



**ΑΝΩΤΑΤΟ ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ
ΚΡΗΤΗΣ ΠΑΡΑΡΤΗΜΑ ΡΕΘΥΜΝΟ ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ
ΜΟΥΣΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΑΚΟΥΣΤΙΚΗΣ**

Τίτλος Διπλωματικής Εργασίας

**Μελέτη και υλοποίηση των μεθόδων γρήγορης αποσυνέλιξης με
κανονικοποίηση και ελαχίστων μέσων τετραγώνων για την
δημιουργία ψηφιακών φίλτρων με σκοπό την αναπαραγωγή
αμφιωτικού σήματος από ηχητικό σύστημα δύο καναλιών.**

Σπουδαστές:

Ανυφαντάκης Κωνσταντίνος (Α.Μ.: 776)

Μουστάκας Θεόδωρος (Α.Μ.: 633)

Επιβλέποντες Καθηγητές:

Ζέρβας Παναγιώτης

Ξενικάκης Δημήτριος

Ρέθυμνο, Ιανουάριος 2015

Περίληψη

Στην παρούσα διπλωματική εργασία εξετάζονται και συγκρίνονται δύο τεχνικές υλοποίησης ανάστροφου ψηφιακού φίλτρου με σκοπό την ακύρωση του crosstalk και της γωνίας της θέσης των ηχείων, προκειμένου να επιτευχθεί αναπαραγωγή αμφιωτικού (binaural) αρχείου ήχου από stereo ζεύγος ηχείων. Το crosstalk και η παρουσία των ηχείων στην γωνία που είναι τοποθετημένα, είναι οι δυο κυριότεροι ανασταλτικοί παράγοντες στην ακρόαση αμφιωτικού ηχητικού σήματος από stereo ζεύγος ηχείων. Οι δύο μέθοδοι οι οποίες εξετάζονται είναι: Γρήγορη αποσυνέλιξη με κανονικοποίηση (Fast deconvolution using regularization) και ελαχίστων μέσων τετραγώνων (Least mean squares). Η βασική διαφορά των δύο αυτών μεθόδων είναι ο τρόπος κατασκευής του φίλτρου. Η κατασκευή του φίλτρου με την πρώτη μέθοδο πραγματοποιείται στο πεδίο των συχνοτήτων (frequency domain) και η δεύτερη στο πεδίο του χρόνου (time domain). Για την υλοποίηση των δύο ανάστροφων φίλτρων χρησιμοποιήθηκε το πρόγραμμα Matlab. Τα ανάστροφα ψηφιακά φίλτρα, τα οποία κατασκευάστηκαν με τις δύο παραπάνω τεχνικές, εφαρμόστηκαν σε binaural αρχεία ήχου, τα οποία στην συνέχεια χρησιμοποιήθηκαν κατά την πειραματική διαδικασία. Κατά την πειραματική διαδικασία πραγματοποιήθηκαν ακροάσεις των αρχείων ήχου που αναφέρθηκαν παραπάνω, από δεκαπέντε τυχαίους ακροατές, από τους οποίους ζητήθηκε να προσδιορίσουν την θέση στην οποία αντιλήφθηκαν την εικονική ηχητική πηγή. Τέλος πραγματοποιήθηκε στατιστική ανάλυση των αποτελεσμάτων και εξαγωγή συμπερασμάτων.

Λέξεις κλειδιά:

Ακύρωση crosstalk, ανάστροφο φίλτρο, αμφιωτικό.

Abstract

The present thesis examines and compares two methods of inverse digital filter design. The aim of the filters is the cancellation of the crosstalk and speakers' position with respect to the listener's head, in order to achieve binaural signal reproduction over stereo speaker system. The crosstalk and the speakers' position are the main inhibitor factors for binaural reproduction over loudspeakers. The methods, this thesis examines, are the Fast deconvolution using regularization and the least means squares (LMS). The main difference between these two methods, is the way of the filter construction. The filter design process, using the method of Fast deconvolution using regularization, is performed in the frequency domain, in difference with the LMS method in which the filter process is performed in time domain. For the implementation of the inverse filters was used Matlab software. The binaural audio files that were used in the experiment, passed through the inverse filters, in order to create virtual audio images. The construction of the filters has been based on the above two methods. The experimental process, in which participated 15 random subjects, comprised a listening test of comparison between a virtual acoustic image and a real source. Finally presented statistical analysis of the listening tests' results and conclusions.

Key words:

Crosstalk cancellation, inverse filter, binaural

Ευχαριστίες

Ευχαριστούμε πολύ για την πολύτιμη βοήθεια τους τον Σίμο Δελικάρη – Μανιά, τον Γιάννη, το φίλο μας Μιχάλη Τερζάκη, τον Αχιλλέα Κάππη, τον Ορέστη Λαμπαδά που μας ανέχτηκε, το Γιάννη Παπαδάτο (εξάδελφος), όλους όσους συμμετείχαν στην πειραματική διαδικασία και τέλος τους επιβλέποντες καθηγητές μας Ζέρβα Παναγιώτη και Ξενικάκη Δημήτρη.

Περιεχόμενα

Κατάλογος σχημάτων.....	vi
Κατάλογος πινάκων.....	ix
1. Βασικά στοιχεία ψυχοακουστικής και συστημάτων αναπαραγωγής.....	1
1.1 Εισαγωγή.....	1
1.2 Ηχητικός εντοπισμός (Sound localization).....	1
1.3 Συναρτήσεις Μεταφοράς του Κεφαλής (HRTF).....	2
1.4 Τεχνητό κεφάλι μικρόφωνο (Dummy Head).....	2
1.5 Stereo συστήματα.....	5
1.6 Η αναπαραγωγή του αμφιωτικού (binaural) ηχητικού σήματος.....	7
1.7 Crosstalk - Ακύρωση του Crosstalk (Crosstalk Cancelation).....	7
1.8 Στέρεο δίπολο (Stereo Dipole).....	8
1.9 Επίλογος.....	9
2. Ιστορική αναδρομή.....	10
2.1 Εισαγωγή.....	10
2.2 Ανασκόπηση των προσπαθειών δημιουργίας ανάστροφων φίλτρων για την ακύρωση του crosstalk των ηχείων.....	10
3. Ανάλυση των μεθόδων δημιουργίας ανάστροφου φίλτρου για την ακύρωση του crosstalk και της θέσης των ηχείων.....	15
3.1 Εισαγωγή.....	15
3.2 Ανάλυση συστήματος ακύρωσης του crosstalk με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση (Fast deconvolution using regularization).....	15
3.2.1 Εισαγωγή.....	15
3.2.2 Αναστροφή συστήματος.....	15
3.2.3 Κανονικοποίηση (Regularization).....	19
3.2.4 Ορισμός του πίνακα του ανάστροφου φίλτρου στο πεδίο των συχνοτήτων....	21
3.2.5 Επίλογος.....	22
3.3 Ανάλυση συστήματος ακύρωσης του crosstalk με την τεχνική ελαχίστων μέσων τετραγώνων (Least mean squares).....	23
3.3.1 Εισαγωγή.....	23
3.3.2 Θεωρητική ανάλυση.....	23
3.3.3 Ακύρωση crosstalk.....	29
3.3.4 Επίλογος.....	31
4. Εφαρμογή των τεχνικών δημιουργίας ανάστροφου φίλτρου για την ακύρωση του crosstalk και της θέσης των ηχείων.....	32
4.1 Εισαγωγή.....	32
4.2 Υλοποίηση ανάστροφου φίλτρου με την τεχνική της γρήγορης αποσυνέλιξης με κανονικοποίηση.....	32
4.2.1 Εισαγωγή.....	32
4.2.2 Εφαρμογή των βημάτων για τη δημιουργία πίνακα ανάστροφου φίλτρου στο πεδίο των συχνοτήτων.....	32

4.2.3 Εφαρμογή τιμών στις παραμέτρους, που επηρεάζουν την έξοδο του ανάστροφου φίλτρου.	35
4.2.4 Ανάλυση των αποτελεσμάτων των φίλτρων.	48
4.2.5 Η εφαρμογή του φίλτρου στο binaural σήμα.	50
4.2.6 Επίλογος.	51
4.3 Υλοποίηση ανάστροφου φίλτρου με την τεχνική των ελαχίστων μέσων τετραγώνων.	52
4.3.1 Εισαγωγή.	52
4.3.2 Ανάλυση αλγορίθμου.	52
4.3.3 Ανάλυση αποτελεσμάτων για διαφορετικές τάξεις φίλτρων.	55
4.3.4 Επίλογος.	71
5 Πειραματική διαδικασία.	72
5.1 Εισαγωγή.	72
5.2 Διάταξη πειραματικής διαδικασίας.	72
5.3 Δημιουργία binaural αρχείου ήχου.	73
5.4 Χαρακτηριστικά των ακροατών που έλαβαν μέρος στην πειραματική διαδικασία.	79
5.5 Ανάλυση αποτελεσμάτων.	80
5.6 Συμπεράσματα πειραματικής διαδικασίας.	87
5.7 Επίλογος.	88
6. Συμπεράσματα.	89
Παράρτημα Α.	95

Κατάλογος σχημάτων.

