρμοστική μείξη και με την εισαγωγή αυτής. Στις κατηγορίες σύμφωνα με διάκριση εμπειρίας, για τους «Αρχάριους», η διαφορά του ποσοστού αγγίζει στο 5.11% και σε σχέση με τον αριθμό επιτυχημένων βολών εμφανίζεται αύξηση 49% από 0.82 σε 1.81 για το μοντέλο προσαρμοστικής μείξης. Στην κατηγορία «Έμπειροι», το μοντέλο προσαρμογής φαίνεται να έχει αυξήσει τον αριθμό των επιτυχημένων βολών από 1.60 σε 2.46 με ποσοστιαία αύξηση 29%. Στην κατηγορία «Ειδικοί», η μέση τιμή επιτυχημένων βολών είναι 1,98 χωρίς το μοντέλο προσαρμοστικής μείξης ενώ με την εισαγωγή του, η μέση τιμή είναι 3,25 βολές με ποσοστό αύξησης 38%. Σύμφωνα με την παραπάνω ανάλυση, το μοντέλο έχει βοηθήσει τις κατηγορίες διάκρισης εμπειρίας από 29% μέχρι 49% ως προς τον αριθμό των επιτυχημένων βολών. Το μεγαλύτερο ποσοστό επιτυχίας εντοπίζεται στους αρχάριους με 49% κατόπιν στους ειδικούς με 38% και τέλος στους έμπειρους με 29%. Κατά συνέπεια φαίνεται πως οι αρχάριοι βοηθήθηκαν από το μοντέλο προσαρμοστικής μείξης σε μεγαλύτερο βαθμό ως προς την αντίληψη του επαυξημένου περιβάλλοντος, βελτιώνοντας τα αποτελέσματα της διάδρασής τους με το ηχητικό παιχνίδι επαυξημένης πραγματικότητας.

| Ηχητικό Παιχνίδι τύπου ARA | Γενικό Σύνολο         |                           | Αρχάριοι              |                           | Έμπειροι              |                           | Ειδικοί               |                           |
|----------------------------|-----------------------|---------------------------|-----------------------|---------------------------|-----------------------|---------------------------|-----------------------|---------------------------|
|                            | Επιτυχία Βολών (Μ.Τ.) | Επιτυχημένες Βολές (Μ.Τ.) | Επιτυχία Βολών (Μ.Τ.) | Επιτυχημένες Βολές (Μ.Τ.) | Επιτυχία Βολών (Μ.Τ.) | Επιτυχημένες Βολές (Μ.Τ.) | Επιτυχία Βολών (Μ.Τ.) | Επιτυχημένες Βολές (Μ.Τ.) |
| Παιχνίδι χωρίς Προσαρμογή  | 8.68 %                | 1.38                      | 5.22 %                | 0.82                      | 10.28 %               | 1.60                      | 11.86 %               | 1.98                      |
| Παιχνίδι με Προσαρμογή     | 14.59 %               | 2.50                      | 10.33 %               | 1.81                      | 14.52 %               | 2.46                      | 19.05 %               | 3.25                      |

Πίνακας 7.7: Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών για κατηγορίες με Διάκριση Εμπειρίας

Ακολούθως για τις κατηγορίες με διαβάθμιση μείξης (βλπ Πίν. 7.9 και 7.10), για την κατηγορία «Διαβάθμιση 0», οι επιτυχημένες βολές παρουσιάζουν μια αύξηση για το μοντέλο προσαρμογής 40% από 8.65% μέση τιμή επιτυχίας βολών σε 14.38%. Στην «Διαβάθμιση 3», η εκδοχή της προσαρμοστικής μείξης καταγράφει μέση τιμή επιτυχίας βολών 15.69% έχο-

| Ηχητικό Παιχνίδι τύπου ARA | Γενικό Σύνολο         |         | Αρχάριοι              |         | Έμπειροι              |         | Ειδικοί               |         |
|----------------------------|-----------------------|---------|-----------------------|---------|-----------------------|---------|-----------------------|---------|
|                            | Επιτυχία Βολών (Μ.Τ.) | Διαφορά | Επιτυχία Βολών (Μ.Τ.) | Διαφορά | Επιτυχία Βολών (Μ.Τ.) | Διαφορά | Επιτυχία Βολών (Μ.Τ.) | Διαφορά |
| Παιχνίδι χωρίς Προσαρμογή  | 8.68 %                | +41%    | 5.22%                 | +49%    | 10.28%                | +29%    | 11.86%                | +38%    |
| Παιχνίδι με Προσαρμογή     | 14.59 %               |         | 10.33%                |         | 14.52%                |         | 19.05%                |         |

Πίνακας 7.8: Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών και Ποσοστιαία Διαφορά για κατηγορίες Εμπειρίας

ντας αύξηση 42% από τη μέση τιμή επιτυχίας βολών 9.08%. Στη «Διαβάθμιση 6», η μέση τιμή επιτυχίας βολών για το ηχητικό παιχνίδι χωρίς προσαρμοστική μείξη, είναι 8.31% και στο αντίστοιχο με προσαρμοστική μείξη είναι 13.69% με ποσοστό αύξησης των βολών 39%. Συγκριτικά, σε όλες τις κατηγορίες διαβάθμισης παρουσιάζεται βελτίωση της μέσης τιμής επιτυχίας βολών σε ποσοστό 39%-42%. Ωστόσο, τη μεγαλύτερη μέση τιμή επιτυχημένων βολών παρουσιάζει η κατηγορία με «Διαβάθμιση 3» και ακολουθεί η κατηγορία με «Διαβάθμιση 0» και τέλος «Διαβάθμιση 6».

| Ηχητικό Παιχνίδι<br>τύπου ARA | Διαβάθμιση 0             |                              | Διαβάθμιση 3             |                              | Διαβάθμιση 6             |                              |
|-------------------------------|--------------------------|------------------------------|--------------------------|------------------------------|--------------------------|------------------------------|
|                               | Επιτυχία<br>Βολών (Μ.Τ.) | Επιτυχημένες<br>Βολές (Μ.Τ.) | Επιτυχία<br>Βολών (Μ.Τ.) | Επιτυχημένες<br>Βολές (Μ.Τ.) | Επιτυχία<br>Βολών (Μ.Τ.) | Επιτυχημένες<br>Βολές (Μ.Τ.) |
| Παιχνίδι χωρίς<br>Προσαρμογή  | 8.65 %                   | 1.36                         | 9.08 %                   | 1.44                         | 8.31 %                   | 1.35                         |
| Παιχνίδι με<br>Προσαρμογή     | 14.38 %                  | 2.47                         | 15.69 %                  | 2.61                         | 13.69 %                  | 2.42                         |

Πίνακας 7.9: Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών για κατηγορίες Διαβαθμισμένης Μείξης

Στον Πίνακα 7.11, παρουσιάζεται η μέση τιμή των αποτελεσμάτων από τις χαμένες ευκαιρίες για κάθε υποκατηγορία του γενικού συνόλου του ηχητικού παιχνιδιού. Με μια πρώτη προσέγγιση, φαίνεται πως σε όλες τις κατηγορίες παιχνιδιού χωρίς προσαρμοστική μείξη, παρουσιάζεται αυξημένη η μέση τιμή χαμένων ευκαιριών σε σύγκριση με το παιχνίδι με προσαρμοστική μείξη. Αυτό δικαιολογείται και από τα προηγούμενα αποτελέσματα επιτυχημένων βολών. Παρόλα αυτά, παρατηρείται πως οι «Ειδικοί» παρουσιάζουν την μικρότερη τιμή χαμένων ευκαιριών από τις υπόλοιπες κατηγορίες διάκρισης εμπειρίας, καθώς και στην περίπτωση της «Διαβάθμισης 6», εμφανίζεται η μικρότερη μέση τιμή σε σχέση με τις υπόλοιπες διαβαθμίσεις.

Η εισαγωγή της διαβάθμισης στην προσαρμοστική μείξη μπορεί να ληφθεί υπόψιν ως ένας παράγοντας μείωσης της δυσκολίας του παιχνιδιού καθώς επηρεάζει θετικά τη διάδραση του παίκτη με το παιγνιακό περιβάλλον του ηχητικού παιχνιδιού. Η αύξηση της διαβάθμισης δεν μειώνει τη δυσκολία του παιχνιδιού ως προς την διάδραση του παίκτη με αυτό. Φαίνεται πως η μέγιστη μείωση της δυσκολίας για τον παίκτη, επιτυγχάνεται στη «Διαβάθμιση 3» που αποτελεί τη μεσαία διαβάθμιση της προσαρμοστικής μείξης.

Στην συνέχεια της ανάλυσης των αποτελεσμάτων, πραγματοποιήθηκε στατιστική ανάλυση τύπου t-Test (βλπ Πίν. 7.12) για την διερεύνηση της στατιστικής σημαντικότητας του ποσοστού επιτυχημένων βολών. Σε αυτή την ανάλυση παρατηρείται πως σε όλες τις κατηγορίες τα αποτελέσματα είναι στατιστικά σημαντικά ( $p < .05$ ). Ωστόσο, στην κατηγορία «Εμπειροί»

| Ηχητικό Παιχνίδι τύπου ARA | Διαβάθμιση 0          |         | Διαβάθμιση 3          |         | Διαβάθμιση 6          |         |
|----------------------------|-----------------------|---------|-----------------------|---------|-----------------------|---------|
|                            | Επιτυχία Βολών (Μ.Τ.) | Διαφορά | Επιτυχία Βολών (Μ.Τ.) | Διαφορά | Επιτυχία Βολών (Μ.Τ.) | Διαφορά |
| Παιχνίδι χωρίς Προσαρμογή  | 8.65%                 | +40%    | 9.08%                 | +42%    | 8.31%                 | +39%    |
| Παιχνίδι με Προσαρμογή     | 14.38%                |         | 15.69%                |         | 13.69%                |         |

Πίνακας 7.10: Αποτελέσματα Παιχνιδιού - Επιτυχία Βολών και Ποσοστιαία Διαφορά για κατηγορίες με Διαβάθμιση μείξης

| Αποτελέσματα Χαμένων Ευκαιριών Ηχητικού Παιχνιδιού τύπου ARA |               |          |          |         |              |              |              |
|--------------------------------------------------------------|---------------|----------|----------|---------|--------------|--------------|--------------|
|                                                              | Γενικό Σύνολο | Αρχάριοι | Έμπειροι | Ειδικοί | Διαβάθμιση 0 | Διαβάθμιση 3 | Διαβάθμιση 6 |
| Παιχνίδι χωρίς Προσαρμογή                                    | 5.67          | 5.81     | 5.60     | 5.27    | 6.15         | 5.92         | 4.92         |
| Παιχνίδι με Προσαρμογή                                       | 4.33          | 4.41     | 4.72     | 3.86    | 4.31         | 4.88         | 3.81         |

Πίνακας 7.11: Αποτελέσματα Χαμένων Ευκαιριών Ηχητικού Παιχνιδιού τύπου ARA

παρατηρείται η τιμή  $p > .05$ , οπότε θεωρείται ως μη σημαντικό το αποτέλεσμα συγκριτικά μεταξύ του παιχνιδιού με προσαρμοστική μείξη και χωρίς.

| Ηχητικό Παιχνίδι | Γενικός Μ.Ο. |           | Αρχάριοι |           | Έμπειροι |           | Ειδικοί |           | Διαβάθμιση 0 |           | Διαβάθμιση 3 |           | Διαβάθμιση 6 |           |
|------------------|--------------|-----------|----------|-----------|----------|-----------|---------|-----------|--------------|-----------|--------------|-----------|--------------|-----------|
|                  | M.T.         | $p < .05$ | M.T.     | $p < .05$ | M.T.     | $p < .05$ | M.T.    | $p < .05$ | M.T.         | $p < .05$ | M.T.         | $p < .05$ | M.T.         | $p < .05$ |
| Χωρίς Προσ/γή    | 8.68 %       |           | 5.22 %   |           | 10.28 %  |           | 11.86 % |           | 8.65 %       |           | 9.08 %       |           | 8.31 %       |           |
| Με Προσ/γή       | 14.59 %      | 0.003     | 10.33 %  | 0.01      | 14.52 %  | 0.12      | 19.05 % | 0.03      | 14.38 %      | 0.02      | 15.69 %      | 0.04      | 13.69 %      | 0.04      |

Πίνακας 7.12: Αποτελέσματα Μέσης Τιμής Ηχητικού Παιχνιδιού - t-Test

Συνοψίζοντας τα αποτελέσματα αξιολόγησης του ηχητικού παιχνιδιού επαυξημένης πραγματικότητας ήχου στην εκδοχή με την εισαγωγή του προσαρμοστικού μοντέλου μείξης και χωρίς γίνονται εμφανή τα παρακάτω στοιχεία. Αρχικά, στο γενικό σύνολο η προσαρμοστική μείξη φαίνεται να λειτουργεί ευεργετικά για την απόδοση του υποκειμένου μέσα στο παιγνιακό περιβάλλον στα πλαίσια της επαυξημένης πραγματικότητας ήχου. Αυξάνει το ποσοστό επιτυχίας του υποκειμένου και κατά συνέπεια εμφανίζεται και αυξημένος αριθμός επιτυχημένων βολών. Οι κατηγορίες διάκρισης εμπειρίας παρουσιάζουν βελτίωση των ποσοστών τους με την χρήση του μοντέλου, ακολουθώντας μεγαλύτερη αύξηση αποτελεσμάτων για τους «Αρχάριους», κατόπιν για τους «Ειδικούς» και τέλος για τους «Έμπειρους». Αυτό δείχνει πως η εμπειρία δημιουργεί ένα επιπλέον παράγοντα απόδοσης στο ηχητικό παιχνίδι επαυξημένης πραγματικότητας ήχου όπου το προσαρμοστικό μοντέλο φαίνεται να ωφελεί περισσότερο τους αρχάριους. Σχετικά με την κατηγορία διαβάθμισης της μείξης, μπορεί να σημειωθεί πως σε όλες τις περιπτώσεις διαβάθμισης φαίνεται μια βελτίωση της απόδοσης του υποκειμένου. Ωστόσο, η «Διαβάθμιση 3» παρουσιάζουν σχετικά βελτιωμένο αποτέλεσμα από την «Διαβάθμιση 0» «Διαβάθμιση 6». Κατά συνέπεια, η εισαγωγή της διαβάθμισης, φαίνεται να αποτελεί σημαντικό κομμάτι της προσαρμοστικής μείξης με ουσιαστικό ρόλο επηρεάζοντας τον τρόπο λήψης απόφασης και διάδρασης του υποκειμένου. Επιπροσθέτως φαίνεται να είναι ένας παράγοντας καθορισμού της δυσκολίας του ηχητικού παιχνιδιού, καθώς φαίνεται πως η αλλαγή της διαβάθμισης μπορεί αυξήσει ή να μειώσει τη δυσκολία του παιχνιδιού ως προς τις αποφάσεις που θα πάρει ο παίκτης κατά τη διάδρασή του με το παιχνίδι. Η «Διαβάθμιση 3» φαίνεται να κάνει το ηχητικό παιχνίδι πιο εύκολο για τους παίκτες σε σύγκριση με τα αποτελέσματα της «Διαβάθμισης 0» και «Διαβάθμισης 6».