Εικόνα 1: Dummy Head KEMAR συμπεριλαμβανομένου του κορμού.	3
Εικόνα 2: Οι καμπύλες του Simonsen οι οποίες απεικονίζουν τρεις τυπικές διευθύνσεις στις οποίες μπορεί να βρίσκεται μία εικονική ηχητική πηγή κατά την αναπαραγωγή σε ένα στερεοφωνικό σύστημα (Ξενικάκης 2006).	6
Εικόνα 3: Φυσικό σύστημα αναπαραγωγής. Στο σχήμα διακρίνονται τα απευθείας ηχητικά σήματα (ipsilateral) HLL και HRR και το crosstalk (contralateral) HLR και HRL (Mouchtaris et al., 2000).	8
Εικόνα 4: Block διάγραμμα αναστροφής συστήματος δύο καναλιών (Manias, 2008).	17
Εικόνα 5: Διάγραμμα του πλάτους (α) και της φάσης (β), του όρου (3.26) (Mouchtaris, et al., 2000).	25
Εικόνα 6: Block διάγραμμα του αλγορίθμου (LMS), για τον υπολογισμό του ανάστροφου φίλτρου (Mouchtaris et al., 2000).	27
Εικόνα 7: Κρουστική απόκριση ανάστροφου φίλτρου ακύρωσης crosstalk. (α) πριν την εφαρμογή του modelling delay, (β) μετά την εφαρμογή του modelling delay.	35
Εικόνα 8: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 1). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	37
Εικόνα 9: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 1) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	38
Εικόνα 10: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 2). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	38
Εικόνα 11: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 2) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	39
Εικόνα 12: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 3). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	39
Εικόνα 13: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 3) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	40
Εικόνα 14: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 4). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	40
Εικόνα 15: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 4) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	41
Εικόνα 16: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 5). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	41
Εικόνα 17: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 5) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	42
Εικόνα 18: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 6). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	42
Εικόνα 19: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 6) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	43
Εικόνα 20: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 7). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	43
Εικόνα 21: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 7) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	44

Εικόνα 22: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 8). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	44
Εικόνα 23: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 8) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	45
Εικόνα 24: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 9). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	45
Εικόνα 25: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 9) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	46
Εικόνα 26: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 10). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.	46
Εικόνα 27: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 10) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.	47
Εικόνα 28: Η κρουστική απόκριση των HRTFs των 30° στο πεδίο του χρόνου (IRCAM database).	48
Εικόνα 29: Η κρουστική απόκριση των HRTFs των 30° στο πεδίο των συχνοτήτων (IRCAM database).	48
Εικόνα 30: Δικάναλο σύστημα ηχείων για την αναπαραγωγή binaural σήματος (Manias, 2008).	51
Εικόνα 31: Η κρουστική απόκριση των 30° στο πεδίο του χρόνου (KEMAR).	53
Εικόνα 32: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – πρώτη δοκιμή.	56
Εικόνα 33: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της πρώτης δοκιμής.	57
Εικόνα 34: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – δεύτερη δοκιμή.	58
Εικόνα 35: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της δεύτερης δοκιμής.	58
Εικόνα 36: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) τρίτη δοκιμή.	59
Εικόνα 37: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της τρίτης δοκιμής.	60
Εικόνα 38: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – τέταρτη δοκιμή.	61
Εικόνα 39: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της τέταρτης δοκιμής.	61
Εικόνα 40: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – πέμπτη δοκιμή.	62
Εικόνα 41: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της πέμπτης δοκιμής.	63
Εικόνα 42: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – έκτη δοκιμή.	64
Εικόνα 43: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της έκτης δοκιμής.	64
Εικόνα 44: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – έβδομη δοκιμή.	65
Εικόνα 45: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της έβδομης δοκιμής.	66
Εικόνα 46: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – όγδοη δοκιμή.	67
Εικόνα 47: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της όγδοης δοκιμής.	67

Εικόνα 48: Οι όροι diagonal και off-diagonal για το αριστερό (α , β) και το δεξί κανάλι (γ , δ) – ένατη δοκιμή.....	68
Εικόνα 49: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της ένατης δοκιμής.....	69
Εικόνα 50: Οι όροι diagonal και off-diagonal για το αριστερό (A-B) και το δεξί κανάλι (Γ - Δ) – δέκατη δοκιμή.....	70
Εικόνα 51: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της δέκατης δοκιμής.....	70
Εικόνα 52: Πειραματική διάταξη. Διακρίνονται τα ηχεία αναπαραγωγής του binaural σήματος σε stereo διάταξη (αριστερό ηχείο, δεξί ηχείο) και το μετακινούμενο ηχείο (ηχείο αναφοράς) στην εκάστοτε γωνία οπού ηχεί η εικονική πηγή.....	74
Εικόνα 53: Η κρουστική απόκριση 0° των HRTFs στο πεδίο του χρόνου (IRCAM database).	75
Εικόνα 54: Η κρουστική απόκριση 30° των HRTFs στο πεδίο του χρόνου (IRCAM database).	75
Εικόνα 55: Η κρουστική απόκριση 45° των HRTFs στο πεδίο του χρόνου (IRCAM database).	76
Εικόνα 56: Η κρουστική απόκριση 60° των HRTFs στο πεδίο του χρόνου (IRCAM database).	76
Εικόνα 57: Η κρουστική απόκριση 75° των HRTFs στο πεδίο του χρόνου (IRCAM database).	77
Εικόνα 58: Η κρουστική απόκριση 90° των HRTFs στο πεδίο του χρόνου (IRCAM database).	77
Εικόνα 59: Η κρουστική απόκριση 135° των HRTFs στο πεδίο του χρόνου (IRCAM database).....	78
Εικόνα 60: Η κρουστική απόκριση 180° των HRTFs στο πεδίο του χρόνου (IRCAM database).....	78
Εικόνα 61: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 0°	80
Εικόνα 62: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 30°	81
Εικόνα 63: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 45°	82
Εικόνα 64: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 60°	83
Εικόνα 65: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 75°	84
Εικόνα 66: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 90°	85
Εικόνα 67: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 135°	86
Εικόνα 68: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 180°	86

Κατάλογος πινάκων

Πίνακας 1: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου.	36
Πίνακας 2: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την πρώτη δοκιμή.	55
Πίνακας 3: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την δεύτερη δοκιμή.	57
Πίνακας 4: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την τρίτη δοκιμή.	59
Πίνακας 5: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την τέταρτη δοκιμή.	60
Πίνακας 6: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την πέμπτη δοκιμή.	62
Πίνακας 7: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την έκτη δοκιμή.	63
Πίνακας 8: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την έβδομη δοκιμή.	65
Πίνακας 9: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την όγδοη δοκιμή.	66
Πίνακας 10: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την ένατη δοκιμή.	68
Πίνακας 11: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την δέκατη δοκιμή.	69
Πίνακας 12: Παρουσίαση των διαστάσεων της κεφαλής του κάθε ακροατή που έλαβε μέρος στην πειραματική διαδικασία και ο μέσος όρος των διαστάσεων του συνόλου των ακροατών.	79

1. Βασικά στοιχεία ψυχοακουστικής και συστημάτων αναπαραγωγής.

1.1 Εισαγωγή.

Σε αυτό το κεφάλαιο γίνεται αναφορά σε βασικές αρχές που διέπουν τη λειτουργία του ανθρώπινου ψυχοακουστικού συστήματος, όσον αφορά τον εντοπισμό των ηχητικών πηγών στο χώρο. Επιπλέον αναλύονται οι αρχές λειτουργίας της stereo και αμφιωτικής αναπαραγωγής, την οποία από δω και στο εξής θα αποκαλούμε binaural αναπαραγωγή, αλλά και των μέσων ηχογράφησης αμφιωτικών (binaural) σημάτων. Τέλος γίνεται αναφορά στο stereo δίπολο ως διάταξη αναπαραγωγής binaural σήματος από δύο ηχεία.

1.2 Ηχητικός εντοπισμός (Sound localization).

Ο άνθρωπος έχει τη δυνατότητα να εντοπίζει την κατεύθυνση από την οποία ηχεί μία ηχητική πηγή λόγω της binaural ακοής. Χρησιμοποιώντας τα δύο αισθητήρια της ακοής μας, δηλαδή τα αυτιά μας, έχουμε τη δυνατότητα να εντοπίσουμε τη θέση και την πορεία μίας ή και περισσότερων ηχητικών πηγών στο χώρο. Αυτό συμβαίνει με τη σύγκριση των σημάτων των δύο αυτιών από τον ανθρώπινο εγκέφαλο.