## 7.5 Σύνοψη Κεφαλαίου

Σε αυτό το κεφάλαιο παρουσιάστηκαν τα αποτελέσματα της αξιολόγησης του μοντέλου προσαρμοστικής μείξης για περιβάλλοντα επαυξημένης πραγματικότητας ήχου. Η αξιολόγηση του μοντέλου, δομήθηκε σε τρεις άξονες. Ο πρώτος άξονας είναι η αντικειμενική αξιολόγηση

του μοντέλου, ο δεύτερος αποτελεί την υποκειμενική αξιολόγηση του μοντέλου και ο τρίτος η εφαρμοσμένη χρήση σε ένα ηχητικό παιχνίδι επαυξημένης πραγματικότητας ήχου. Για τον πρώτο άξονα, ακολουθήθηκε η μεθοδολογία αξιολόγησης που περιγράφηκε στην Ενότητα 6.2.1. Σύμφωνα με την επιλογή τεσσάρων σεναρίων επαυξημένης μείξης ήχου, χρησιμοποιήθηκαν δύο μέθοδοι αντικειμενικής αξιολόγησης της προσαρμοστικής μείξης. Με την πρώτη μέθοδο, έγινε υπολογισμός της συσχέτισης του συχνотικού φάσματος μεταξύ του ανεπεξέργαστου εικονικού περιβάλλοντος, του προσαρμοσμένου εικονικού περιβάλλοντος και του πραγματικού περιβάλλοντος για όλα τα επίπεδα διαβάθμισης της προσαρμοστικής μείξης (Διαβάθμιση 0, 3, 6). Τα αποτελέσματα συσχέτισης για τα τέσσερα σενάρια μεταξύ του προσαρμοσμένου εικονικού περιβάλλοντος και του πραγματικού περιβάλλοντος, είναι μεικτά, καθώς στο 50% των περιπτώσεων διαπιστώνεται αύξηση της συσχέτισης με την αύξηση του συντελεστή διαβάθμισης και στο υπόλοιπο 50% παρατηρείται υψηλότερη συσχέτιση στην προσαρμοστική μείξη με «Διαβάθμιση 6». Η δεύτερη μέθοδος, χρησιμοποιεί το μοντέλο ακουστότητας GM για τον υπολογισμό της ακουστότητας κάθε περιβάλλοντος για κάθε διαβάθμιση προσαρμοστικής μείξης. Σύμφωνα με το αποτέλεσμα των μετρήσεων, φαίνεται πως η προσαρμοστική μείξη επιτυγχάνει την ελάττωση της διαφοράς μεταξύ της ακουστότητας του πραγματικού περιβάλλοντος με το προσαρμοσμένο εικονικό σε σύγκριση με το εικονικό χωρίς προσαρμογή. Σαν αποτέλεσμα των παραπάνω μετρήσεων, ενώ η ακουστότητα μπορεί να τείνει να ισοσκελιστεί μεταξύ του προσαρμοσμένου εικονικού περιβάλλοντος και πραγματικού, δεν σημαίνει ότι φασματικά τα περιβάλλοντα θα γίνουν ίδια. Ο δεύτερος άξονας αξιολόγησης, αποτελεί την υποκειμενική αξιολόγηση του προσαρμοστικού μοντέλου. Τα αποτελέσματα της υποκειμενικής αξιολόγησης παρουσιάστηκαν μέσα από την ανάλυση των τεσσάρων τέστ τύπου MUSHRA. Τα αποτελέσματα της υποκειμενικής αξιολόγησης, έδειξαν μια αισθητική προτίμηση στην προσαρμοστική μείξη από το γενικό σύνολο των υποκειμένων. Η μέση τιμή αξιολόγησης της προσαρμοστικής μείξης ανάμεσα στις τιμές των υπόλοιπων εκδοχών μείξης, παρουσιάζει υψηλότερο αποτέλεσμα. Ωστόσο, στην στατιστική ανάλυση τύπου t-Test, στο γενικό σύνολο η διαφορά που υπάρχει εμφανίζει στατιστική σημαντικότητα. Στις κατηγορίες με διάκριση εμπειρίας, η κατηγορία των ειδικών εμφανίζει στατιστική σημαντικότητα για την προσαρμοστική μείξη στα περισσότερα σενάρια. Επίσης, στις κατηγορίες διαβάθμισης, στο 50% των σεναρίων η διαβάθμιση 0 και 3 εμφανίζει για την προσαρμοστική μείξη στατιστικά σημαντική διαφορά από τις υπόλοιπες εκδοχές. Πρέπει να σημειωθεί πως στην υποκειμενική αξιολόγηση, το κάθε υποκείμενο έκρινε με βάση την προσωπική του ακουστική αισθητική, δίνοντας στην προσαρμοστική μείξη προτίμηση χωρίς να υπάρχει πριν την αξιολόγηση κάποιο είδος εκπαίδευσης. Επιπλέον λόγω του τύπου του τεστ MUSHRA η μείξη αναφοράς στην προκειμένη περίπτωση η προσαρμοστική μείξη, ήταν κρυφή για το εκάστοτε υποκείμενο. Ο τρίτος άξονας αξιολόγησης, αποτελείται από την αξιολόγηση της προσαρμοστικής μείξης μέσα από την εφαρμογή της σε ένα ηχητικό παιχνίδι επαυξημένης πραγματικότητας ήχου. Τα αποτελέσματα του ηχητικού παιχνιδιού παρατίθενται μέσα από ένα σύνολο μεταβλητών που καταγράφονται κατά την διάρκεια της παιγνιακής εμπειρίας. Οι μεταβλητές που απεικονίζουν το αποτέλεσμα του ηχητικού παιχνιδιού αποτελούν οι επιτυχημένες βολές, το ποσοστό επιτυχίας και χαμένων ευκαιριών και η σύγκριση για τις δύο εκδοχές ηχητικού παιχνιδιού (με την χρήση προσαρμο-

στικού μοντέλου και χωρίς την χρήση αυτού). Το αποτέλεσμα επιτυχίας των βολών για την προσαρμοστική μείξη έδειξε σημαντική διαφορά υπέρ της προσαρμοστικής μείξης δείχνοντας την επίδραση που έχει στην διάδραση του παίκτη με το ηχητικό περιβάλλον, καθώς και στην γρηγορότερη αντίληψη των ακουστικών περιβαλλόντων. Επιπροσθέτως, τα αποτελέσματα κατηγοριοποιήθηκαν με βάση της διάκριση της εμπειρίας του κάθε παίκτη και την διαβάθμιση της προσαρμοστικής μείξης. Σχετικά με τη διάκριση εμπειρίας, παρουσιάζεται η μεγαλύτερη αύξηση των επιτυχημένων αποτελεσμάτων για την προσαρμοστική μείξη στην κατηγορία των «Αρχάριων» και κατόπιν για τους «Ειδικούς» και τους «Εμπειρους». Επιπροσθέτως, τα αποτελέσματα εμφανίζουν στατιστική σημαντικότητα για την κατηγορία των «Αρχάριων» και των «Ειδικών». Για τις κατηγορίες διαβαθμισμένης μείξης, παρουσιάζεται μια μικρή μείωση των επιτυχημένων βολών στην «Διαβάθμιση 6» σε σύγκριση με την «Διαβάθμιση 0» και «Διαβάθμιση 3». Η εισαγωγή της διαβάθμισης φαίνεται να λειτουργεί ως ένας παράγοντας μεταβολής της δυσκολίας του ηχητικού παιχνιδιού. Η αύξηση της διαβάθμισης δεν παρουσιάζει ανάλογη αύξηση στην επιτυχία των βολών καθώς φαίνεται πως η «Διαβάθμιση 3», μειώνει τη δυσκολία του παιχνιδιού σε σύγκριση με τη «Διαβάθμιση 6» και «Διαβάθμιση 0».

Συνοψίζοντας, πρέπει να σημειωθεί πως τα αποτελέσματα της αξιολόγησης, δείχνουν πως η προσαρμοστική μείξη ήχου σε περιβάλλοντα επαυξημένης πραγματικότητας ήχου, μπορεί να επηρεάσει την αισθητική αντίληψη της ποιότητας της μείξης για τα επαυξημένα περιβάλλοντα ήχου καθώς και να έχει σημαντική επίδραση στην διάδραση του ακροατή με το επαυξημένο περιβάλλον ήχου και στην βελτίωση της ακουστικής του αντίληψης ειδικά σε περιπτώσεις σεναρίων που χρειάζεται η αλληλεπίδραση ακροατή περιβάλλοντος, όπως είναι το παράδειγμα του ηχητικού παιχνιδιού επαυξημένης πραγματικότητας ήχου.

## Κεφάλαιο 8

# Συμπεράσματα

Σε αυτό το κεφάλαιο, παρουσιάζονται τα συμπεράσματα της εφαρμογής της προσαρμοστικής μείξης στα περιβάλλοντα επαυξημένης πραγματικότητας ήχου υπό το πρίσμα της μεθοδολογικής προσέγγισης που περιγράφηκε στην παρούσα διατριβή και σύμφωνα με τα αποτελέσματα της πειραματικής διαδικασίας. Η επαυξημένη πραγματικότητα ήχου αποτελεί ένα πεδίο το οποίο τα τελευταία χρόνια έχει επιδείξει μια στασιμότητα. Πιο αναλυτικά, η αρχή λειτουργίας της επαυξημένης πραγματικότητας ήχου στηρίζεται στην μείξη του πραγματικού περιβάλλοντος με το αντίστοιχο εικονικό περιβάλλον και μεταδίδεται στον ακροατή ως ηχητική επαύξηση του πραγματικού περιβάλλοντος. Λέγοντας στασιμότητα προηγουμένως, θεωρείται πως η εξέλιξη της επαυξημένης πραγματικότητας ήχου έχει περιοριστεί α) στην τεχνολογία αναπαραγωγής του πραγματικού περιβάλλοντος (ως ψευδοακουστικό) και την βελτίωση της διαφάνειας του ακουστικού εξοπλισμού, β) στις τεχνολογίες χωρικής τοποθέτησης των εικονικών πηγών ήχου γ) στις τεχνολογίες ηχητικής διάδρασης. Επιπροσθέτως, η αρχιτεκτονική του μοντέλου μείξης επαυξημένων περιβαλλόντων ήχου παραμένει στατική χωρίς να έχει παρουσιαστεί κάποια σημαντική εξέλιξη στην αρχή λειτουργίας της. Λαμβάνοντας υπόψιν τα παραπάνω, η παρούσα διατριβή στόχευσε στην ανάπτυξη του πυρήνα της επαυξημένης πραγματικότητας ήχου όπου είναι η επαυξημένη ηχητική μείξη εστιάζοντας στην αναβάθμιση του ρόλου της και την παρουσίαση της επίδρασης που έχει στον ακροατή.

Παραδείγματα εφαρμογών επαυξημένης πραγματικότητας συναντώνται στον τομέα της επικοινωνίας με εφαρμοσμένη χρήση σε συστήματα τηλεπαρουσίας εισάγοντας στον ομιλητή το επαυξημένο περιβάλλον του συνομιλητή [24, 47]. Εκτός από την τηλεπαρουσία, υπάρχουν εφαρμογές βελτίωσης της κατανόησης με τη χρήση τεχνολογίας τύπου ARA [7, 105]. Το εργασιακό περιβάλλον αποτελεί ένα ακόμα πεδίο χρήσης εφαρμογών τύπου ARA οι οποίες στοχεύουν κυρίως στην ενημέρωση του ακροατή με πληροφορίες για την εργασία του όπως για παράδειγμα η ηχητική παρουσίαση μηνυμάτων ή ειδοποιήσεων [137, 113, 114]. Η κατηγορία των αυτόματων οδηγών τύπου ARA αποτελεί παράδειγμα χρήσης της ARA τομείς όπως είναι η τέχνη και ο πολιτισμός ενσωματώνοντας τέτοιους οδηγούς σε εκθέσεις [9, 59, 49], ή ο τομέας του τουρισμού [14]. Άλλη μια κατηγορία αποτελεί η πλοήγηση στο δρόμο για τη βελτίωση της μετακίνησης των χρηστών [132, 3, 78]. Στον τομέα των εφαρμογών σοβαρού σκοπού, εντάσσονται εφαρμογές ARA οι οποίες βοηθούν την πλοήγηση ατόμων με δυσκολία

όρασης [62, 13, 146], καθώς και οι εφαρμογές εκπαίδευσης [54, 67]. Ο τομέας των ηχητικών παιχνιδιών τύπου ARA, αποτελεί μια μεγάλη κατηγορία εφαρμογών επαυξημένης πραγματικότητας ήχου τα οποία αποτελούν ηλεκτρονικά παιχνίδια με έμφαση στη διάδραση του παίκτη με το σύστημα μέσα από την ηχητική παρουσίαση του σεναρίου του παιχνιδιού. Η διάδραση του παίκτη με το εκάστοτε σύστημα για την επίτευξη ενός σκοπού, αποτελεί ένα απαιτητικό περιβάλλον για τον δημιουργών τέτοιων παιχνιδιών καθώς και για τον παίκτη [74, 29, 20, 111, 94]. Η δυσκολία έγκειται στον τρόπο ηχητικής απόδοσης της εικονικής πληροφορίας στον τρισδιάστατο ηχητικό χώρο, οι έντονες διακυμάνσεις που συναντώνται στο πραγματικό περιβάλλον, στον ακουστικό εξοπλισμό, στους τρόπους διάδρασης, και στη μείξη με το πραγματικό περιβάλλον. Στην παρούσα διατριβή, τα ηχητικά παιχνίδια επαυξημένης πραγματικότητας ήχου, αποτελούν την εφαρμογή στόχου για την οποία γίνεται εστίαση στην επαυξημένη μείξη ήχου παρουσιάζοντας την επίδραση που μπορεί να επιφέρει στο αποτέλεσμα της διάδρασης του παίκτη μέσα σε ένα τέτοιο παιχνίδι. Η πρωτότυπη συμβολή της παρούσας διατριβής συνοψίζεται στα ακόλουθα:

- Νέο μοντέλο ARA μείξης
- Εισαγωγή προσαρμοστικότητας στη λειτουργία του ARA μείκτη, μέσα από λειτουργία εξόρυξης δεδομένων ακουστικής αντίληψης
- Αξιολόγηση εφαρμοσμένης χρήσης σε τυπική εφαρμογή ARA μέσα από ηχητικό παιχνίδι τύπου ARA
- Θετική η επίδραση του νέου μοντέλου προσαρμοστικής μείξης ως προς την ακουστική αντίληψη και τη διάδραση του παίκτη με το σύστημα, μέσα από τη βελτίωση του ποσοστού επιτυχίας των παικτών σε σύγκριση με το παραδοσιακό μοντέλο ARA μείξης.

Η τεχνολογία μείξης επαυξημένων περιβαλλόντων ήχου, καλείται να δώσει λύση στον τρόπο μείξης του πραγματικού περιβάλλοντος με το εικονικό. Η στατικότητα της μείξης, δημιουργεί αρκετούς περιορισμούς πλέον ως προς την ευρύτερη χρήση της. Η δυναμική μεταβολή του πραγματικού περιβάλλοντος με συνεχείς εναλλαγές στην ενέργεια, και η έλλειψη προσαρμογής του επαυξημένου περιβάλλοντος σε αυτές, δημιουργούν συνθήκες προβληματικής ακουστικής αντίληψης. Η προσαρμοστικότητα φαίνεται ότι είναι η λύση στο παραπάνω πρόβλημα. Με την παρούσα εργασία, ερευνήθηκε για πρώτη φορά η εισαγωγή της προσαρμοστικότητας στον τομέα των επαυξημένων περιβαλλόντων ήχου με σκοπό τη δυναμική εξισορρόπηση των συνισταμένων περιβαλλόντων. Για να γίνει εφικτή η έρευνα της προσαρμοστικότητας στην επαυξημένη μείξη ήχου, έγινε η ανάπτυξη ενός μοντέλου προσαρμοστικής μείξης για περιβάλλοντα επαυξημένης πραγματικότητας ήχου λαμβάνοντας υπόψιν τα πιο πρόσφατα δεδομένα της ARA τεχνολογίας σχετικά με τη μέθοδο αναπαραγωγής του ψευδοακουστικού περιβάλλοντος, την χωρική τοποθέτηση του εικονικού περιβάλλοντος και των μεθόδων διάδρασης.

Λαμβάνοντας υπόψιν το χαρακτηριστικό της αμφιωτικότητας της τεχνολογίας τύπου ARA, έγινε έρευνα πάνω στην μέθοδο υλοποίησης του μοντέλου προσαρμοστικής μείξης. Ζητούμενο ήταν η αναζήτηση του χαρακτηριστικού που θα εξορυχθεί από το κάθε περιβάλλον και

σύμφωνα με το οποίο θα πραγματοποιηθεί η προσαρμογή του εικονικού περιβάλλοντος στο πραγματικό. Το χαρακτηριστικό που επιλέχθηκε, ήταν η ακουστότητα του εκάστοτε περιβάλλοντος. Για τον υπολογισμό της ακουστότητας, κατά το παρελθόν έχουν υπάρξει ποικίλες υλοποιήσεις και πρότυπα, ωστόσο στην περίπτωση της επαυξημένης μείξης λόγω του αμφιωτικού χαρακτηριστικού της, επιλέχθηκε η υλοποίηση ενός μοντέλου πρόβλεψης αμφιωτικής ακουστότητας με το οποίο δημιουργήθηκαν κανόνες μείξης και συσχέτισης ενέργειας συχνοτήτων ως προς της αμφιωτική ακουστότητά τους. Σημαντική πρόκληση για το μοντέλο ήταν η λειτουργία του σε πραγματικό χρόνο και η δοκιμή του από υποκείμενα.

Λόγω της ανάγκης για λειτουργία του μοντέλου σε πραγματικό χρόνο, σχεδιάστηκε και υλοποιήθηκε τεχνική πρόβλεψης ακουστότητας, με χρήση προϋπολογισμένου πίνακα συσχέτισης ακουστότητας (*Phons*), συχνότητας (*Hz*) και ακουστικής ενέργειας (*dB SPL*). Για την υλοποίηση του πίνακα, επιλέχθηκαν τόνοι συγκεκριμένης συχνότητας οι οποίοι αποτελούν τις κεντρικές συχνότητες των επιλεγμένων συχνοτικών ζωνών, και έγινε εκτίμηση της ακουστότητας τους σύμφωνα με το υφιστάμενο μοντέλο του Glasberg-Moore. Ο πίνακας λειτούργησε ως ένα σύνολο κανόνων για το μοντέλο πρόβλεψης ακουστότητας που υλοποιήθηκε για το προσαρμοστικό μοντέλο μείξης επαυξημένης πραγματικότητας ήχου.

Η αξιολόγηση του μοντέλου πραγματοποιήθηκε σε τρεις φάσεις. Η πρώτη φάση περιέχει τις μετρήσεις απόδοσης του μοντέλου σε πειραματικά σενάρια χρήσης. Σύμφωνα με τις μετρήσεις, το μοντέλο φαίνεται να έχει τάση ισοστάθμισης της ακουστότητας μεταξύ του εικονικού περιβάλλοντος και του πραγματικού σε υψηλό επίπεδο. Ωστόσο, εκτός από τις μετρήσεις ακουστότητας, διενεργήθηκαν και μετρήσεις συσχέτισης του συχνοτικού φάσματος των εκδοχών μείξης και ανάμεσά τους και της προσαρμοστικής. Στην περίπτωση της προσαρμοστικής μείξης, παρατηρήθηκε ότι η προσαρμογή της ακουστότητας μεταξύ των περιβαλλόντων δεν σημαίνει και συσχέτιση του συχνοτικού φάσματος μεταξύ των ηχητικών περιβαλλόντων.

Στην δεύτερη φάση της αξιολόγησης, πραγματοποιήθηκε υποκειμενική αξιολόγηση του μοντέλου σε πραγματικό χρόνο μέσα σε εργαστηριακό επίπεδο. Για την αξιολόγηση χρησιμοποιήθηκε η μεθοδολογία τέστ τύπου MUSHRA χωρίς προηγούμενη εκπαίδευση των υποκειμένων σε καμιά φάση της αξιολόγησης. Τα αποτελέσματα, δίνουν προβάδισμα ως προς την ποιότητα της μείξης στην προσαρμοστική μείξη έναντι των υπόλοιπων εκδοχών προσδίδοντας ιδιαίτερη σημασία στο μοντέλο προσαρμοστικής μείξης.

Στην τρίτη φάση της αξιολόγησης, το μοντέλο προσαρμοστικής μείξης εισάγεται μέσα σε ένα ηχητικό παιχνίδι επαυξημένης πραγματικότητας ήχου με τα υποκείμενα να καλούνται να βάλουν σε έναν κινούμενο ηχητικό στόχο. Τα υποκείμενα δοκιμάστηκαν σε δύο εκδοχές παιχνιδιού. Η μια εκδοχή περιλαμβάνει την παραδοσιακή επαυξημένη μείξη όπως υπάρχει μέχρι στιγμής στην βιβλιογραφία και η επόμενη εκδοχή περιλαμβάνει την προσαρμοστική μείξη ήχου για επαυξημένα περιβάλλοντα ήχου. Τα αποτελέσματα μεταξύ των δύο εκδοχών παιχνιδιού αποτελούν σημαντική ένδειξη για την επίδραση της προσαρμοστικής μείξης στη διάδραση του υποκειμένου μέσα σε επαυξημένα περιβάλλοντα ήχου.

Σύμφωνα με την αξιολόγηση, το μοντέλο προσαρμοστικής μείξης, αποτελεί σημαντικό παράγοντα στα περιβάλλοντα τύπου ARA με δυνατότητα σημαντικής επιρροής στις αποφάσεις του θεατή σε περίπτωση ηχητικών σεναρίων όπως είναι ένα ηχητικό παιχνίδι επαυξημένης

πραγματικότητας ήχου. Ωστόσο σε αυτή τη νέα προσέγγιση της ARA μείξης φαίνεται πως δημιουργείται ένα νέο πεδίο στην επαυξημένη πραγματικότητα ήχου το οποίο στοχεύει στην βελτίωση της μείξης μεταξύ των περιβαλλόντων και χρήζει περαιτέρω έρευνας. Στο προτεινόμενο μοντέλο που υλοποιήθηκε, περιβάλλον αναφοράς αποτέλεσε το πραγματικό περιβάλλον πάνω στο οποίο προσαρμόζεται το εικονικό. Ωστόσο θα μπορούσε να υπάρξει η δυνατότητα αυτή η ισορροπία να αλλάζει ανάλογα με την σημαντικότητα του εκάστοτε περιβάλλοντος. Πιθανότατα στο μέλλον η ανάπτυξη ενός νευρωνικού δικτύου με δυνατότητα αναγνώρισης συμβάντων στο εκάστοτε περιβάλλον να μεταβάλλει αυτή την ισορροπία.

Μέσα από την υποκειμενική αξιολόγηση της προσαρμοστικής μείξης, φαίνεται πως υπάρχει ένα συνεχές πάνω στο οποίο τοποθετείται η προσαρμοστική μείξη. Στα δύο άκρα του συνεχές τοποθετείται η αισθητική της επαυξημένης μείξης και η βελτίωση της διάδρασης. Στην παρούσα διατριβή, φάνηκε πως η μείξη με βάση το στοιχείο της ακουστότητας, φαίνεται να λειτουργεί ξεκάθαρα στην περίπτωση της διάδρασης και της λήψης αποφάσεων στο ηχητικό παιχνίδι τύπου ARA. Ωστόσο, σε θέματα αισθητικής όπως είναι ένα σενάριο ακρόασης μουσικής χωρίς την αποκοπή του χρήστη από το πραγματικό περιβάλλον, ίσως θα πρέπει να διερευνηθεί η μέτρηση του φάσματος και η συσχέτιση μεταξύ των περιβαλλόντων. Υπάρχει μια πρώτη επισήμανση με βάση τα αποτελέσματα αντικειμενικής και υποκειμενικής αξιολόγησης, ότι η προσαρμοστική μείξη παρουσιάζει αισθητικά καλύτερο αποτέλεσμα αυξάνοντας την συσχέτιση μεταξύ εικονικού περιβάλλοντος και πραγματικού ενώ για την βελτίωση της διάδρασης ο παράγοντας που επηρεάζει το υποκείμενο είναι η ακουστότητα. Επιπλέον, η ισότητα μεταξύ των περιβαλλόντων ή η διαφοροποίησή τους ως επιθυμητή μείξη για τον ακροατή, φαίνεται να είναι ένα πολύπλοκο ερώτημα το οποίο χρήζει επιπλέον ερευνητικής εργασίας.

# Βιβλιογραφία

- [1] ISO 532: 2017. Acoustics—methods for calculating loudness—part 2: Moore-glasberg method, 2017.
- [2] Robert Albrecht, Tapio Lokki και Lauri Savioja. A mobile augmented reality audio system with binaural microphones. Στο *Proceedings of Interacting with Sound Workshop: Exploring Context-Aware, Local and Social Audio Applications*, σελίδες 7–11, 2011.
- [3] Robert Albrecht, Riitta Väänänen και Tapio Lokki. Guided by music: pedestrian and cyclist navigation with route and beacon guidance. *Personal and Ubiquitous Computing*, 20(1):121–145, 2016.
- [4] Francesco Aletta, Timothy Van Renterghem και Dick Botteldooren. Influence of personal factors on sound perception and overall experience in urban green areas. a case study of a cycling path highly exposed to road traffic noise. *International journal of environmental research and public health*, 15(6):1118, 2018.
- [5] Amit Barde, William S Helton, Gun Lee και Mark Billingham. Binaural spatialization over a bone conduction headset: minimum discernable angular difference. Στο *Audio Engineering Society Convention 140*. Audio Engineering Society, 2016.
- [6] Amit Barde, Gun Lee, William S Helton και Mark Billingham. Binaural spatialisation over a bone conduction headset: Elevation perception. International Community on Auditory Display, 2016.
- [7] Amit Barde, Matt Ward, William S Helton, Mark Billingham και Gun Lee. Attention redirection using binaurally spatialised cues delivered over a bone conduction headset. Στο *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, τόμος 60, σελίδες 1534–1538. SAGE Publications Sage CA: Los Angeles, CA, 2016.
- [8] Søren Bech και Nick Zacharov. *Perceptual audio evaluation: theory, method and application*. Wiley Online Library, 2006.
- [9] Benjamin B Bederson. Audio augmented reality: a prototype automated tour guide. Στο *Conference companion on Human factors in computing systems*, σελίδες 210–211, 1995.