Σύμφωνα με το παράδειγμα του Vorlander, υποθέτοντας έναν άνθρωπο, σε ελεύθερο πεδίο (απαλλαγμένο από ανακλάσεις) και μια ηχητική πηγή η οποία παράγει ένα επίπεδο ηχητικό κύμα, από συγκεκριμένη θέση στον χώρο, τότε κύμα αυτό, προτού φτάσει στα τύμπανα του ακροατή, ανακλάται και περιθλάται στους ώμους, το κεφάλι, τα πτερύγια και το ακουστικό κανάλι του εκάστοτε αυτιού. Το επίπεδο ηχητικό κύμα έχει πλέον παραμορφωθεί σε σχέση με το αρχικό (Vorlander, 2008). Η παραμόρφωση είναι άμεσα εξαρτώμενη από την γωνία της πηγής σε σχέση με τον ακροατή. Για κάθε διαφορετική θέση στον χώρο, η παραμόρφωση που υπόκειται το κύμα είναι μοναδική. Στην περίπτωση που το κύμα ήταν σφαιρικό η παραμόρφωση του θα εξαρτώνταν επίσης και από την απόσταση της πηγής σε σχέση με τον ακροατή. Η διαφορά των δύο σημάτων που καταλήγουν στα αυτιά του ακροατή έγκειται στην φάση και το πλάτος, δηλαδή την διαφορά του χρόνου άφιξης των δύο σημάτων και την διαφορά έντασης αντίστοιχα. Εάν για παράδειγμα μία πηγή ηχούσε στις 90 μοίρες δεξιά από τον ακροατή, το δεξιό αυτί θα αντιλαμβάνονταν πρώτο το ερέθισμα του ηχητικού κύματος επειδή βρίσκεται πιο κοντά στην πηγή και

σε μεγαλύτερη ένταση από το αριστερό εξαιτίας της ηχητικής σκιάς που προκαλείται από το κεφάλι. Οι διαφορές αυτές μεταξύ των σημάτων των δύο αυτιών είναι γνωστές ως Interaural Time Difference (ITD) και Interaural Level Difference (ILD). Όταν το ηχητικό κύμα έχει την ίδια απόσταση από τα αυτιά του ακροατή προέρχεται δηλαδή από ευθεία μπροστά, πίσω ή πάνω τότε η διαφορές των παραπάνω τιμών ελαχιστοποιούνται και η παραμόρφωση που γίνεται στα δύο σήματα που φθάνουν στα αυτιά είναι πανομοιότυπη. Η παραμόρφωση που προκαλείται στα ηχητικά σήματα των δύο αυτιών περιγράφεται από τις Συναρτήσεις Μεταφοράς του Κεφαλής, τις Head Related Transfer Functions (HRTF).

1.3 Συναρτήσεις Μεταφοράς του Κεφαλής (HRTF).

HRTF είναι η συνάρτηση μεταφοράς η οποία περιγράφει την διαμόρφωση που υπόκειται ένα ηχητικό κύμα, από τον κορμό και το ανθρώπινο κεφάλι, αναλόγως την γωνία από την οποία εκπέμπει η πηγή σε σχέση με τον ακροατή. Χαρακτηρίζει δηλαδή, το πώς ένα αυτί λαμβάνει ένα ηχητικό ερέθισμα από ένα συγκεκριμένο σημείο στον χώρο.

Ένα ζεύγος HRTFs για τα δύο αυτιά, μπορούν να χρησιμοποιηθούν για να συνθέσουν ένα αμφιωτικό ήχο, που δημιουργεί την ψευδαίσθηση ότι ένα ηχητικό σήμα προέρχεται από το σημείο εκείνο που περιγράφει η συνάρτηση μεταφοράς. Τα HRTFs διαφοροποιούνται από άνθρωπο σε άνθρωπο, ανάλογα με το σχήμα και τις διαστάσεις των παραγόντων που τα επηρεάζουν, δηλαδή το κεφάλι, τους ώμους και τα πτερύγια των αυτιών (Vorlander, 2008).

1.4 Τεχνητό κεφάλι μικρόφωνο (Dummy Head).

Το 1960 ξεκίνησαν οι πρώτες προσπάθειες με σκοπό τη δημιουργία ενός συγκεκριμένου Dummy Head μικροφώνου μετρήσεων και προσομοίωσης, από τους Burkhard, Suchs και Shaw. Το 1972 δημιουργείται το KEMAR, Εικόνα 1, το οποίο ήταν η πρώτη εφαρμογή μέτρησης βοηθημάτων ακοής, και στη συνέχεια χρησιμοποιήθηκε σαν Dummy Head αναφοράς σε ευρύτερο φάσμα. Χρησιμοποιείται μάλιστα μέχρι σήμερα ως μία από τις βάσεις δεδομένων αναφοράς των HRTFs.



Εικόνα 1: Dummy Head KEMAR συμπεριλαμβανομένου του κορμού.

Το κυρίως πρόβλημα στην έρευνα και ανάπτυξη των Dummy Heads που έπρεπε να λυθεί ήταν ο καθορισμός και η υλοποίηση ενός Dummy Head μέσου όρου, το οποίο θα αναπαριστά ένα ακροατή μέσου όρου. Ο στόχος δηλαδή ήταν η υλοποίηση ενός Dummy Head, με διαστάσεις κατάλληλες, ούτως ώστε να αναπαριστά το μέσο όρο των διαστάσεων του κεφαλιού και των ώμων των ακροατών. Το αποτέλεσμα, παρ' όλα αυτά, ίσως να μην ήταν το ιδανικό στα πειράματα εντοπισμού της θέσης της ηχητικής πηγής από τυχαίους ακροατές. Λύση αυτού του προβλήματος θεωρήθηκε η χρησιμοποίηση διαφόρων γεωμετρικών και κατασκευαστικών μέσων όρων.

Ο Genuit ανέπτυξε ένα μαθηματικό μοντέλο περίθλασης, βασισμένο σε ελλειπτικά και κυλινδρικά στοιχεία. Το πτερύγιο είναι επίσης απλουστευμένο ως προς την είσοδο του ακουστικού καναλιού (Genuit, 1984). Η ακριβής θέση του σημείου εισόδου του ακουστικού καναλιού, χρησιμοποιείται περισσότερο ως σημείο αναφοράς για μετρήσεις πηγών, κοντά στο αυτί διότι είναι πολύ σημαντικό για τις σωστές καμπύλες εντοπισμού.

Όταν χρησιμοποιούνται Dummy Heads για μετρήσεις, έρευνα και ηχογραφήσεις, στην αντίληψη της κατεύθυνσης και της απόστασης, ένα μεγάλο μέρος των ακροατών που συμμετέχουν, αναφέρουν προβλήματα ως προς το αποτέλεσμα που ακούν. Τα αποτελέσματα των πειραμάτων έχουν δείξει ότι οι

περισσότεροι ακροατές έχουν πρόβλημα στον προσδιορισμό του μπροστά - πίσω και του προσδιορισμού της θέσης της ηχητικής πηγής στο διάμεσο επίπεδο. Αν συγκρίνουμε ένα αριθμό ακροατών με ένα Dummy Head, θα διαπιστώσουμε ότι κάποιοι ακροατές ακούν το επιθυμητό και κάποιοι άλλοι όχι. Αυτοί που ακούν το επιθυμητό είναι αυτοί που τα χαρακτηριστικά τους ταιριάζουν περισσότερο με το συγκεκριμένο Dummy Head.

Προσεγγίσεις για κατασκευή Dummy Heads, τα οποία να ταιριάζουν καλύτερα με τους ακροατές, πραγματοποιήθηκαν από τους (Schmitz, 1995) και (Christensen, 2000). Η διαδικασία αυτή τη φορά δεν ήταν η μέτρηση και δημιουργία Dummy Head μέσου όρου, αλλά η μέτρηση μεγάλου αριθμού HRTFs, μεγάλου αριθμού διαφορετικών ανθρώπων, με σκοπό τη δημιουργία μιας βάσης δεδομένων την οποία στη συνέχεια θα χρησιμοποιούσαν σε test ηχητικού εντοπισμού. Τα HRTFs της βάσης δεδομένων, που δημιουργήθηκε, αποδείχθηκε από τα tests ηχητικού εντοπισμού που πραγματοποιήθηκαν, ότι είναι η καλύτερη επιλογή HRTFs, αφού το μεγαλύτερο μέρος των ακροατών άκουσε το επιθυμητό. Αυτή η προσέγγιση δεν στηρίζεται σε γεωμετρικούς ή κατασκευαστικούς μέσους όρους. Στηρίζεται σε επιλεγμένα HRTFs συγκεκριμένων ακροατών, στις καμπύλες των οποίων περιλαμβάνονται σημαντικές πληροφορίες, όπως είναι οι κορυφές και τα βυθίσματα στο φάσμα του ήχου. Πληροφορίες οι οποίες είναι καθοριστικές για τον εντοπισμό της πηγής από τον εκάστοτε ακροατή. Πρόσφατα αποτελέσματα έδειξαν ότι η σωστή επιλογή ανθρώπινου κεφαλιού, στο οποίο θα είναι τοποθετημένα μικρόφωνα προσαρμοσμένα στα αυτιά έχουν πολύ καλύτερα αποτελέσματα σε σχέση με τα Dummy Heads. Τα Dummy Heads τα οποία σχεδιάστηκαν και δημιουργήθηκαν με οδηγό έναν άνθρωπο (όχι από μέσο όρο ατόμων), λειτουργούν εξίσου καλά όσο και ένα ανθρώπινο κεφάλι.