- [10] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [11] Jens Blauert. *Communication acoustics*, τόμος 2. Springer, 2005.
- [12] Jens Blauert και P Laws. Group delay distortions in electroacoustical systems. *The Journal of the Acoustical Society of America*, 63(5):1478–1483, 1978.
- [13] Jeffrey R Blum, Mathieu Bouchard και Jeremy R Cooperstock. What’s around me? spatialized audio augmented reality for blind users with a smartphone. Στο *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, σελίδες 49–62. Springer, 2011.
- [14] Costas Boletsis και Dimitra Chasanidou. Smart tourism in cities: Exploring urban destinations with audio augmented reality. Στο *Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference*, σελίδες 515–521, 2018.
- [15] J Borwick. Loudspeaker and headphone handbook, ed. *Reed Educational and Professional Publishing Ltd*, 2001.
- [16] Bose. Bose ar headset, 2020.
- [17] Donald E Broadbent και CM Harris. Handbook of noise control. 1979.
- [18] Douglas S Brungart, Alex J Kordik, Cynthia S Eades και Brian D Simpson. The effect of microphone placement on localization accuracy with electronic pass-through earplugs. Στο *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No. 03TH8684)*, σελίδες 149–152. IEEE, 2003.
- [19] Thomas Chatzidimitris, Damianos Gavalas και Vlasios Kasapakis. Pacmap: transferring pacman to the physical realm. Στο *International Internet of Things Summit*, σελίδες 139–144. Springer, 2014.
- [20] Thomas Chatzidimitris, Damianos Gavalas και Despina Michael. Soundpacman: Audio augmented reality in location-based games. Στο *2016 18th Mediterranean Electrotechnical Conference (MELECON)*, σελίδες 1–6. IEEE, 2016.
- [21] Miguel A Chavez και Camille Huin. Sigmastudio. a user friendly, intuitive and expandable, graphical development environment for audio/dsp applications. Στο *Audio Engineering Society Convention 120*. Audio Engineering Society, 2006.
- [22] David Y Chung, Keith Mason, R Patrick Gannon και Glenn N Willson. The ear effect as a function of age and hearing loss. *The Journal of the Acoustical Society of America*, 73(4):1277–1282, 1983.
- [23] Tuck Siong Chung, Roland T Rust και Michel Wedel. My mobile music: An adaptive personalization system for digital audio players. *Marketing Science*, 28(1):52–68, 2009.

- [24] Michael Cohen, Shigeaki Aoki και Nobuo Koizumi. Augmented audio reality: Telepresence/vr hybrid acoustic environments. Στο *Proceedings of 1993 2nd IEEE International Workshop on Robot and Human Communication*, σελίδες 361–364. IEEE, 1993.
- [25] Etienne Corteel. Synthesis of directional sources using wave field synthesis, possibilities, and limitations. *EURASIP Journal on Advances in Signal Processing*, 2007(1):090509, 2007.
- [26] Daniela D’Auria, Dario Di Mauro, Davide Maria Calandra και Francesco Cutugno. A 3d audio augmented reality system for a cultural heritage management and fruition. *Journal of Digital Information Management*, 13(4), 2015.
- [27] Hallowell Davis και Sol Richard Silverman. *Hearing and deafness*. Holt, Rinehart & Winston of Canada Ltd, 1970.
- [28] Simulink Documentation. Simulation and model-based design, 2017.
- [29] Thomas M Drewes, Elizabeth Mynatt, Maribeth Gandy και others. Sleuth: An audio experience. Στο *Proceedings of The International Conference on Auditory Display*, 2000.
- [30] Gerhard Eckel. Immersive audio-augmented environments: the listen project. Στο *Proceedings fifth international conference on information visualisation*, σελίδες 571–573. IEEE, 2001.
- [31] Mohammed Fakhour, Ahmed Azough, Fatima Zahra Kaghat και Mohammed Meknassi. A cultural scavenger hunt serious game based on audio augmented reality. Στο *International Conference on Advanced Intelligent Systems for Sustainable Development*, σελίδες 1–8. Springer, 2019.
- [32] Hugo Fastl. The psychoacoustics of sound-quality evaluation. *Acta Acustica united with Acustica*, 83(5):754–764, 1997.
- [33] Harvey Fletcher και Wilden A Munson. Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, 12(4):377–430, 1933.
- [34] Harvey Fletcher και Wilden A Munson. Relation between loudness and masking. *The Journal of the Acoustical Society of America*, 9(1):1–10, 1937.
- [35] Bill Gardner, Keith Martin και others. Hrft measurements of a kemar dummy-head microphone, 1994.
- [36] William G Gardner. 3d audio and acoustic environment modeling. *Wave Arts, Inc*, 99, 1999.
- [37] William Grant Gardner. *Transaural 3-D audio*. Citeseer, 1995.

- [38] SA Genesis. Loudness toolbox. for matlab, 2009.
- [39] Christian Giguère και Sharon M Abel. Sound localization: Effects of reverberation time, speaker array, stimulus frequency, and stimulus rise/decay. *The Journal of the Acoustical Society of America*, 94(2):769–776, 1993.
- [40] Stephen B Gilbert. Perceived realism of virtual environments depends on authenticity. *Presence: Teleoperators and Virtual Environments*, 24(4):322–324, 2016.
- [41] Brian R Glasberg και Brian CJ Moore. Derivation of auditory filter shapes from notched-noise data. *Hearing research*, 47(1-2):103–138, 1990.
- [42] Brian R Glasberg και Brian CJ Moore. A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5):331–342, 2002.
- [43] Rishabh Gupta, Rishabh Ranjan, Jianjun He και Woon Seng Gan. Parametric hear through equalization for augmented reality audio. Στο *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, σελίδες 1587–1591. IEEE, 2019.
- [44] Rishabh Gupta, Rishabh Ranjan, Jianjun He και Gan Woon-Seng. On the use of closed-back headphones for active hear-through equalization in augmented reality applications. Στο *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2018.
- [45] Dorte Hammersho/i και Henrik Mo/lker. Sound transmission to and within the human ear canal. *The Journal of the Acoustical Society of America*, 100(1):408–427, 1996.
- [46] Aki Härmä, Julia Jakka, Miikka Tikander, Matti Karjalainen, Tapio Lokki, Jarmo Hiipakka και Gaëtan Lorho. Augmented reality audio for mobile and wearable appliances. *Journal of the Audio Engineering Society*, 52(6):618–639, 2004.
- [47] Aki Harma, Julia Jakka, Miikka Tikander, Matti Karjalainen, Tapio Lokki και Heli Nironen. Techniques and applications of wearable augmented reality audio. Στο *Audio Engineering Society Convention 114*. Audio Engineering Society, 2003.
- [48] Karen S Helfer. Gender, age, and hearing. Στο *Seminars in Hearing*, τόμος 22, σελίδες 271–286. Copyright© 2001 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ... , 2001.
- [49] Florian Heller και Jan Borchers. Corona: Audio augmented reality in historic sites. Στο *MobileHCI 2011 Workshop on Mobile Augmented Reality: Design Issues and Opportunities*, σελίδες 51–54, 2011.

- [50] Florian Heller, Jayan Jevanesan, Pascal Dietrich και Jan Borchers. Where are we? evaluating the current rendering fidelity of mobile audio augmented reality systems. Στο *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, σελίδες 278–282, 2016.
- [51] Kenneth R Henry. Age-related auditory loss and genetics: an electrocochleographic comparison of six inbred strains of mice. *Journal of Gerontology*, 37(3):275–282, 1982.
- [52] Robert Höldrich και Katharina Vogt. Augmented audification. Georgia Institute of Technology, 2015.
- [53] A James Hudspeth. How the ear’s works work. *Nature*, 341(6241):397–404, 1989.
- [54] Dean P Inman, Ken Loge και Aaron Cram. Teaching orientation and mobility skills to blind children using computer generated 3d sound environments. Georgia Institute of Technology, 2000.
- [55] B ISO. 226: 2003: Acoustics—normal equal-loudness-level contours. *International Organization for Standardization*, 63, 2003.
- [56] Finn Jacobsen, Torben Poulsen, Jens Holger Rindel, Anders Christian Gade και Mogens Ohlrich. Fundamentals of acoustics and noise control, 2011.
- [57] Giordano Jacuzzi. ” augmented audio”: An overview of the unique tools and features required for creating ar audio experiences. Στο *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2018.
- [58] Rämö Jussi. Evaluation of an augmented reality audio headset and mixer. *Master’s Thesis, Helsinki University of Technology*, 2009.
- [59] Fatima Zahra Kaghat, Ahmed Azough και Mohammed Fakhour. Sarim: a gesture-based sound augmented reality interface for visiting museums. Στο *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, σελίδες 1–9. IEEE, 2018.
- [60] Matti Karjalainen, Tapio Lokki, Heli Nironen, Aki Harma, Lauri Savioja και Sampo Vesa. Application scenarios of wearable and mobile augmented reality audio. Στο *Audio Engineering Society Convention 116*. Audio Engineering Society, 2004.
- [61] Matti Karjalainen, Ville Riikonen και Miikka Tikander. An augmented reality audio mixer and equalizer. Στο *Audio Engineering Society Convention 124*. Audio Engineering Society, 2008.
- [62] Brian FG Katz, Slim Kammoun, Gaëtan Parseihian, Olivier Gutierrez, Adrien Brihlhault, Malika Auvray, Philippe Truillet, Michel Denis, Simon Thorpe και Christophe

- Jouffrais. Navig: augmented reality guidance system for the visually impaired. *Virtual Reality*, 16(4):253–269, 2012.
- [63] Shin Kim, Kun pyo Lee και Tek Jin Nam. Sonic-badminton: audio-augmented badminton game for blind people. Στο *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, σελίδες 1922–1929, 2016.
- [64] Katie King και Dafydd Stephens. Auditory and psychological factors in ‘auditory disability with normal hearing’. *Scandinavian audiology*, 21(2):109–114, 1992.
- [65] Kazuhiro Kondo, Naoya Anazawa και Yosuke Kobayashi. Characteristics comparison of two audio output devices for augmented audio reality. Στο *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, σελίδες 1–6. IEEE, 2013.
- [66] Gregory Kramer. *Auditory display: sonification, audification and auditory interfaces*. Addison-Wesley Longman Publishing Co., Inc., 2000.
- [67] Myron W Krueger και Deborah Gilden. Knowwhere: an audio/spatial interface for blind people. Georgia Institute of Technology, 1997.
- [68] Josh Lando, Alex Brandmeyer, Phil Brown, Alan Seefeldt και Andy Jaspar. A perceptually-motivated headphone transparency algorithm. Στο *Audio Engineering Society Convention 147*, 2019.
- [69] Eric Lee, Henning Kiel, Saskia Dedenbach, Ingo Grüll, Thorsten Karrer, Marius Wolf και Jan Borchers. isymphony: an adaptive interactive orchestral conducting system for digital audio and video streams. Στο *CHI’06 Extended Abstracts on Human Factors in Computing Systems*, σελίδες 259–262, 2006.
- [70] Robert W Lindeman, Haruo Noma και Paulo Goncalvesde Barros. Hear-through and mic-through augmented reality: Using bone conduction to display spatialized audio. Στο *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, σελίδες 173–176. IEEE, 2007.
- [71] Robert W Lindeman, Haruo Noma και Paulo Gonçalvesde Barros. An empirical study of hear-through augmented reality: Using bone conduction to deliver spatialized audio. Στο *2008 IEEE Virtual Reality Conference*, σελίδες 35–42. IEEE, 2008.
- [72] Peter H Lindsay και Donald A Norman. *Human information processing: An introduction to psychology*. Academic press, 2013.
- [73] Juho Liski, Riitta Väänänen, Sampo Vesa και Vesa Välimäki. Adaptive equalization of acoustic transparency in an augmented-reality headset. Στο *Audio Engineering*

- Society Conference: 2016 AES International Conference on Headphone Technology*. Audio Engineering Society, 2016.
- [74] Kent Lyons, Maribeth Gandy και Thad Starner. Guided by voices: An audio augmented reality system. Georgia Institute of Technology, 2000.
  - [75] Justin A MacDonald, Paula P Henry και Tomasz R Letowski. Spatial audio through a bone conduction interface: Audición espacial a través de una interfase de conducción ósea. *International journal of audiology*, 45(10):595–599, 2006.
  - [76] Ewan A Macpherson και John C Middlebrooks. Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, 111(5):2219–2236, 2002.
  - [77] David G Malham και Anthony Myatt. 3-d sound spatialization using ambisonic techniques. *Computer music journal*, 19(4):58–70, 1995.
  - [78] Jessica Mantell, Jan Rod, Yuichiro Kage, Fabien Delmotte και Johnson Leu. Navinko: Audio augmented reality-enabled social navigation for city cyclists. Στο *Programme, Workshop Pervasive 2010 Proc*, 2010.
  - [79] MATLAB. *version 9.3.0 (R2017b)*. The MathWorks Inc., Natick, Massachusetts, 2017.
  - [80] James H McClellan και Thomas W Parks. A personal history of the parks-mcclellan algorithm. *IEEE signal processing magazine*, 22(2):82–86, 2005.
  - [81] Mark McGill, Stephen Brewster, David McGookin και Graham Wilson. Acoustic transparency and the changing soundscape of auditory mixed reality. 2020.
  - [82] RF McLean, SH Alsop και JS Fleming. Nyquist—overcoming the limitations. *Journal of sound and vibration*, 280(1-2):1–20, 2005.
  - [83] Paul Milgram και Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.
  - [84] Allen William Mills. On the minimum audible angle. *The Journal of the Acoustical Society of America*, 30(4):237–246, 1958.
  - [85] David Miranda και Kiju Lee. Music blocks: Audio-augmented block games for play-based cognitive assessment. Στο *2018 IEEE Games, Entertainment, Media Conference (GEM)*, σελίδες 375–381. IEEE, 2018.
  - [86] Henrik Møller. Fundamentals of binaural technology. *Applied acoustics*, 36(3-4):171–218, 1992.
  - [87] Brian CJ Moore. *An introduction to the psychology of hearing*. Brill, 2012.