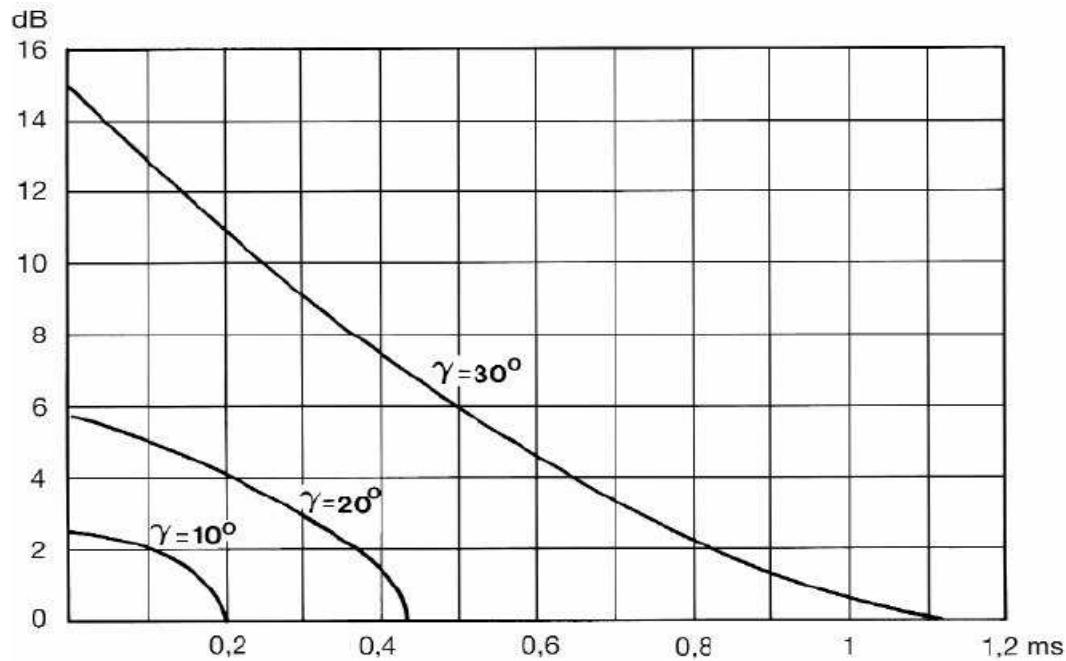
Είναι πολύ σημαντική η επιλογή του Dummy Head σε σχέση με τον ακροατή. Όλες οι έρευνες, μέχρι σήμερα, υστερούν στην κατά κάποιο τρόπο τυχαία επιλογή του Dummy Head σε σχέση με τον ακροατή και στο ότι υπάρχει περιορισμένος αριθμός μετρημένων HRTFs ξεχωριστών ατόμων. Η πιο περιεκτική μελέτη παραμένει αυτή των (Barkhard & Sachs, 1975). Σήμερα θα μπορούσε να υποστηριχτεί ότι ο πληθυσμός έχει αλλάξει τα τελευταία 30 χρόνια. Οι αλλαγές αυτές παρατηρούνται κυρίως στο μέγεθος του κεφαλιού και των ώμων, τα οποία έχουν μεγαλύτερες διαστάσεις. Επίσης αλλαγές παρατηρούνται και στο μέγεθος των

περυγίων. Θα μπορούσε να συζητηθεί το ερώτημα, αν ένα συγκεκριμένο Dummy Head μπορεί να είναι επαρκές στο να καλύψει όλους τους ανθρώπους.

Σήμερα, τα Dummy Heads και η binaural τεχνολογία είναι υψηλής ποιότητας. Είναι διαθέσιμα διεθνή πρότυπα στην τεχνολογία τηλεπικοινωνιών και ακουστικών μετρήσεων. Από την άλλη οι ηχογραφήσεις με Dummy Head και η αναπαραγωγή τους σε οικιακή χρήση δεν ήταν τόσο επιτυχημένες εμπορικά. Παρόλα αυτά, δημιουργούνται και αναπτύσσονται διάφορα πεδία, όπως το 3-D audio, εικονικά περιβάλλοντα (virtual environments), εξέλιξη των ακουστικών (headphone development). Παρόλα αυτά, κάποια ερωτήματα σχετικά με τη δυνατότητα εφαρμογής Dummy Head μέσου όρου και για τα δύο φύλα σε όλο τον κόσμο αλλά και στα παιδιά, παραμένουν αναπάντητα.

1.5 Stereo συστήματα.

Τα πρώτα συστήματα αναπαραγωγής ήχου, τα οποία φιλοδοξούσαν να πετύχουν στερεοφωνική αναπαραγωγή είχαν τις ρίζες τους στη θεωρητική εργασία του A. D. Blumlein στις αρχές της δεκαετίας του 1930 και εμφανίστηκαν ως εμπορική εφαρμογή το 1950 με το όνομα «Stereophonic systems». Όπως υποδηλώνει το όνομα, η νέα πτυχή εισήχθη μέσω δύο καναλιών, σε σύγκριση με τα τότε γνωστά μονοφωνικά συστήματα. Η λειτουργία τους βασίζεται στις καμπύλες του Simonsen, Εικόνα 2, και στα ITD και ILD που αναφέρθηκαν παραπάνω. Ο στόχος ήταν η δημιουργία ενός συστήματος που θα έδινε στον ακροατή την αίσθηση της θέσης της ηχητικής πηγής σε οποιοδήποτε σημείο στον χώρο που περιβάλλει τον ακροατή. Όπως έχει αποδειχθεί στη σχεδόν καθολική χρήση των, παραγωγών, ηχοληπτών, μηχανικών ήχου αλλά και στην ιδιωτική χρήση, με την κατάλληλη χρήση της τεχνικής του stereo panning, είναι δυνατή η τοποθέτηση της εικονικής ηχητικής πηγής σε οποιαδήποτε θέση ανάμεσα στα δύο ηχεία.



Εικόνα 2: Οι καμπύλες του Simonsen οι οποίες απεικονίζουν τρεις τυπικές διευθύνσεις στις οποίες μπορεί να βρίσκεται μία εικονική ηχητική πηγή κατά την αναπαραγωγή σε ένα στερεοφωνικό σύστημα (Ξενικάκης 2006).

Η τεχνική του stereo panning μπορεί να λειτουργήσει είτε με τη χρήση της διαφοράς έντασης ανάμεσα στα δύο κανάλια του stereo συστήματος, είτε εισάγοντας μία καθυστέρηση στο ένα από τα δύο κανάλια. Από την άλλη πλευρά, δείχνει να μη λειτουργεί, όσον αφορά τη δημιουργία μιας πλήρους τρισδιάστατης ηχητικής εικόνας. Οι εικονικές ηχητικές πηγές έξω από τα όρια της γωνίας, στην οποία είναι τοποθετημένα τα ηχεία, αλλά και όσον αφορά στις εικονικές ηχητικές πηγές που θα μπορούσαν να αναπαρασταθούν καθ' ύψος, εμφανίζονται ασταθώς αν όχι καθόλου.

Πρέπει να σημειωθεί ότι, παρά την ύπαρξη ενός ολοκληρωμένου θεωρητικού υπόβαθρου για τη στερεοφωνική τεχνική, οι μέθοδοι που χρησιμοποιούνται στην καθημερινή εφαρμογή του είναι σε πολλές περιπτώσεις, με βάση την πρακτική εμπειρία και την προσωπική αίσθηση του εκάστοτε ηχολήπτη και παραγωγού. Ακόμα και η τοποθέτηση των ηχείων από τους χρήστες - ακροατές γίνεται εμπειρικά και σε σχέση με το χώρο αναπαραγωγής και όχι με τον κανόνα του ισοσκελούς τριγώνου.

Κατά τη διάρκεια της «στερεοφωνικής εποχής» τα τελευταία 50 χρόνια, έχουν πραγματοποιηθεί πολλές θεωρητικές διατυπώσεις και πειραματικά σχέδια, προκειμένου να αντιμετωπιστεί η ανεπάρκεια των στερεοφωνικών συστημάτων να αναπαράγουν πλήρως τα χωρικά χαρακτηριστικά του ηχογραφημένου ηχητικού πεδίου.

1.6 Η αναπαραγωγή του αμφιωτικού (binaural) ηχητικού σήματος.

Η binaural αναπαραγωγή παρέχει στον ακροατή την αίσθηση ότι βρίσκεται «βυθισμένος» σε ένα περιβάλλον στο οποίο έχει τη δυνατότητα να ακούσει ήχους να ηχούν από συγκεκριμένες κατευθύνσεις στον τρισδιάστο χώρο. Μελέτες σχετικά με την αντίληψη και τον εντοπισμό των ήχων από τα δύο αυτιά, δηλαδή την binaural ακρόαση, οδήγησαν στην ανάπτυξη συστημάτων που βασίζονται στην ανθρώπινη ικανότητα ακρόασης. Τα συστήματα αυτά, έχουν την δυνατότητα να «ξεγελούν» το ανθρώπινο ψυχοακουστικό σύστημα και συγκεκριμένα τον μηχανισμό εντοπισμού, με τον έλεγχο των HRTFs και κατ' επέκταση του ITD και του ILD μεταξύ των δύο αυτιών.

Binaural ηχογράφηση είναι η διαδικασία της λήψης ενός ηχητικού γεγονότος με τον ίδιο τρόπο που θα το λάμβανε ένας άνθρωπος και πιο συγκεκριμένα τα αυτιά του. Αυτό επιτυγχάνεται με δύο μικρόφωνα σε διάταξη παρόμοια με τα αυτιά του ανθρώπου. Ο πιο εύκολος τρόπος να επιτευχθεί αυτό είναι με ένα Dummy Head.