- [88] Brian CJ Moore και Brian R Glasberg. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The journal of the acoustical society of America*, 74(3):750–753, 1983.
- [89] Brian CJ Moore και Brian R Glasberg. Modeling binaural loudness. *The Journal of the Acoustical Society of America*, 121(3):1604–1612, 2007.
- [90] Brian CJ Moore, Brian R Glasberg και Thomas Baer. A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4):224–240, 1997.
- [91] Masayuki Morimoto και Hitoshi Aokata. Localization cues of sound sources in the upper hemisphere. *Journal of the Acoustical Society of Japan (E)*, 5(3):165–173, 1984.
- [92] Nikolaos Moustakas, Andreas Floros και Nikolaos Kanellopoulos. Eidola: An interactive augmented reality audio-game prototype. Στο *Audio Engineering Society Convention 127*. Audio Engineering Society, 2009.
- [93] Nikolaos Moustakas, Emmanouel Rovithis, Konstantinos Vogklis και Andreas Floros. Adaptive audio mixing for enhancing immersion in augmented reality audio games. Στο *Companion Publication of the 2020 International Conference on Multimodal Interaction*, ICMI '20 Companion, σελίδα 220–227, New York, NY, USA, 2020. Association for Computing Machinery.
- [94] Nikos Moustakas, Andreas Floros και Nicolas Grigoriou. Interactive audio realities: An augmented/mixed reality audio game prototype. Στο *Audio Engineering Society Convention 130*. Audio Engineering Society, 2011.
- [95] Nikos Moustakas, Andreas Floros και Bill Kapralos. An augmented reality audio live network for live electroacoustic music concerts. Στο *Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society, 2016.
- [96] Nikos Moustakas, Andreas Floros, Emmanouel Rovithis και Konstantinos Vogklis. Augmented audio-only games: A new generation of immersive acoustic environments through advanced mixing. Στο *Audio Engineering Society Convention 146*. Audio Engineering Society, 2019.
- [97] Elizabeth D Mynatt, Maribeth Back, Roy Want και Ron Frederick. Audio aura: Light-weight audio augmented reality. Στο *Proceedings of the 10th annual ACM symposium on User interface software and technology*, σελίδες 211–212, 1997.
- [98] Sean Olive, Todd Welti και Omid Khonsaripour. The preferred low frequency response of in-ear headphones. Στο *Audio Engineering Society Conference: 2016 AES International Conference on Headphone Technology*. Audio Engineering Society, 2016.

- [99] Helen Papagiannis. *Augmented human: How technology is shaping the new reality*. ” O’Reilly Media, Inc.”, 2017.
- [100] Roy D Patterson. Auditory filter shape. *The Journal of the Acoustical Society of America*, 55(4):802–809, 1974.
- [101] Ignacio Pecino και Ricardo Climent. *SonicMaps: connecting the ritual of the concert hall with a locative audio urban experience*. Ann Arbor, MI: Michigan Publishing, University of Michigan Library, 2013.
- [102] Jussi Rämö και Vesa Välimäki. Digital augmented reality audio headset. *Journal of Electrical and Computer Engineering*, 2012, 2012.
- [103] Jussi Rämö και Vesa Välimäki. Signal processing framework for virtual headphone listening tests in a noisy environment. Στο *Audio Engineering Society Convention 132*. Audio Engineering Society, 2012.
- [104] Jussi Rämö και Vesa Välimäki. An allpass hear-through headset. Στο *2014 22nd European Signal Processing Conference (EUSIPCO)*, σελίδες 1123–1127. IEEE, 2014.
- [105] Jussi Rämö, Vesa Välimäki και Miikka Tikander. Live sound equalization and attenuation with a headset. Στο *Audio Engineering Society Conference: 51st International Conference: Loudspeakers and Headphones*. Audio Engineering Society, 2013.
- [106] Rishabh Ranjan και Woon Seng Gan. Natural listening over headphones in augmented reality using adaptive filtering techniques. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(11):1988–2002, 2015.
- [107] Rishabh Ranjan και Woon Seng Gan. Adaptive equalization of natural augmented reality headset using non-stationary virtual signals. Στο *Audio Engineering Society Conference: 2016 AES International Conference on Headphone Technology*, 2016.
- [108] ITUR Recommendation. 1534-1, “. *Method for the subjective assessment of intermediate quality level of coding systems*, 2003, 2001.
- [109] Joshua D Reiss. Intelligent systems for mixing multichannel audio. Στο *2011 17th International Conference on Digital Signal Processing (DSP)*, σελίδες 1–6. IEEE, 2011.
- [110] Emmanouel Rovithis, Andreas Floros, Nikos Moustakas, Konstantinos Vogklis και Lily Kotsira. Bridging audio and augmented reality towards a new generation of serious audio-only games. *Electronic Journal of e-Learning*, 17(2):144–156, 2019.
- [111] Emmanouel Rovithis, Nikolaos Moustakas, Andreas Floros και Kostas Vogklis. Audio legends: Investigating sonic interaction in an augmented reality audio game. *Multimodal Technologies and Interaction*, 3(4):73, 2019.

- [112] Gregory J Sandell. Auditory display: Sonification, audification, and auditory interfaces, 1996.
- [113] Nitin Sawhney και Chris Schmandt. Design of spatialized audio in nomadic environments. Georgia Institute of Technology, 1997.
- [114] Nitin Sawhney και Chris Schmandt. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM transactions on Computer-Human interaction (TOCHI)*, 7(3):353–383, 2000.
- [115] B Scharf και D Fishken. Binaural summation of loudness: Reconsidered. *Journal of Experimental Psychology*, 86(3):374, 1970.
- [116] Roman Schlieper, Song Li, Stephan Preihs και Jürgen Peissig. The relationship between the acoustic impedance of headphones and the occlusion effect. Στο *Audio Engineering Society Conference: 2019 AES INTERNATIONAL CONFERENCE ON HEADPHONE TECHNOLOGY*, 2019.
- [117] David Shier, Jackie Butler και Ricki Lewis. *Hole's essentials of human anatomy & physiology*. McGraw-Hill Education, 2018.
- [118] Russell D Shilling, Tomasz Letowski και Russell Storms. Spatial auditory displays for use within attack rotary wing aircraft. Georgia Institute of Technology, 2000.
- [119] Julius O Smith και Jonathan S Abel. Bark and erb bilinear transforms. *IEEE Transactions on speech and Audio Processing*, 7(6):697–708, 1999.
- [120] Stefan Stenfelt και Bo Håkansson. Air versus bone conduction: an equal loudness investigation. *Hearing research*, 167(1-2):1–12, 2002.
- [121] SS Stevens. Sensory power functions and neural events. Στο *Principles of receptor physiology*, σελίδες 226–242. Springer, 1971.
- [122] Stanley S Stevens. On the psychophysical law. *Psychological review*, 64(3):153, 1957.
- [123] Stanley S Stevens. Perceived level of noise by mark vii and decibels (e). *The Journal of the Acoustical Society of America*, 51(2B):575–601, 1972.
- [124] Gabriel M Swords, Lydia T Nguyen, Raksha A Mudar και Daniel A Llano. Auditory system dysfunction in alzheimer disease and its prodromal states: A review. *Ageing research reviews*, 44:49–59, 2018.
- [125] PC Thomas και WM David. Augmented reality: An application of heads-up display technology to manual manufacturing processes. Στο *Hawaii International Conference on System Sciences*, σελίδες 659–669, 1992.

- [126] Miikka Tikander. Modeling the attenuation of a loosely-fit insert headphone for augmented reality audio. Στο *Audio Engineering Society Conference: 30th International Conference: Intelligent Audio Environments*. Audio Engineering Society, 2007.
- [127] Miikka Tikander. Usability issues in listening to natural sounds with an augmented reality audio headset. *Journal of the Audio Engineering Society*, 57(6):430–441, 2009.
- [128] Miikka Tikander, Matti Karjalainen και Ville Riikonen. An augmented reality audio headset. Στο *Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFx-08)*, Espoo, Finland, 2008.
- [129] Miikka Tikander και V Riikonen. Sound quality of an augmented reality audio headset. Στο *Proc. of the 8th Int. Conf. on Digital Audio Effects (DAFx-05)*, Madrid, Spain, 2005.
- [130] Joseph Timoney, Thomas Lysaght και Marc Schoenwiesner. Implementing loudness models in matlab. DAFX, 2004.
- [131] Olli Tuomi και Nick Zacharov. A real-time binaural loudness meter. Στο *139th meeting of the Acoustical Society of America, Atlanta, USA*, 2000.
- [132] Riitta Väänänen, Sampo Vesa και Matti Hämäläinen. Testing the user experience of an augmented reality headset and 3d audio-guided pedestrian navigation. Στο *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*. Audio Engineering Society, 2014.
- [133] Vesa Valimaki, Andreas Franck, Jussi Ramo, Hannes Gamper και Lauri Savioja. Assisted listening using a headset: Enhancing audio perception in real, augmented, and virtual environments. *IEEE Signal Processing Magazine*, 32(2):92–99, 2015.
- [134] Vincent Verfaille, Udo Zolzer και Daniel Arfib. Adaptive digital audio effects (adafx): A new class of sound transformations. *IEEE Transactions on audio, speech, and language processing*, 14(5):1817–1831, 2006.
- [135] J Virolainen. *Design and implementation of a stereo audio conferencing system*. Διδακτορική Διατριβή, Master’s thesis, Helsinki University of Technology, 2001.
- [136] Ron Wakkary, Kenneth Newby, Marek Hatala, Dale Evernden και Milena Droumeva. Interactive audio content: an approach to audio content for a dynamic museum experience through augmented audio reality and adaptive information retrieval. 2004.
- [137] Ashley Walker, Stephen Brewster, David McGookin και Adrian Ng. Diary in the sky: A spatial audio display for a mobile calendar. Στο *People and Computers XV—Interaction without Frontiers*, σελίδες 531–539. Springer, 2001.
- [138] Bruce N Walker και Jeffrey Lindsay. Navigation performance in a virtual environment with bonephones. Georgia Institute of Technology, 2005.

- [139] Bruce N Walker και Raymond M Stanley. Thresholds of audibility for bone-conduction headsets. Georgia Institute of Technology, 2005.
- [140] Dominic Ward, Sean Enderby, Cham Athwal και Joshua D Reiss. Real-time excitation based binaural loudness meters. Στο *18th Int. Conf. DAFx*, 2015.
- [141] Dominic Ward και Joshua D Reiss. Loudness algorithms for automatic mixing. Στο *Proceedings of the 2nd AES Workshop on Intelligent Music Production*, τόμος 13, 2016.
- [142] Richard M Warren. *Auditory perception: A new synthesis*, τόμος 109. Elsevier, 2013.
- [143] Francis M Wiener. Sound diffraction by rigid spheres and circular cylinders. *The Journal of the Acoustical Society of America*, 19(3):444–451, 1947.
- [144] John Willett. Design challenges for compact headphones. Στο *Audio Engineering Society Conference: UK 23rd Conference: Music Everywhere*. Audio Engineering Society, 2008.
- [145] Paul N Wilson, Nigel Foreman, Raphael Gillett και Danae Stanton. Active versus passive processing of spatial information in a computer-simulated environment. *Ecological Psychology*, 9(3):207–222, 1997.
- [146] Farrah Wong, R Nagarajan, Sazali Yaacob, Ali Chekima και NE Belkhamza. Electronic travel aids for visually impaired—a guided tour. Στο *Proceedings of Conference in Engineering in Sarawak*, σελίδες 377–382, 2000.
- [147] Andreas Zimmermann και Andreas Lorenz. Listen: a user-adaptive audio-augmented museum guide. *User Modeling and User-Adapted Interaction*, 18(5):389–416, 2008.
- [148] E Zwicker και H Fastl. A portable loudness-meter based on iso 532 b. Στο *Proc. 11th International Congress on Acoustics*, σελίδες 135–137, 1983.
- [149] Eberhard Zwicker και Hugo Fastl. *Psychoacoustics: Facts and models*, τόμος 22. Springer Science & Business Media, 2013.
- [150] Eberhard Zwicker, Hugo Fastl, Ulrich Widmann, Kenji Kurakata, Sonoko Kuwano και Seiichiro Namba. Program for calculating loudness according to din 45631 (iso 532b). *Journal of the Acoustical Society of Japan (E)*, 12(1):39–42, 1991.
- [151] Eberhard Zwicker, Georg Flottorp και Stanley Smith Stevens. Critical band width in loudness summation. *The Journal of the Acoustical Society of America*, 29(5):548–557, 1957.

## Παράρτημα Α΄

# Πίνακας Μετατροπής Αμφιωτικής Ακουστότητας - Συχνότητας - Στάθμης Ακουστικής Πίεσης

Σε αυτό το παράρτημα, παρουσιάζεται η μήτρα συσχέτισης «Αμφιωτικής Ακουστότητας» *Phons* για προκαθορισμένες συχνότητες (*Hz*) για μεταβαλλόμενη διαφορά στη στάθμη ηχητικής πίεσης *1 dB SPL*.

| <i>Phons</i> | 0.125kHz | 0.25kHz | 0.5kHz | 1kHz | 2kHz | 4kHz | 8kHz | 16kHz |
|--------------|----------|---------|--------|------|------|------|------|-------|
| 0            | -        | -       | -      | -    | -    | -    | 10   | 37    |
| 1            | 20       | 10      | -      | -    | -    | -    | 11   | 39    |
| 2            | 21       | 11      | -      | -    | -    | -    | 12   | 40    |
| 3            | 22       | 12      | -      | -    | -    | -    | 13   | 41    |
| 4            | 23       | 13      | -      | -    | -    | -    | 14   | 43    |
| 5            | 24       | 14      | -      | -    | -    | -    | 15   | 44    |
| 6            | 25       | 15      | -      | -    | -    | -    | 16   | 45    |
| 7            | 26       | 16      | -      | -    | -    | -    | 17   | 46    |
| 8            | 27       | 17      | 10     | -    | -    | -    | 18   | 47    |
| 9            | 28       | 18      | 11     | -    | -    | -    | 19   | 48    |
| 10           | 29       | 19      | 12     | 10   | -    | -    | 20   | 49    |
| 11           | 30       | 20      | 13     | 11   | -    | -    | 21   | 50    |
| 12           | 31       | 21      | 14     | 12   | -    | -    | 22   | 51    |
| 13           | 32       | 22      | 15     | 13   | -    | -    | 23   | 52    |
| 14           | 33       | 23      | 16     | 14   | 10   | -    | 24   | 53    |
| 15           | 34       | 24      | 17     | 15   | 11   | -    | 25   | 54    |
| 16           | 35       | 25      | 18     | 16   | 12   | -    | 26   | 55    |
| 17           | 36       | 26      | 19     | 17   | 13   | -    | 27   | 57    |
| 18           | 36       | 27      | 20     | 18   | 14   | -    | 28   | 58    |