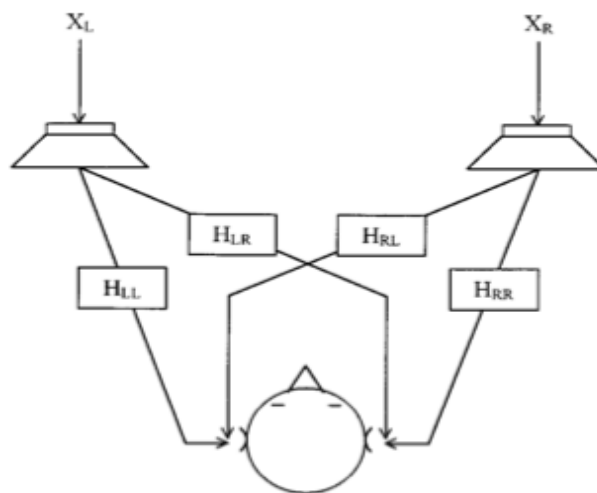
Η ιδέα της binaural τεχνολογίας είναι να δώσει στους ακροατές την αντίληψη της ακουστικής εμπειρίας με την παρουσίαση ηχητικών σημάτων στα αυτιά τους, τα οποία προσεγγίζουν τα ηχητικά σήματα ενός πραγματικού ακουστικού περιβάλλοντος. Τα ηχητικά συστήματα απεικόνισης, εικονικών ηχητικών πηγών βασίζονται σ' αυτήν την τεχνολογία. Ένα σύστημα που βασίζεται στην binaural τεχνολογία μπορεί να αναπαράγει μια ακριβή ψευδαίσθηση ενός εικονικού, ακουστικού χώρου, συμπεριλαμβανομένης της κατεύθυνσης και της απόστασης. Binaural σήματα μπορούν να αναπαραχθούν από ακουστικά ή από ηχεία. Το βασικό πλεονέκτημα στη χρήση ακουστικών είναι ότι τα δύο σήματα μεταφέρονται απευθείας στα αυτιά του ακροατή. Η αναπαραγωγή binaural σημάτων από τα ηχεία παρουσιάζει κάποια προβλήματα. Το σημαντικότερο πρόβλημα είναι η παρουσία του crosstalk των ηχείων σε σχέση με τον ακροατή (Alastair Sibbald, 1999).

1.7 Crosstalk - Ακύρωση του Crosstalk (Crosstalk Cancelation).

Όταν αναπαράγεται ένα αρχείο ήχου αποτελούμενο από δύο κανάλια, ακούγοντάς το ένας ακροατής από ένα ζεύγος ακουστικών, τα σήματα τα οποία φτάνουν στα αυτιά του είναι δύο. Από το αριστερό αυτί ακούει το αριστερό

ακουστικό και από το δεξιό αντί ακούει το δεξιό ακουστικό. Αυτή η συνθήκη ευνοεί την binaural αναπαραγωγή, αφού στο εκάστοτε αυτί, φτάνει μόνο το σήμα του αντίστοιχου μικροφώνου του Dummy Head. Τι γίνεται όμως στην περίπτωση που η αναπαραγωγή του ίδιου binaural σήματος πραγματοποιείται από δύο ηχεία τοποθετημένα στις ± 30 μοίρες;

Αυτό που συμβαίνει στην περίπτωση της αναπαραγωγής binaural σήματος από δύο ηχεία τοποθετημένα στις ± 30 μοίρες από το κεφάλι του ακροατή, είναι ότι εισάγονται στο σύστημα ακρόασης δύο νέα δεδομένα, εικόνα 3. Αυτά είναι η φυσική θέση των ηχείων και το crosstalk των ηχείων σε σχέση με το κεφάλι του ακροατή. Ο ακροατής δηλαδή ακούει το αριστερό ηχείο από το αριστερό και από το δεξιό αυτί και αντιστρόφως. Για να καταφέρει ο ακροατής να πραγματοποιήσει ακρόαση binaural σήματος από τα ηχεία, είναι απαραίτητη η ακύρωση του crosstalk αλλά και της θέσης των ηχείων (Alastair Sibbald, 1999). Αυτό μπορεί να πραγματοποιηθεί με την εφαρμογή ανάστροφου φίλτρου στο σήμα. Το ανάστροφο φίλτρο αναιρεί τη γωνία του ηχείου σε σχέση με το κεφάλι του ακροατή και εξαλείφει την παρουσία του crosstalk.



Εικόνα 3: Φυσικό σύστημα αναπαραγωγής. Στο σχήμα διακρίνονται τα απευθείας ηχητικά σήματα (ipsilateral) H_{LL} και H_{RR} και το crosstalk (contralateral) H_{LR} και H_{RL} (Mouhtaridis et al., 2000).

1.8 Στέρεο δίπολο (Stereo Dipole).

Το Stereo δίπολο είναι ένα σύστημα ηχητικής απεικόνισης, ηχητικών εικονικών πηγών από δύο ηχεία. Η τοποθέτηση των ηχείων είναι στις 10 μοίρες, όπως κοιτά ο ακροατής, και σε κοντινή απόσταση το ένα από το άλλο. Το πλεονέκτημα αυτού του συστήματος είναι ότι το ηχητικό πεδίο που δημιουργεί είναι πολύ πιο απλό

από το κλασικό stereo των ± 30 μοιρών. Προσομοιώσεις στο stereo δίπολο, έχουν δείξει ότι το hot spot, στο οποίο γίνεται αντιληπτή, από τον ακροατή, η εικονική θέση της πηγής στο τρισδιάστατο πεδίο, είναι διευρυμένο σε σχέση με το stereo των ± 30 μοιρών. Επιπροσθέτως, η τοποθέτηση των ηχείων σε κοντινή μεταξύ τους απόσταση, προκαλεί μια αισθητή μείωση της «συχνότητας κουδουνίσματος» (ringing frequency). Η συχνότητα κουδουνίσματος μπορεί να προκληθεί από πολλές αιτίες, όπως από την εφαρμογή ψηφιακών φίλτρων στο σήμα.

Επιπλέον, οι μελέτες για την επίδραση των θέσεων των ηχείων που σχετίζονται με το κατά πόσο ανταποκρίνεται το stereo δίπολο στο συχνοτικό φάσμα λειτουργίας του, δείχνουν ότι σε σύγκριση με την τοποθέτηση των ηχείων στις 30 μοίρες και στις 15 μοίρες, το crosstalk cancellation στη συχνοτική περιοχή, από τα 250 Hz και κάτω, είναι μικρότερο, αλλά στη συχνοτική περιοχή από τα 500 Hz και πάνω το crosstalk cancellation είναι μεγαλύτερο (Lopez & Gonzalez, 2001). Απαιτείται μια αναγκαία αύξηση της στάθμης των χαμηλών συχνοτήτων για να επιτευχθεί μια επαρκής ακύρωση του crosstalk των ηχείων, σε σχέση με τον ακροατή, σ' αυτή τη συχνοτική περιοχή (Nelson et al., 1997).

1.9 Επίλογος

Στο πρώτο κεφάλαιο έγινε αναφορά στην λειτουργία του ψυχοακουστικού συστήματος όσον αφορά τον εντοπισμό της ηχητικής πηγής. Δόθηκε ο ορισμός των head related transfer functions (HRTFs). Επιπροσθέτως έγινε αναφορά στα dummy heads και στην εξέλιξη τους αλλά και στην εξέλιξη των stereo συστημάτων αναπαραγωγής και της binaural αναπαραγωγής. Ακόμα παρουσιάστηκαν τα προβλήματα που προκύπτουν κατά την αναπαραγωγή binaural ηχητικού σήματος από stereo ζεύγος ηχείων (crosstalk, θέση των ηχείων). Τέλος έγινε μία μικρή αναφορά στο stereo δίπολο, το οποίο παρουσιάζει πολλά πλεονεκτήματα κατά την αναπαραγωγή binaural αρχείου ήχου από δύο ηχεία.

2. Ιστορική αναδρομή.

2.1 Εισαγωγή.

Σε αυτό το κεφάλαιο γίνεται αναφορά στις απόπειρες δημιουργίας συστημάτων αναπαραγωγής binaural ηχητικού σήματος από ηχεία, από τις πρώτες απόπειρες των Atal και Schroeder μέχρι και σήμερα.

2.2 Ανασκόπηση των προσπαθειών δημιουργίας ανάστροφων φίλτρων για την ακύρωση του crosstalk και των ηχείων.

Ο Bauer ήταν ο πρώτος, που εισήγαγε την ιδέα για την ακύρωση του crosstalk (Bauer, 1961). Οι πρώτοι όμως που ξεκίνησαν να ερευνούν ουσιαστικά για τη δημιουργία ενός ανάστροφου φίλτρου, το οποίο θα ακύρωνε την ύπαρξη του crosstalk, ήταν οι Atal και Schroeder (Atal & Schroeder, 1962). Αυτή ήταν η αρχή μιας μεγάλης προσπάθειας, που έγινε, για τη δημιουργία μεθόδων οι οποίες θα ήταν αποτελεσματικότερες στην δημιουργία ανάστροφων φίλτρων τα οποία θα ακύρωναν το crosstalk. Η ακύρωση του crosstalk των ηχείων μπορεί να επιτευχθεί εφαρμόζοντας στο binaural ηχητικό σήμα, ανάστροφα φίλτρα, τα οποία υπολογίζονται αναστρέφοντας ένα πίνακα τιμών, ο οποίος εμπεριέχει τις ηλεκτροακουστικές συναρτήσεις μεταφοράς, από τα ηχεία στα αυτιά του ακροατή. Το σήμα με το οποίο τροφοδοτούνται τα ηχεία έχει φιλτραριστεί από τα ανάστροφα φίλτρα, τα οποία θα ακυρώσουν το crosstalk. Διάφοροι αλγόριθμοι είναι διαθέσιμοι, οι οποίοι μπορούν να χωριστούν σε δύο κατηγορίες, τους frequency domain και τους time domain. Οι time domain μέθοδοι υλοποιούνται στο πεδίο του χρόνου, ενώ οι frequency domain υλοποιούνται στο πεδίο των συχνοτήτων.

Ο Damaske χρησιμοποίησε για τη δημιουργία του φίλτρου του, τη μέθοδο 90-degree (Damaske, 1971). Σ' αυτή τη μέθοδο χρησιμοποιείται ένα stereo ζεύγος ηχείων, τα οποία είναι τοποθετημένα στις ± 36 μοίρες και αναπαράγουν το ίδιο σήμα, το οποίο είναι λευκός θόρυβος. Πραγματοποιήθηκαν υποκειμενικά tests στα οποία ο εκάστοτε ακροατής είχε τη δυνατότητα να ελέγξει τη στάθμη και την καθυστέρηση του δεξιού ηχείου, μέχρι η κατεύθυνση του ήχου να είναι εμφανώς από το αριστερό ηχείο. Αυτές οι διαμορφώσεις στο σήμα γίνονταν και για το έτερο ηχείο αντίστοιχα με σκοπό τη συμμετρική ακύρωση του crosstalk. Τα tests που πραγματοποιήθηκαν

έδειξαν ότι ο ηχητικός εντοπισμός στο οριζόντιο και στο κάθετο επίπεδο μπορεί να επιτευχθεί με την προϋπόθεση της ακρόασης σε ανηχοϊκό θάλαμο. Σε αντηχητικά πεδία ο ηχητικός εντοπισμός της εικονικής πηγής δίνει την αίσθηση στον ακροατή ότι η θέση της ηχητικής πηγής δεν είναι ξεκάθαρη, είναι θολή.

Ο Sakamoto μίλησε για ένα συμμετρικό σύστημα ακύρωσης του crosstalk, το οποίο αποτελείται από δύο φίλτρα. Το πρώτο φίλτρο το ονομάζει common filter. Αυτό το φίλτρο πραγματοποιεί την ισοστάθμιση του συστήματος. Το δεύτερο φίλτρο το ονομάζει ratio filter. Το φίλτρο αυτό ακυρώνει το crosstalk των ηχείων (Sakamoto, 1981). Πειράματα που πραγματοποιήθηκαν σε ακροατές, δηλαδή πειράματα τα οποία είναι υποκειμενικά, δείχνουν ότι η χρήση των common filters είναι απαραίτητη για τη δημιουργία εικονικών ηχητικών πηγών. Πραγματοποιήθηκαν επίσης πειράματα σε μη ανηχοϊκά περιβάλλοντα, τα οποία έδειξαν ότι το φαινόμενο δεν λειτουργεί τόσο καλά, όσο σε ανηχοϊκό περιβάλλον. Οι Cooper και Bauck, εφάρμοσαν φίλτρα ελάχιστης φάσης (minimum phase filters), χρησιμοποιώντας το άθροισμα και τη διαφορά της αριστερής και της δεξιάς εισόδου (Cooper & Bauck, 1989). Οι Koring και Wightman πραγματοποίησαν υποκειμενικά tests και παρατήρησαν την απόδοση του ανάστροφου φίλτρου ερευνώντας τη χρήση ανάστροφων φίλτρων, τα οποία δεν ήταν ξεχωριστά για κάθε ακροατή και αξιολογώντας τη χρήση ενός κλασικού μοντέλου αυτού. Η αποδοτικότητα αυτού του φιλτραρίσματος αξιολογήθηκε μετά από ακροάσεις σε μη ανηχοϊκό περιβάλλον ακρόασης (Koring & Wightman, 1992).

Οι Miyoshi και Kaneda ανέπτυξαν μία μέθοδο η οποία στηρίζεται στην αρχή του ανάστροφου θεωρήματος πολλαπλής εισόδου - εξόδου. Σ' αυτή τη μέθοδο η αναστροφή πραγματοποιείται μέσω του συνδυασμού FIR φίλτρων με την πρόσθεση ενός επιπλέον καναλιού ηχητικού σήματος το οποίο αναπαράγεται από τα ηχεία του συστήματος. Οι συντελεστές (coefficients) αυτών των FIR φίλτρων μπορούν να υπολογιστούν λύνοντας ένα γραμμικό σύστημα. Το σύστημα στο οποίο θα εφαρμοστεί αυτή η μέθοδος δεν πρέπει να είναι ελάχιστης φάσης (minimum phase), το αποτέλεσμα της αναστροφής όμως πρέπει να είναι αιτιατό (causal) και σταθερό (stable) (Miyoshi & Kaneda, 1988).

Ο Nelson σύγκρισε τη μέθοδο των Miyoshi και Kaneda με τον αλγόριθμο του ME-LMS (Multiple Error LMS). Στη μέθοδο του ME-LMS, τα FIR φίλτρα σχεδιάστηκαν ξεχωριστά. Το πραγματικό delay στο ανεστραμμένο σύστημα εξαρτάται από την χρήση του modeling delay. Επιπλέον, ο Nelson απέδειξε ότι με τη

σωστή επιλογή του modeling delay και του αριθμού των συντελεστών (coefficients) του ανάστροφου φίλτρου, ο σχεδιασμός των φίλτρων χρησιμοποιώντας MINT ή ME-LMS έχει πανομοιότυπα αποτελέσματα με τη μέθοδο των Miyoshi και Kaneda, (Nelson, 1995). Ο Nelson πρότεινε η αναστροφή να επιτυγχάνεται με τη χρησιμοποίηση προσαρμοστικών φίλτρων (adaptive filters). Αυτή η προσέγγιση είναι μία γενικευμένη εκδοχή του συστήματος ακύρωσης του crosstalk των ηχείων των Atal και Schroeder (Atal & Schroeder, 1962). Μπορεί να εφαρμοστεί σε τεχνητό κεφάλι ή coincident μικρόφωνα και σε μικρόφωνα ηχογράφησης ηχητικών πεδίων (Nelson, 1996). Μπορεί δηλαδή να εφαρμοστεί σε όλους τους τύπους ηχογράφησης. Αυτό το σύστημα εξαρτάται κατά πολύ από την επιλογή του modeling delay. Ο Nelson έδειξε επίσης ότι τα καλύτερα αποτελέσματα, όσον αφορά την ισοστάθμιση της απόκρισης του συστήματος, βγήκαν με χρήση modeling delay μήκους ίσου με το μισό του μήκους του φίλτρου. Η μελέτη αυτή του Nelson αφορά προβλήματα αναστροφής, χωρίς να εξειδικεύει σε συστήματα που εξαρτώνται από τα HRTFs.

Ο Kirkeby πρότεινε μία μέθοδο γρήγορης αποσυνέλιξης η οποία βασίζεται στον γρήγορο μετασχηματισμό Fourier (Fast Fourier Transform - FFT). Χρησιμοποιώντας αυτή τη μέθοδο κατάφερε να κατασκευάσει FIR φίλτρα για πολυκάναλα συστήματα. Ο FFT βρίσκει την ακριβή λύση των ελαχίστων τετραγώνων, η οποία δεν είναι πραγματοποιήσιμη στην πράξη. Η κανονικοποίηση (Regularization), εισήχθη στη διαδικασία της αναστροφής για να καθυστερήσει τη λειτουργία του ιδανικού ανάστροφου φίλτρου, ούτως ώστε να διασφαλιστεί ότι η ύπαρξη του circular convolution effect δεν θα είναι αρκετή για να προκαλέσει προβλήματα στο σήμα. Η κανονικοποίηση (Regularization) χρησιμοποιείται για να αποδώσει την απόκριση συχνότητας σε έναν επαρκή αριθμό διακριτών σημείων για τη διασφάλιση της ακριβούς αναπαραγωγής των επιθυμητών σημάτων. (Kirkeby, 1998a). Ο Kirkeby πρότεινε επίσης ένα σύστημα αναπαραγωγής εικονικών ηχητικών πηγών, το οποίο αποτελείται από δύο ηχεία τοποθετημένα κοντά μεταξύ τους. Τα δύο ηχεία είναι τοποθετημένα μπροστά από τον ακροατή με γωνία ± 10 μοιρών όπως κοιτάει ο ακροατής. Αυτή η διάταξη των ηχείων είναι το stereo δίπολο (stereo dipole), που αναφέρθηκε προηγουμένως. Το αναπαραγόμενο ηχητικό πεδίο που δημιουργεί το stereo δίπολο σε σύγκριση με διάφορες άλλες γωνίες τοποθέτησης των ηχείων είναι το πιο αποτελεσματικό στην ακύρωση του crosstalk. Ο Kirkeby παρατήρησε ότι η αναστροφή πειραματικά μετρημένων δεδομένων προκαλεί μια

ενίσχυση των συχνοτήτων που βρίσκονται ακριβώς κάτω από τη συχνότητα του Nyquist (Kirkeby & Nelson, 1999). Ο Kirkeby μελετά μία μέθοδο σχεδιασμού ψηφιακού φίλτρου, η οποία επιτρέπει τον έλεγχο της τιμής της κανονικοποίησης ανά συχνοτική περιοχή. Αυτό επιτρέπει η απόδοση των ιδανικών FIR φίλτρων να είναι η βέλτιστη δυνατή στις ορισμένες συχνοτικές μπάντες και είναι υπολογισμένα λύνοντας ένα σύστημα γραμμικής ισότητας και επιλέγοντας τις κατάλληλες τιμές για τον παράγοντα κανονικοποίησης και το μήκος του φίλτρου. Η μέθοδος χρησιμεύει στην ακύρωση της απόκρισης των ηχείων, της απόκρισης του χώρου και των HRTFs, χρησιμοποιώντας μία μέτρηση κρουστικής απόκρισης.