Συνέχεια στην επόμενη σελίδα

Πίνακας Α'.1 – Συνέχεια από προηγούμενη σελίδα

| <i>Phons</i> | 0.125kHz | 0.25kHz | 0.5kHz | 1kHz | 2kHz | 4kHz | 8kHz | 16kHz |
|--------------|----------|---------|--------|------|------|------|------|-------|
| 19           | 37       | 28      | 21     | 19   | 15   | 10   | 29   | 59    |
| 20           | 38       | 29      | 22     | 20   | 16   | 11   | 30   | 60    |
| 21           | 39       | 30      | 23     | 21   | 17   | 12   | 31   | 61    |
| 22           | 40       | 31      | 24     | 22   | 18   | 13   | 32   | 62    |
| 23           | 41       | 32      | 25     | 23   | 19   | 14   | 33   | 62    |
| 24           | 41       | 33      | 26     | 24   | 20   | 15   | 34   | 63    |
| 25           | 42       | 34      | 27     | 25   | 21   | 16   | 35   | 64    |
| 26           | 43       | 34      | 28     | 26   | 22   | 17   | 36   | 65    |
| 27           | 44       | 35      | 29     | 27   | 23   | 18   | 37   | 66    |
| 28           | 45       | 36      | 30     | 28   | 24   | 19   | 38   | 67    |
| 29           | 46       | 37      | 31     | 29   | 25   | 20   | 39   | 68    |
| 30           | 47       | 38      | 32     | 30   | 26   | 21   | 40   | 69    |
| 31           | 48       | 39      | 33     | 31   | 27   | 22   | 41   | 69    |
| 32           | 48       | 40      | 34     | 32   | 28   | 23   | 42   | 70    |
| 33           | 49       | 41      | 35     | 33   | 29   | 24   | 42   | 71    |
| 34           | 50       | 42      | 36     | 34   | 30   | 25   | 43   | 72    |
| 35           | 51       | 43      | 37     | 35   | 31   | 26   | 44   | 73    |
| 36           | 52       | 44      | 38     | 36   | 32   | 27   | 45   | 74    |
| 37           | 53       | 45      | 39     | 36   | 33   | 28   | 47   | 75    |
| 38           | 54       | 45      | 40     | 37   | 34   | 29   | 47   | 76    |
| 39           | 54       | 46      | 41     | 38   | 35   | 30   | 48   | 77    |
| 40           | 55       | 47      | 42     | 39   | 36   | 31   | 49   | 78    |
| 41           | 56       | 48      | 43     | 40   | 37   | 32   | 50   | 79    |
| 42           | 57       | 49      | 44     | 41   | 38   | 33   | 51   | 80    |
| 43           | 58       | 50      | 45     | 42   | 39   | 34   | 52   | 80    |
| 44           | 59       | 51      | 46     | 43   | 40   | 35   | 53   | 81    |
| 45           | 60       | 52      | 47     | 44   | 41   | 36   | 54   | 82    |
| 46           | 61       | 53      | 48     | 45   | 42   | 37   | 55   | 83    |
| 47           | 62       | 54      | 49     | 46   | 43   | 38   | 56   | 84    |
| 48           | 63       | 55      | 50     | 47   | 44   | 39   | 57   | 85    |
| 49           | 64       | 56      | 51     | 48   | 45   | 40   | 58   | 86    |
| 50           | 65       | 57      | 52     | 49   | 46   | 41   | 59   | 87    |
| 51           | 66       | 58      | 53     | 50   | 47   | 42   | 60   | 88    |
| 52           | 66       | 59      | 54     | 51   | 48   | 43   | 61   | 89    |
| 53           | 67       | 60      | 55     | 52   | 49   | 44   | 62   | 90    |
| 54           | 68       | 61      | 56     | 53   | 50   | 45   | 63   | 91    |
| 55           | 69       | 62      | 57     | 54   | 51   | 46   | 64   | -     |
| 56           | 70       | 63      | 58     | 55   | 52   | 47   | 65   | -     |

Συνέχεια στην επόμενη σελίδα

Πίνακας Α'.1 – Συνέχεια από προηγούμενη σελίδα

| <i>Phons</i> | 0.125kHz | 0.25kHz | 0.5kHz | 1kHz | 2kHz | 4kHz | 8kHz | 16kHz |
|--------------|----------|---------|--------|------|------|------|------|-------|
| 57           | 71       | 64      | 59     | 57   | 53   | 48   | 66   | -     |
| 58           | 72       | 65      | 60     | 57   | 54   | 49   | 67   | -     |
| 59           | 73       | 66      | 61     | 58   | 55   | 50   | 68   | -     |
| 60           | 74       | 67      | 62     | 59   | 56   | 51   | 69   | -     |
| 61           | 75       | 68      | 63     | 60   | 57   | 52   | 70   | -     |
| 62           | 76       | 69      | 64     | 61   | 58   | 53   | 71   | -     |
| 63           | 76       | 70      | 65     | 62   | 59   | 54   | 72   | -     |
| 64           | 77       | 71      | 66     | 63   | 60   | 55   | 73   | -     |
| 65           | 78       | 72      | 67     | 64   | 61   | 56   | 74   | -     |
| 66           | 79       | 73      | 68     | 65   | 62   | 57   | 75   | -     |
| 67           | 80       | 74      | 69     | 66   | 63   | 58   | 76   | -     |
| 68           | 81       | 75      | 70     | 67   | 64   | 59   | 77   | -     |
| 69           | 82       | 76      | 71     | 68   | 65   | 60   | 78   | -     |
| 70           | 83       | 77      | 72     | 69   | 66   | 61   | 79   | -     |
| 71           | 84       | 78      | 73     | 70   | 67   | 62   | 80   | -     |
| 72           | 85       | 79      | 74     | 71   | 68   | 63   | 81   | -     |
| 73           | 86       | 80      | 75     | 72   | 69   | 64   | 83   | -     |
| 74           | 86       | 81      | 76     | 73   | 70   | 65   | 84   | -     |
| 75           | 87       | 82      | 77     | 74   | 71   | 66   | 85   | -     |
| 76           | 88       | 83      | 78     | 75   | 72   | 67   | 86   | -     |
| 77           | 89       | 84      | 79     | 76   | 73   | 68   | 87   | -     |
| 78           | 90       | 85      | 80     | 77   | 74   | 69   | 89   | -     |
| 79           | -        | 86      | 81     | 78   | 75   | 70   | 90   | -     |
| 80           | -        | 87      | 82     | 79   | 76   | 71   | -    | -     |
| 81           | -        | 88      | 83     | 80   | 77   | 73   | -    | -     |
| 82           | -        | 89      | 84     | 81   | 78   | 74   | -    | -     |
| 83           | -        | 90      | 85     | 82   | 79   | 75   | -    | -     |
| 84           | -        | -       | 86     | 83   | 80   | 76   | -    | -     |
| 85           | -        | -       | 87     | 84   | 81   | 77   | -    | -     |
| 86           | -        | -       | 88     | 86   | 82   | 78   | -    | -     |
| 87           | -        | -       | 90     | 87   | 83   | 79   | -    | -     |
| 88           | -        | -       | 91     | 88   | 84   | 80   | -    | -     |
| 89           | -        | -       | -      | 89   | 85   | 81   | -    | -     |
| 90           | -        | -       | -      | 90   | 86   | 83   | -    | -     |
| 91           | -        | -       | -      | 91   | 87   | 84   | -    | -     |
| 92           | -        | -       | -      | -    | 88   | 85   | -    | -     |
| 93           | -        | -       | -      | -    | 89   | 86   | -    | -     |
| 94           | -        | -       | -      | -    | 90   | 87   | -    | -     |

Συνέχεια στην επόμενη σελίδα

Πίνακας Α'.1 – Συνέχεια από προηγούμενη σελίδα

| <i>Phons</i> | 0.125kHz | 0.25kHz | 0.5kHz | 1kHz | 2kHz | 4kHz | 8kHz | 16kHz |
|--------------|----------|---------|--------|------|------|------|------|-------|
| 95           | -        | -       | -      | -    | -    | 88   | -    | -     |
| 96           | -        | -       | -      | -    | -    | 89   | -    | -     |
| 97           | -        | -       | -      | -    | -    | 90   | -    | -     |

Πίνακας Α'.1: Μήτρα Συσχέτισης Αμφιωτικής Ακουστότητας (*Phons*), Συχνότητας (*Hz*) και Ενέργειας (*dB SPL*)

## Παράρτημα Β΄

# Πηγαίος Κώδικας Υλοποίησης

Στο παρόν Παράρτημα, παρουσιάζεται τα τμήματα Πηγαίου Κώδικα που δημιουργήθηκαν για την υλοποίηση του Προσαρμοστικού Μοντέλου Επαυξημένης Μείξης ήχου.

```
1 function [s_calibrated]=Pascalize(sinput, target_pascal)
2
3 Pref = 20e-6;
4 s_refscaled= (sinput./Pref);
5 RMS=sqrt(mean(s_refscaled.^2));
6 SPLmat=20*log10(RMS);
7
8 % dB SPL in matlab
9 c=10.^((target_pascal-SPLmat)/20);
10 s_calibrated=c*(s_refscaled);
11
12 end
```

Πηγαίος Κώδικας B.1: Βαθμονόμηση Σήματος σε κλίμακα Pascal

```
1 function [signal_spectrum_dB]=CalculateSpectrum(s_calibrated)
2
3 N = length(s_calibrated);
4 Pref=20e-6;
5 signal_spectrum_calibrated_FFT = fft(s_calibrated);
6 signal_spectrum = signal_spectrum_calibrated_FFT * ((2/N)/sqrt(2));
7 signal_spectrum_dB = 20*log10(abs(signal_spectrum./Pref));
8
9 end
```

Πηγαίος Κώδικας B.2: Φασματική Ανάλυση σήματος και μετατροπή σε κλίμακα dB SPL

```
1 Loops=10000; %Loops for Calibrated Sine Constructor
2 p=0;
3 i=0;
4 j= Loops;
5
6 f=5000; %sine Freq
7
8 Amp = 9*(10e-4); %Amplitude initialization
9
10 for i= p:j
11
12     Amp = Amp + (10e-4);
13     disp(i)
14     Fs=48000;
15     T=5; %Seconds
16     t=0:1/Fs:T;
17     y= Amp * sin((2*pi*f*t)); %Sine Generator
18
19     rms_y=sqrt(sum(y.^2)/length(y)); %RMS
20     rms_dB=20*log10(rms_y); %RMS dB
21     Tone_dB_SPL = 94+(rms_dB); %dB SPL
22
23     Tone_dB_SPL = round(Tone_dB_SPL);
24
25     filename = ['TONE_' num2str(f) 'Hz_' num2str(Tone_dB_SPL) 'dB.wav'];
26     wavwrite(y,Fs, filename);
27
28 end
```

Πηγαίος Κώδικας B.3: Δημιουργία βαθμονομημένων Τόνων σε κλίμακα dB SPL

```

1 %Import Loudness Matrix
2 [PhonsToDbMat]=PhonsToDbMatrix();
3
4 %Excitation Matrix of Real Environment Input
5 ExcBandReal = [37 46 50 60 55 47 53];
6
7 %Excitation Matrix of Virtual Environment Input
8 ExcBandVirtual = [35 46 60 80 45 50 40];
9
10 for i=1:7
11     BLConvReal = find(ExcBandReal(i)≤ PhonsToDbMat(:,i),1,'first');
12     if (isempty(BLConvReal))
13         MatSearchRealIdx(i)=find(PhonsToDbMat(:,i)≥10,1,'first');
14     else
15         MatSearchRealIdx(i)=BLConvReal;
16     end
17
18     BLConvVirtual = find(ExcBandVirtual(i)≤ PhonsToDbMat(:,i),1,'first');
19     if (isempty(BLConvVirtual))
20         MatSearchVirtualIdx(i)=find(PhonsToDbMat(:,i)≥10,1,'first');
21     else
22         MatSearchVirtualIdx(i)=BLConvVirtual;
23     end
24 end

```

Πηγαίος Κώδικας B.4: Συνάρτηση υπολογισμού αμφιωτικής ακουστότητας ζωνών από καμπύλη διέγερσης

```

1 function VirAdaptOut = BLAdapt(SignalFiltered7Bands, FlagSmooth, ...
    LoudMatReal_Phons_Input, LoudMatVirtual_Phons_Input, DifferCoef)
2
3 %Input Loudness Matrix for Real and Virtual Environment
4 LoudMatReal_Phons=LoudMatReal_Phons_Input;
5 LoudMatVirtual_Phons=LoudMatVirtual_Phons_Input;
6
7 %Calculate AVG of Virtual Loudness Pattern
8 VirtLoudMatAVG = round(sum(LoudMatVirtual_Phons)/7);
9
10 %Define the Coefficient Matrix for synthesis virtual signal
11 LoudnessRatioMat=zeros(7,1);
12
13 for i=1:7
14     PhonsDiafora = LoudMatReal_Phons(i)-LoudMatVirtual_Phons(i)+DifferCoef;
15
16     %Define method with smoothing (FlagSmooth=0) or without (FlagSmooth>0)
17     if FlagSmooth==0
18         VirtualToVirtualAVGPhonsRatio= PhonsDiafora/ VirtLoudMatAVG;
19         PhonsDiaforaFinal = abs(round(VirtualToVirtualAVGPhonsRatio* ...
20             PhonsDiafora));
21     else
22         PhonsDiaforaFinal=PhonsDiafora;
23     end
24     dBDiafora(i) = PhonsDiaforaFinal;
25
26     % Convert dB SPL to linear value
27     LoudnessRatioMat(i) = (10.^(dBDiafora(i)/20));
28
29     SignalFiltered7BandsAdapt(:,i) = SignalFiltered7Bands(:,i) ...
30         .* LoudnessRatioMat(i);
31 end

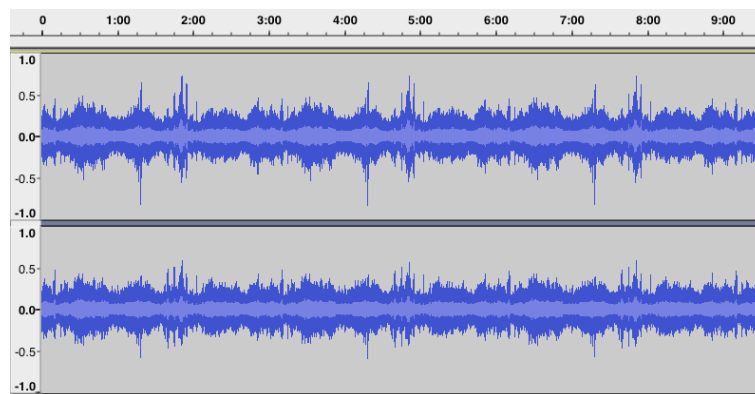
```

Πηγαίος Κώδικας B.5: Συνάρτηση υπολογισμού αμφιωτικής ακουστότητας ζωνών από καμπύλη διέγερσης