Οι Nelson και Rose πραγματοποίησαν έρευνα, της οποίας τα αποτελέσματα δείχνουν πως σχετίζονται τα σφάλματα του προβλήματος αντιστροφής με τη διεύρυνση του φάσματος, στο οποίο πραγματοποιείται ακύρωση του crosstalk και με την απόκριση του συστήματος αναπαραγωγής στο πεδίο του χρόνου. Οι προσομοιώσεις που πραγματοποιήθηκαν χρησιμοποιώντας τη μέθοδο Singular Value Decomposition (SVD), στον πίνακα των συναρτήσεων μεταφοράς σε ελεύθερο πεδίο, έδειξαν τη σχέση του αναπαραγόμενου από τα ηχεία σήματος, με την ηχητική πίεση που δημιουργείται στις θέσεις όπου έχουν τοποθετηθεί τα δύο μικρόφωνα, τα οποία αντιπροσωπεύουν τα αυτιά του ακροατή. Το πείραμα έδειξε ότι στις συχνοτικές περιοχές, στις οποίες οι τιμές του φίλτρου είναι επιλεγμένες σωστά, η ακύρωση του crosstalk είναι αποτελεσματικότερη και η απόκριση συχνότητας στο πεδίο του χρόνου έχει πεπερασμένη διάρκεια.

Ο Μουχτάρης εξέτασε δύο μεθόδους δημιουργίας ανάστροφων φίλτρων για την ακύρωση του crosstalk των ηχείων. Συνεχίζοντας την εργασία του Nelson (Nelson, 1996), εξέτασε τις μεθόδους: Least Mean Square (LMS) και LMS διασταυρωμένη (transversal). Η εργασία του Μουχτάρη, περιλαμβάνει επεξεργασία με σκοπό τη δημιουργία χωρικού ηχητικού σήματος (spatial audio) βασισμένο στα HRTFs. Ο απώτερος σκοπός της έρευνας του, είναι η επίτευξη προσαρμοζόμενης μεθόδου για διαδραστικές εφαρμογές, δηλαδή real time rendering. Αποδεικνύεται ότι με τη χρήση αυτών των μεθόδων μπορεί να επιτευχθεί ακύρωση του crosstalk της τάξης των 30 dB, όμως η χρήση της LMS διασταυρωμένης (transversal) προσφέρει ένα πιο σημαντικό πλεονέκτημα ως προς της υπολογιστική αποτελεσματικότητα. Μία παραγοντοποίηση του αντίστροφου πίνακα C προτείνεται επίσης σε δύο όρους. Ο ένας όρος χρησιμοποιείται για την ακύρωση του crosstalk. Ο δεύτερος όρος

χρησιμοποιείται για την ισοστάθμιση του απευθείας σήματος, το λεγόμενο ipsilateral (Mouchtaris et al., 2000).

Ο Kim παρουσίασε ένα διαφορετικό σύστημα για την αναπαραγωγή binaural σήματος από δύο ηχεία. Ένα σύστημα αποτελούμενο από δύο ηχεία και δύο μικρόφωνα (Kim & Wang, 2003). Η μέθοδος του Kim βασίζεται στη συστηματική αποσυνέλιξη στο πεδίο του χρόνου με τη χρήση φίλτρων Wiener, τα οποία είναι από τη φύση τους αιτιατά και σταθερά. Αυτό σημαίνει ότι δεν χρειάζεται περαιτέρω επεξεργασία σταθεροποίησης, σε αντίθεση με τις μεθόδους στο πεδίο των συχνοτήτων. Τα σφάλματα που προκύπτουν χωρίζονται σε τρεις κατηγορίες. Τα σφάλματα μοντελοποίησης (modeling errors), τα οποία προκύπτουν κατά τη διαδικασία της μοντελοποίησης. Τα αντίστροφα σφάλματα, τα οποία προκύπτουν κατά τη διαδικασία της αποσυνέλιξης και τα σφάλματα επαναληψιμότητας, τα οποία προκύπτουν λόγω των χρονομεταβλητών χαρακτηριστικών. Αυτά τα σφάλματα χρησιμοποιούνται για τη μέτρηση της απόδοσης του συστήματος αναπαραγωγής. Πειράματα τα οποία συμπεριλάμβαναν όχι μόνο binaural ηχητικά σήματα, αλλά και μονοφωνικά, πραγματοποιήθηκαν σε ανηχοϊκό περιβάλλον και έδειξαν ότι το αποτέλεσμα της αναπαραγωγής βελτιώθηκε, για το συγκεκριμένο σύστημα, όχι εξαιτίας του σφάλματος αναστροφής, αλλά εξαιτίας του modeling σφάλματος.

3. Ανάλυση των μεθόδων δημιουργίας ανάστροφου φίλτρου για την ακύρωση του crosstalk και της θέσης των ηχείων.

3.1 Εισαγωγή

Στο τρίτο κεφάλαιο θα αναλυθούν οι μέθοδοι δημιουργίας ανάστροφου φίλτρου για την ακύρωση του crosstalk και της θέσης των ηχείων, που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία. Οι δύο αυτές μέθοδοι είναι η γρήγορη αποσυνέλιξη με κανονικοποίηση, η οποία πραγματοποιείται στο πεδίο των συχνοτήτων και η μέθοδος των ελαχίστων μέσων τετραγώνων, η οποία πραγματοποιείται στο πεδίο του χρόνου.

3.2 Ανάλυση συστήματος ακύρωσης του crosstalk με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση (Fast deconvolution using regularization).

3.2.1 Εισαγωγή.

Στο κεφάλαιο που ακολουθεί θα αναλυθεί η μέθοδος της γρήγορης αποσυνέλιξης με κανονικοποίηση (Fast Deconvolution Using Regularization) για τη δημιουργία ανάστροφου φίλτρου, με σκοπό την ακύρωση του crosstalk και της φυσικής θέσης των ηχείων σε σχέση με το κεφάλι του ακροατή. Η μέθοδος χρησιμοποιείται για τη δημιουργία συστήματος ακύρωσης crosstalk. Αυτά τα συστήματα χρησιμοποιούνται για την αναπαραγωγή binaural σημάτων από δύο ηχεία.

3.2.2 Αναστροφή συστήματος.

Έχουμε ένα σύστημα δύο καναλιών το οποίο θέλουμε να αναστρέψουμε. Το σύστημα αυτό αποτελείται από δύο ηχεία τοποθετημένα σε συγκεκριμένες θέσεις σε σχέση με το κεφάλι του ακροατή, είναι δηλαδή το ηλεκτροακουστικό σύστημα το οποίο χρησιμοποιείται για την αναπαραγωγή του binaural σήματος.

Η κρουστική απόκριση του συστήματος είναι $\alpha(n)$ και περιέχει N παράγοντες. Το σήμα το οποίο θέλουμε να αναπαραχθεί από τα ηχεία και να φτάσει στα αυτιά του ακροατή είναι το $d(n)$. Το επιθυμητό αυτό σήμα είναι το binaural σήμα. Το πρόβλημα είναι ότι το σήμα το οποίο τελικά φτάνει στα αυτιά του ακροατή

δεν είναι το $d(n)$ αλλά το $\hat{d}(n)$. Το $\hat{d}(n)$ είναι μία παραλλαγή του επιθυμητού σήματος $d(n)$. Αυτό συμβαίνει εξαιτίας της ύπαρξης του ηλεκτροακουστικού συστήματος, το οποίο είναι λογικό να επηρεάζει το σήμα. Το $\hat{d}(n)$ ορίζεται ως εξής.

$$\hat{d}(n) = \alpha(n) * d(n) \quad \text{Εξίσωση 3.1}$$

Ο μετασχηματισμός-Z του $\alpha(n)$ είναι:

$$C(z) = \sum_{n=0}^{N-1} \alpha(n) z^{-n} = \alpha(0) + \alpha(1)z^{-1} + \dots + \alpha(N-1)z^{-(N-1)} \quad \text{Εξίσωση 3.2}$$