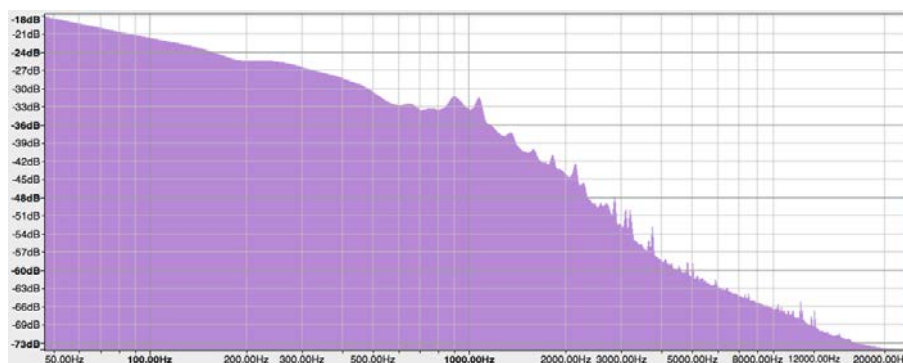
## Παράρτημα Γ΄

# Γραφική Παρουσίαση Ήχων Αξιολόγησης

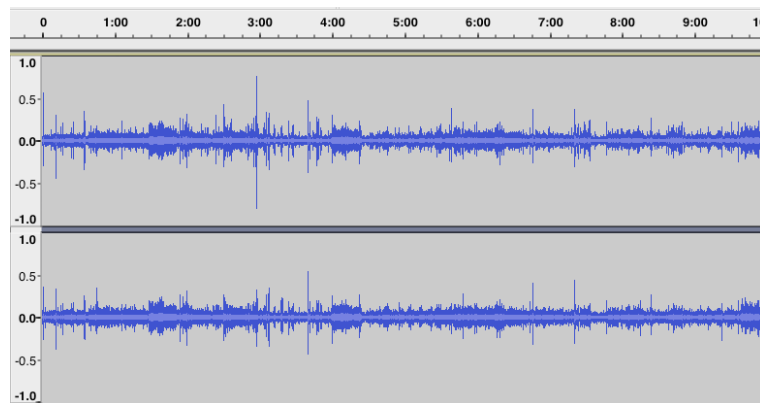
Στο παρόν παράρτημα, παρουσιάζονται οι αναλύσεις των ήχων που χρησιμοποιήθηκαν για την αξιολόγηση του προσαρμοστικού μοντέλου. Για κάθε ήχο παρουσιάζεται η κυματομορφή του καθώς και η φασματική ανάλυσή του.



Σχήμα Γ΄.1: Κυματομορφή ήχου Κίνησης Πόλης



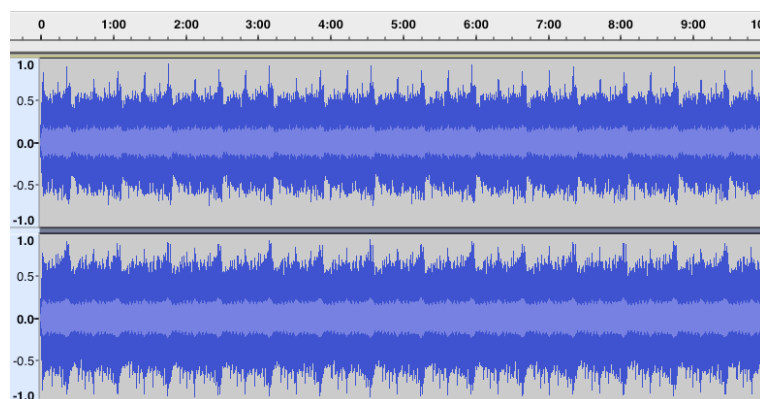
Σχήμα Γ΄.2: Φασματική Ανάλυση ήχου Κίνησης Πόλης



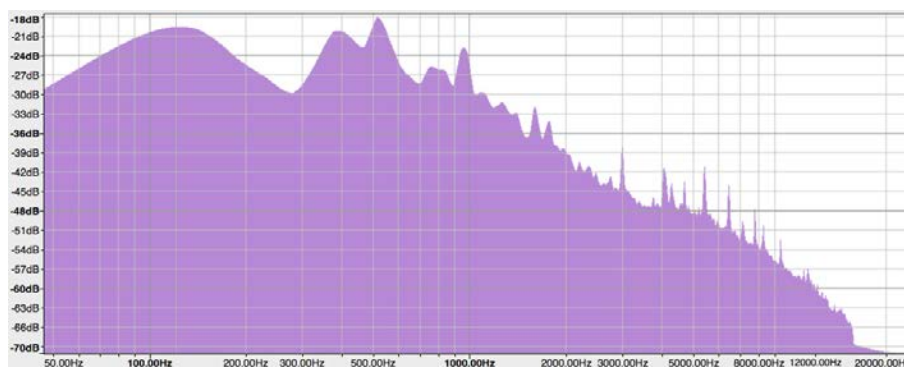
Σχήμα Γ'.3: Κυματομορφή ήχου Καφέ



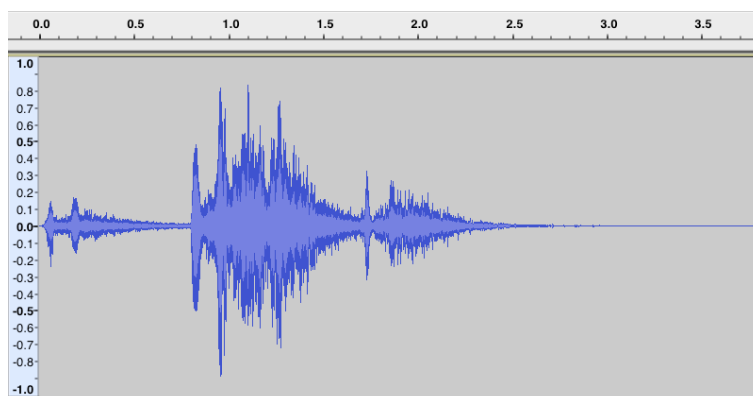
Σχήμα Γ'.4: Φασματική Ανάλυση ήχου Καφέ



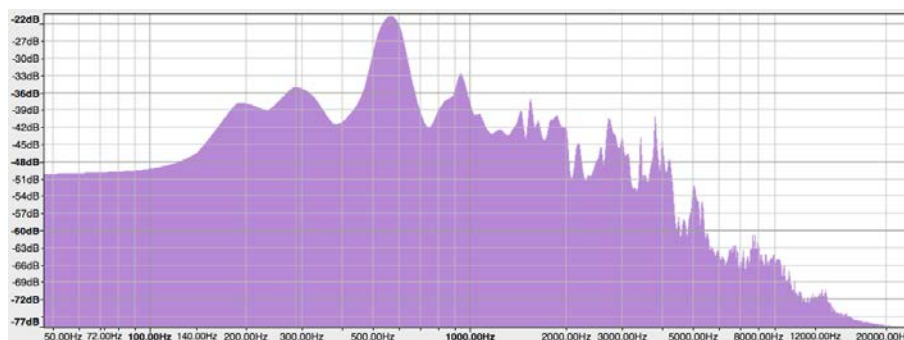
Σχήμα Γ'.5: Κυματομορφή ήχου Βιομηχανίας



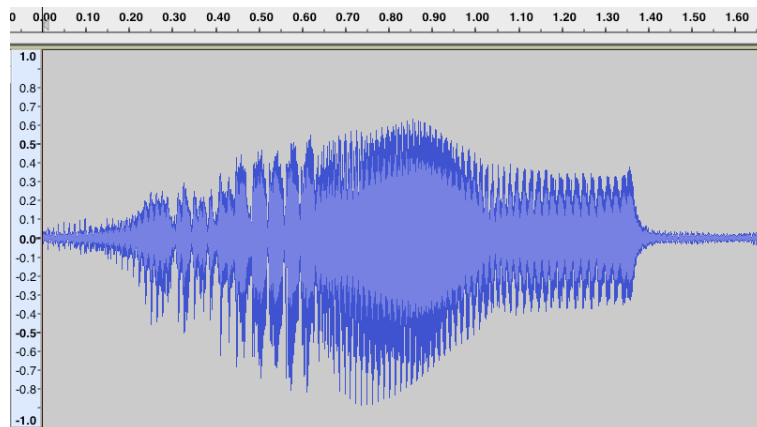
Σχήμα Γ'.6: Φασματική Ανάλυση ήχου Βιομηχανίας



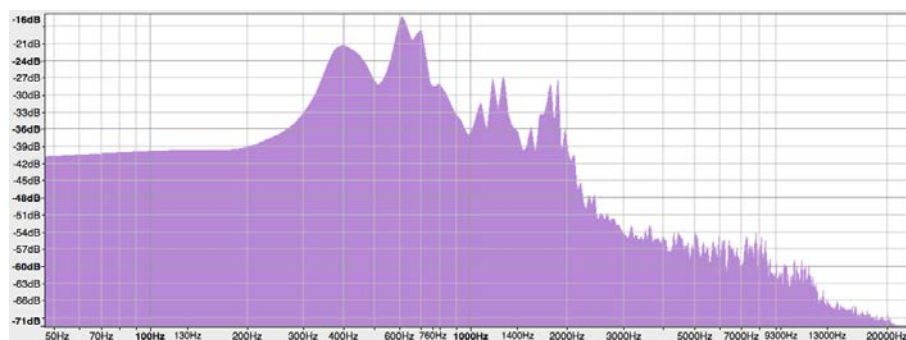
Σχήμα Γ'.7: Κυματομορφή ήχου Miss



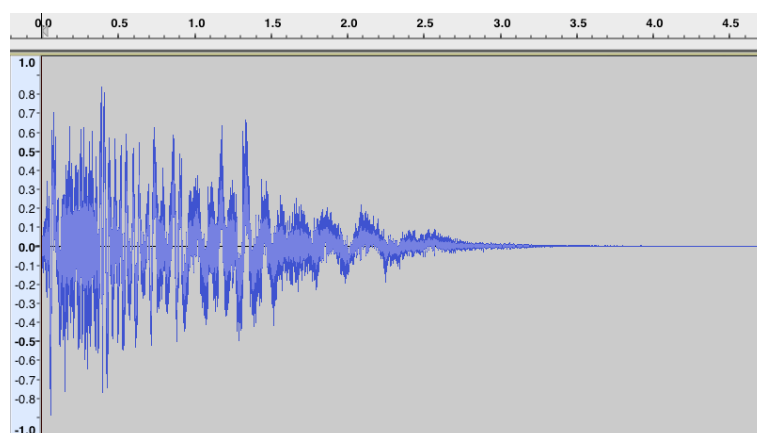
Σχήμα Γ'.8: Φασματική Ανάλυση ήχου Miss



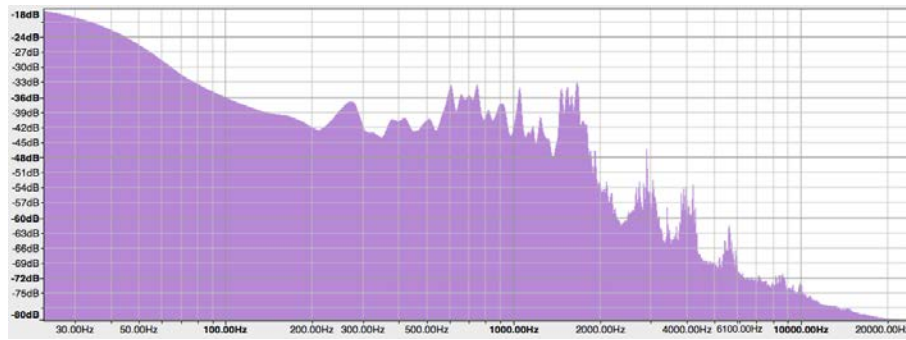
Σχήμα Γ'.9: Κυματομορφή ήχου Roboscream



Σχήμα Γ'.10: Φασματική Ανάλυση ήχου Roboscream



Σχήμα Γ'.11: Κυματομορφή ήχου Score



Σχήμα Γ'.12: Φασματική Ανάλυση ήχου Score





## Παράρτημα Δ'

# Έγγραφα Ερευνητικής Μελέτης

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                    |                                        |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|--|
| <br><b>ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ</b><br><b>ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                    | Κέρκυρα, 13/12/2019<br>Αρ. Πρωτ.: 4542 |  |
| <b>ΕΠΙΤΡΟΠΗ ΔΕΟΝΤΟΛΟΓΙΑΣ</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                    |                                        |  |
| Ταχ. Δ/ση : Ακαδημίας 1<br>Ταχ. Κώδικας : 49100, Κέρκυρα<br>Πληροφορίες : Νικολέττα Ρούσσου<br>Τηλέφωνο : 26610.87106<br>Τηλεομοιότυπο : 26610.87175<br>E-mail : roussou@ionio.gr                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Προς: Κο Νικόλαο Μουστάκα<br>Υπ. Διδάκτορα<br>1 <sup>η</sup> πάροδος Κολοκοτρώνη 34<br>49100 Κέρκυρα<br>a11mous@ionio.gr<br><br>Τηλ.: 6973919708<br>Κοιν.: Γραμματεία Μεταπτυχιακών<br>Σπουδών Τμήματος Τεχνών Ήχου<br>και Εικόνας |                                        |  |
| <b>ΘΕΜΑ: Αίτημα εκπόνησης ερευνητικής μελέτης.</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                    |                                        |  |
| Αξιότιμε κύριε Μουστάκα,<br><br>Η Επιτροπή Δεοντολογίας του Ιονίου Πανεπιστημίου, κατόπιν μελέτης του υπ.αριθμ. 4542/24-09-2019 αιτήματός σας, στο πλαίσιο της διδακτορικής διατριβής που εκπονείτε στο Τμήμα Τεχνών Ήχου και Εικόνας του Ιονίου Πανεπιστημίου με θέμα «Νοήμονα περιβάλλοντα ηχητικής επαυξημένης πραγματικότητας με χρήση φυσικής υπολογιστικής διάδρασης και ηχητικής διαδραστικής σχεδίασης» με επιβλέποντα τον Καθηγητή κ. Ανδρέα Φλώρο κατέληξε στα εξής:<br><br><b>Σύμφωνα με το κατατιθέμενο ερευνητικό πρωτόκολλο μελέτης, κρίνεται απαραίτητη η διενέργεια έρευνας, ώστε να δοκιμαστεί το σχεδιαστικό μοντέλο σε φυσικό περιβάλλον ενηλίκων ατόμων τυπικής ανάπτυξης.</b><br><br><b>Η μεθοδολογία του ερευνητικού σχεδίου ακολουθεί τις προτάσεις της επιστημονικής έρευνας, εντάσσοντας την καταγραφή των συμπερασμάτων της πειραματικής εφαρμογής του πρωτότυπου μοντέλου επαυξημένης ηχητικής μίξης μέσω ψυχοακουστικών τεστ και ηλεκτρονικού ερωτηματολογίου στον βασικό άξονα σχεδιασμού της μελέτης με τα επισυναπτόμενα τεστ να κρίνονται επαρκή για την διεξαγωγή της.</b><br><br><b>Η συμμετοχή στην έρευνα θα γίνει μόνο κατόπιν σύμφωνης γνώμης συμμετεχόντων, είναι προαιρετική και διατηρούν το δικαίωμα υπαναχώρησης σε οποιοδήποτε στάδιο διεξαγωγής της. Σχετικό έντυπο ενημέρωσης και συγκατάθεσης θα επιδοθεί σε όλους τους συμμετέχοντες και θα τηρηθεί στο αρχείο.</b><br><br>Η Επιτροπή Δεοντολογίας κατόπιν διαλογικής συζήτησης επί του φακέλου της αίτησής σας απεφάνθη ότι <b>συμφωνεί ομόφωνα</b> με το Αναλυτικό Σχέδιο Έρευνάς σας και <b>εγκρίνει</b> την |                                                                                                                                                                                                                                    |                                        |  |

διεξαγωγή αυτής στις προτεινόμενες σχολικές μονάδες, για την περίοδο Ιανουάριος 2019 – Ιούνιος 2020 με δυνατότητα επέκτασής της.

Σας ευχόμαστε καλή επιτυχία στο έργο σας.

Με τιμή

Ο Αναπληρωτής Πρύτανη  
Ακαδημαϊκών Υποθέσεων και Προσωπικού  
και Πρόεδρος της Επιτροπής Δεοντολογίας  
του Ιονίου Πανεπιστημίου

Καθηγητής Κωνσταντίνος Αγγελάκος

Εσωτερική διανομή:

- Πρύτανη Ι.Π.
- Συντονιστή Δ.Υ. Ι.Π.
- Μέλη Επιτροπής Δεοντολογίας Ι.Π.

Σχήμα Δ'.1: Αίτηση Εκπόνησης Ερευνητικής Μελέτης

### Δήλωση Συγκατάθεσης

Με την παρούσα δήλωση, την οποία υπογράφω παρακάτω, αποδέχομαι τους όρους συμμετοχής στην ερευνητική δραστηριότητα που αφορά την διδακτορική έρευνα του Νικόλαου Μουστάκα με τίτλο «Νοήμονα περιβάλλοντα ηχητικής επαυξημένης πραγματικότητας με χρήση φυσικής υπολογιστικής διάδρασης και ηχητικής διαδραστικής σχεδίασης», όπως αυτοί διαβιβάστηκαν σε εμένα από τον υπεύθυνο για τη διεξαγωγή του πειράματος.

Η προκείμενη έρευνα απευθύνεται σε εθελοντές 1) ενήλικες και 2) ανήλικους, αποσκοπώντας στην αξιολόγηση της προσαρμοσμένης ηχητικής μίξης σε περιβάλλον επαυξημένης πραγματικότητας. Η αξιολόγηση αυτή θα λάβει χώρα μέσω της συμπλήρωσης ηλεκτρονικού ερωτηματολογίου με ποιοτικές και ποσοτικές ερωτήσεις. Η διαδικασία είναι σύμφωνη με την εθνική και ευρωπαϊκή νομοθεσία, καθώς και με τις ηθικές αρχές της έρευνας, και η διαδικασία που ακολουθείται συνάδει με τον σχετικό εσωτερικό κανονισμό του Ιονίου Πανεπιστημίου. Συγκεκριμένα:

1) Η συμμετοχή των υποκειμένων είναι εξ'ολοκλήρου εθελοντική και η συγκατάθεσή τους έχει εξασφαλιστεί εκ προοιμίου. Στην περίπτωση των ανήλικων, η συγκατάθεση έχει παρασχεθεί από τους γονείς ή νόμιμους κηδεμόνες τους.

2) Όλα τα υποκείμενα θα ενημερωθούν σχετικά με τον σκοπό και τις μεθόδους της έρευνας, και θα λάβουν τα στοιχεία επικοινωνίας του υπεύθυνου ερευνητή, καθώς και αντίγραφο της παρούσας δήλωσης συγκατάθεσης.

3) Ο ερευνητής θα φροντίσει, έτσι ώστε οι συμμετέχοντες να μην νιώσουν οποιαδήποτε πίεση σχετική με τη συγκατάθεση ή την παροχή των σχολιασμών τους. Οι συμμετέχοντες διατηρούν το δικαίωμα να σταματήσουν το πείραμα και να αποσύρουν τη συμμετοχή τους εάν και όποτε αυτοί κρίνουν και χωρίς περαιτέρω συνέπειες.

4) Η συμπλήρωση των ερωτηματολογίων είναι ανώνυμη. Σε κάθε συμμετέχοντα ανατίθεται ένας μοναδικός κωδικός για την ανάλυση των αποτελεσμάτων. Μόνο οι δηλώσεις συγκατάθεσης θα είναι επώνυμες και θα αποθηκευτούν σε ειδικό χώρο φύλαξης στο γραφείο του εργαστηρίου ΕΠ.ΟΑ.ΣΗ στο κτίριο του Τ.Τ.Η.Ε., διαθέσιμες αποκλειστικά στους υπεύθυνους ερευνητές.

5) Πριν τη διεξαγωγή του πειράματος, οι συμμετέχοντες θα συμπληρώσουν φόρμα με δημογραφικά και τεχνολογικά στοιχεία, όπως η ηλικία και το φύλο τους, ή ο βαθμός εξοικείωσης με ηχητικές εφαρμογές και την επαυξημένη πραγματικότητα. Έπειτα, θα ενημερωθούν προφορικά σχετικά με την ερευνητική και πειραματική διαδικασία, και θα συμμετάσχουν σε ατομικά ηλεκτρονικά ψυχοακουστικά ηχητικά τεστ επαυξημένης πραγματικότητας συμπληρώνοντας το ηλεκτρονικό ερωτηματολόγιο.

6) Μετά την ανάλυση των αποτελεσμάτων τα ερωτηματολόγια θα αποθηκευτούν σε ειδικό χώρο φύλαξης στο γραφείο του εργαστηρίου ΕΠ.ΟΑ.ΣΗ στο κτίριο του Τ.Τ.Η.Ε., προσβάσιμο μόνο από τους υπεύθυνους ερευνητές.

ο/η συμμετέχ... \_\_\_\_\_

### Statement of Consent

I hereby accept the terms of participation in the research experiment regarding the PhD research project of Nikolaos Moustakas, "Intelligent Audio Augmented Environments with the use of physical computational interaction and interactive sound design", as they were delivered to me by the researcher, responsible for conducting the experiment.

The particular research is addressing volunteers 1) adults, and 2) minors, aiming to evaluate the quality of audio interaction in an augmented reality environment. The evaluation will occur through the completion of questionnaires with qualitative and quantitative questions. The process to be followed complies with the national and european laws regarding research ethics, as well as with the internal regulations of the Ionian University. More specifically:

- 1) The participation is voluntary and consent has been provided in advance. In the case of minors, the consent has been sought from their parents or legal guardians.
- 2) All subjects will be informed about the aims and methods of the research, and provided with correspondence data of the responsible researchers, as well as with a copy of the current statement of consent.
- 3) The researchers will ensure that no pressure whatsoever is exercised on the subjects in providing their consent and comments. The subjects may withdraw their participation at any time without further consequences.
- 4) Completing the questionnaires is anonymous. Each participant is assigned with an ID-code facilitating the analysis of the results. Only the statements of consent are signed by name and kept in the office of EP.OA.SI Lab in the Dept. of Audio & Visual Arts, accessible only to the respective researchers.
- 5) At the beginning of the experiment the participants will complete a questionnaire regarding demographic and technographic data, such as age and sex, or their previous experience with similar projects. Then they will be orally informed about the research process, and subsequently participate in electronic psychoacoustic tests of Audio Augmented Reality and filling out the electronic questionnaire.
- 6) After analysing the results the questionnaires will be kept in the office of EP.OA.SI Lab in the Dept. of Audio & Visual Arts, accessible only to the respective researchers.

the participant \_\_\_\_\_

Σχήμα Δ'.3: Δήλωση Συγκατάθεσης Υποκειμένου στην Πειραματική Διαδικασία - Αγγλική Έκδοση



## Παράρτημα Ε΄

# Συντομογραφίες - Αρκτικόλεξα - - Ακρωνύμια

|                           |                                           |
|---------------------------|-------------------------------------------|
| βλπ                       | βλέπε                                     |
| κ.λπ.                     | και λοιπά                                 |
| κ.ο.κ                     | και ούτω καθεξής                          |
| Μ.Τ.                      | Μέση Τιμή                                 |
| Σχ.                       | Σχήμα                                     |
| Εξ.                       | Εξίσωση                                   |
| Πηγ. Κώδ.                 | Πηγαίος Κώδικας                           |
| Παρ.                      | Παράρτημα                                 |
| Πιν.                      | Πίνακας                                   |
| Εν.                       | Ενότητα                                   |
| ARA                       | Augmented Reality Audio                   |
| HT                        | Hear Through                              |
| GM Model                  | Glasberg - Moore Binaural Loudness Model  |
| AR-ION Model              | Augmened Reality-Ionian Model             |
| MUSHRA Test               | Multi Stimulus Test with Hidden Reference |
| ΕΛΟΤ                      | Ελληνικός Οργανισμός Τυποποίησης          |
| ITD                       | Interaural Time Difference                |
| ILD                       | Interaural Level Difference               |
| HRTF                      | Head Related Transfer Functions           |
| HRIR                      | Head Related Impulse Responce             |
| GPS                       | Global Positioning System                 |
| CB                        | Critical Band Rate                        |
| ERB                       | Equivalent Rectangular Bandwidth          |
| Finite Impulse Response   | FIR                                       |
| Infinite impulse response | IIR                                       |
| FFT                       | Fast Fourier Transformation               |
| SPL                       | Sound Pressure Level                      |

|        |                                |
|--------|--------------------------------|
| RMS    | Round Mean Square              |
| A-DAFX | Adaptive Digital Audio Effects |
| BL     | Binaural Loudness              |

## Παράρτημα Γ'

# Απόδοση Ξενόγλωσσων Όρων

### Απόδοση

Επαυξημένη Πραγματικότητα Ήχου

Προσαρμοστικός

Αμφιωτική Ακουστότητα

Προσαρμοστικό Μοντέλο

Τεστ πολλαπλών ερεθισμάτων

με κρυμμένο το ερέθισμα αναφοράς

Σχεδίαση Ηχητικής Διάδρασης

Ηχοποίηση δεδομένων

Καμπύλες Ίσης Ακουστότητας

Ειδική Ακουστότητα

Ενεργητική Ακύρωση Θορύβου

Ακουστικά Εσωτερικού Αυτιού

Επίδραση έμφραξης

ακουστικού πόρου

Επίδραση διαρροής//ακουστικού πόρου

Αρχή Πίεσης Ακουστικού Θαλάμου

Ακουστικά κλειστού τύπου

Ακουστικά ανοικτού τύπου

Ακουστικά διέγερσης οστού

Υποβοηθούμενη ακρόαση

Συνακρόαση

Βέλτιστο σημείο ακρόασης

Χωρική Ακουστική

Θεωρία ακοής με δύο αυτιά

Χρονικές διαφορές ακοής

Διαφορές ηχητικής στάθμης ακοής

Κώνος Αβεβαιότητας

### Ξενόγλωσσος όρος

Augmented Reality Audio

Adaptive

Binaural Loudness

Adaptive Model

MUSHRA Test

Sonic Interaction Design

Audification

Equal Loudness Contours

Specific Loudness

Active Noise Cancellation

In-Ear

Occlusion Effect

Leakage Effect

Pressure Chamber Principle

Comb Filter Effect

Closed-Back headphones

Open-Back headphones

Bone-Conduction headphones

Assisted Listening

Cross-talk

Sweet Spot

Spatial Hearing

Duplex Theory

Interaural Time Differences

Interaural Level Differences

Cone of Confusion

|                                   |                                 |
|-----------------------------------|---------------------------------|
| Συναρτήσεις μοντελοποίησης        |                                 |
| εξωτερικού αυτιού                 | Head Related Transfer Functions |
| Κρουστική διέγερση αυτιού         | Head Related Impulse Responce   |
| Bluetooth πομπός                  | Beacon                          |
| Ανίχνευση Θέσης Κεφαλιού          | Head Tracking                   |
| Ισοκυματικό Φίλτρο                | Equiripple Filter               |
| Συσχέτιση                         | Correlation                     |
| Προσαρμοστικά Ψηφιακά Ηχητικά Εφέ | Adaptive Digital Audio Effects  |

# Ευρετήριο όρων

Sonic Interaction Design, 7

Ήχος, 4

Ακουστότητα, 8

Ελάχιστο Κατώφλι Ακουστότητας, 8

Εξωτερικό αυτί, 4

Επαυξημένη Πραγματικότητα Ήχου, 2

Εσωτερικό αυτί, 5

Ηχητική σχεδίαση, 7

Ηχητική σχεδίαση διάδρασης, 7

Ηχητικό Σύστημα, 9

Ηχοποίηση δεδομένων, 7

Καμπύλη Ίσης Ακουστότητας, 8

Μέσο αυτί, 5

Μοντέλα πρόβλεψης αμφιωτικής ακουστότητας, 9

Πραγματικό Περιβάλλον, 3

Προσαρμοστικά Ψηφιακά Ηχητικά Εφέ, 10

Προσαρμοστικό Ηχητικό Σύστημα ARA, 12

Προσαρμοστικότητα, 10

Φυσιολογία της ακοής, 7

Ψευδοακουστικό Περιβάλλον, 4