όπου N είναι ένας θετικός ακέραιος. Ο μετασχηματισμός-z της εξίσωσης 3.1, μας δίνει:

$$\hat{D}(z) = C(z)D(z) \quad \text{Εξίσωση 3.3}$$

Παρατηρώντας την εξίσωση 3.3, συμπεραίνουμε ότι για να είναι το $\hat{D}(z)$ ίσο με το $D(z)$ πρέπει το $C(z)$ να γίνει 1. Για να γίνει αυτό θα εισαχθεί στην εξίσωση ένας νέος όρος.

$$H(z) = \frac{1}{C(z)} \quad \text{Εξίσωση 3.4}$$

Με την εισαγωγή του παραπάνω όρου στο σύστημα, η έξοδος του ηλεκτροακουστικού συστήματος γίνεται:

$$\hat{D}(z) = H(z)C(z)D(z) = D(z) \quad \text{Εξίσωση 3.5}$$

Άρα το ιδανικό ανάστροφο φίλτρο είναι:

$$H(z) = C(z)^{-1} \quad \text{Εξίσωση 3.6}$$

Άρα:

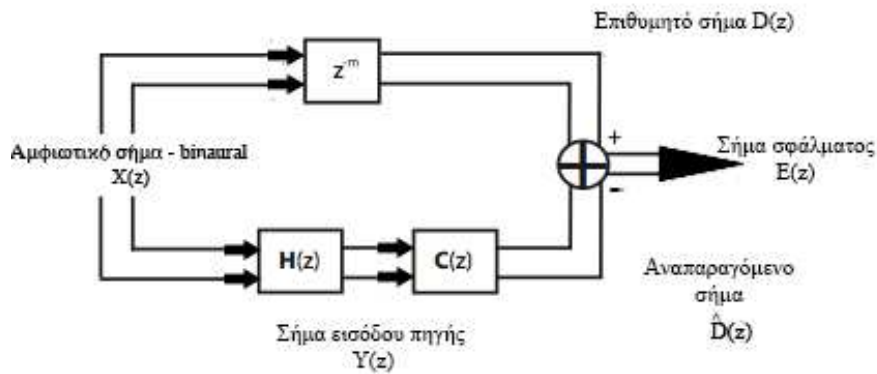
$$H(z) = \frac{1}{C(0) + C(1)z^{-1} + \dots + C(N-1)z^{-(N-1)}} \quad \text{Εξίσωση 3.7}$$

Με το μετασχηματισμό z μπορούμε να ελέγξουμε αν το φίλτρο είναι αιτιατό (causal). Για να είναι ένα σύστημα αιτιατό πρέπει κάθε έξοδος του συστήματος να εξαρτάται από εισόδους που βρίσκονται χρονικά στο παρόν και στο παρελθόν. Δηλαδή για $C(n)$ πρέπει $n \geq 0$. Για να διασφαλιστεί η αιτιότητα του φίλτρου, εισάγεται μία καθυστέρηση. Επίσης, αν το σύστημα μας δίνει φραγμένη έξοδο, όταν του εισάγουμε φραγμένη είσοδο, τότε είναι ευσταθές (stable). Όταν δηλαδή του εισάγουμε μία είσοδο με πεπερασμένο πλάτος και μας δώσει έξοδο με επίσης πεπερασμένο πλάτος.

Το block διάγραμμα της εικόνας 4 μας δείχνει την αναστροφή ενός συστήματος δύο καναλιών. Σύμφωνα με τον Kirkeby, (Kirkeby 1998), τα σήματα που υπάρχουν στο σύστημα είναι τα παρακάτω (εξίσωση 3.8)

$$\begin{aligned} x(n) &= [x_1(n), x_2(n)]^T \\ y(n) &= [y_1(n), y_2(n)]^T \\ d(n) &= [d_1(n), d_2(n)]^T \\ \hat{d}(n) &= [\hat{d}_1(n), \hat{d}_2(n)]^T \\ e(n) &= [e_1(n), e_2(n)]^T \end{aligned}$$

Εξίσωση 3.8



Εικόνα 4: Block διάγραμμα αναστροφής συστήματος δύο καναλιών (Manias, 2008).

Το διάνυσμα $X(z)$ αντιπροσωπεύει το binaural σήμα, το διάνυσμα $Y(z)$ αντιπροσωπεύει το ζεύγος σημάτων εισόδου πηγής, το διάνυσμα $D(z)$ αντιπροσωπεύει το επιθυμητό ζεύγος σημάτων, το διάνυσμα $\hat{D}(z)$ αντιπροσωπεύει το ζεύγος σημάτων τα οποία φτάνουν στα αυτιά του ακροατή, και το διάνυσμα $E(z)$ αντιπροσωπεύει το ζεύγος σημάτων σφάλματος, το οποίο είναι ίσο με τη διαφορά του $\hat{D}(z)$ με το $D(z)$. Τα παραπάνω σήματα είναι ουσιαστικά πίνακες τιμών.

Εισάγεται επίσης ο πίνακας $C(z)$ ο οποίος έχει διαστάσεις 2×2 και περιέχει τα HRTFs τα οποία περιέχουν πληροφορίες σχετικές με την ακουστική πίεση στα αυτιά του ακροατή, η οποία εξαρτάται από το σήμα εισόδου στην πηγή. Ο πίνακας τιμών $C(z)$ είναι της μορφής (εξίσωση 3.9).

Επιπλέον εισάγεται ο πίνακας $H(z)$ με διαστάσεις 2×2 , ο οποίος περιέχει τις συναρτήσεις μεταφοράς που χρησιμοποιούνται για την δημιουργία της κατάλληλης εισόδου στην πηγή και είναι η αναστροφή του πίνακα $C(z)$.

$$C(z) = \begin{bmatrix} C_{11}(z) & C_{1S}(z) \\ \vdots & \vdots \\ C_{R1}(z) & \cdots & C_{RS}(z) \end{bmatrix} \quad \text{Εξίσωση 3.9}$$

Κάθε στοιχείο $C_{RS}(z)$ του πίνακα $C(z)$, είναι ο μετασχηματισμός-z της αντίστοιχης κρουστικής c_{RS} , όπου S είναι το σήμα εισόδου πηγής και όπου R είναι ο δέκτης.

$$C_{RS}(z) = \sum_{n=0}^{N_c-1} c_{RS}(n) z^{-n} \quad \text{Εξίσωση 3.10}$$

Έστω ότι το $c(n)$ είναι μία ακολουθία πεπερασμένου μήκους με N_c παράγοντες. Επίσης υποθέτουμε ότι, το ιδανικό φίλτρο $h(n)$ είναι ένα FIR φίλτρο με N_h παράγοντες. Είναι απαραίτητο να εφαρμοστεί στο επιθυμητό ζεύγος σημάτων $d(n)$, μία καθυστέρηση με μέγεθος m δείγματα, ούτως ώστε να αποκατασταθούν τα στοιχεία του $c(n)$, τα οποία δεν είναι ελάχιστης φάσης (minimum phase). Αυτή η καθυστέρηση ονομάζεται “modeling delay”.

Αν $D(z)$ είναι ο μετασχηματισμός-z που αντιστοιχεί στο ζεύγος επιθυμητών σημάτων $d(n)$ και αν $\hat{D}(z)$ είναι ο μετασχηματισμός-z που αντιστοιχεί στο ζεύγος αναπαραγόμενων σημάτων $\hat{d}(n)$, τότε ο μετασχηματισμός-z των σημάτων σφάλματος είναι:

$$E(z) = D(z) - \hat{D}(z) \quad \text{Εξίσωση 3.11}$$

Όταν το $D(z)$ είναι όμοιο με το $X(z)$, όπου $X(z)$ είναι ο μετασχηματισμός-z του binaural σήματος, το $E(z)$ γίνεται:

$$\begin{aligned}
E(z) &= X(z) - \hat{D}(z) \\
&= X(z) - C(z)H(z)X(z) \\
&= [I_{2 \times 2} - C(z)H(z)]X(z)
\end{aligned}
\tag{Εξίσωση 3.12}$$

Όπου $I_{2 \times 2}$ είναι ένας πίνακας ταυτότητας (identity matrix). Από την εξίσωση 3.12, προκύπτει ότι το $E(z)$ γίνεται μηδέν, όταν το $I_{2 \times 2} - C(z)H(z)$ γίνεται μηδέν:

$$I_{2 \times 2} - C(z)H(z) = 0 \tag{Εξίσωση 3.13}$$

Άρα

$$C(z)H(z) = I_{2 \times 2} \tag{Εξίσωση 3.14}$$

Άρα

$$H(z) = [C(z)]^{-1} \tag{Εξίσωση 3.15}$$

Το αποτέλεσμα του πολλαπλασιασμού των πινάκων $C(z)$ και $H(z)$ περιγράφει την ακύρωση του crosstalk και την ισοστάθμιση (equalization). Αν ονομάσουμε αυτό το αποτέλεσμα X_{eq} συμπεραίνουμε ότι:

Στο πεδίο των συχνοτήτων (frequency domain)

$$X_{eq} = C(z)H(z) = \begin{bmatrix} C_{11}H_{11} + C_{12}H_{21} & C_{11}H_{12} + C_{12}H_{22} \\ C_{21}H_{11} + C_{22}H_{21} & C_{21}H_{12} + C_{22}H_{22} \end{bmatrix} \tag{Εξίσωση 3.16}$$

Στο πεδίο του χρόνου (time domain)

$$x_{eq} = c(n) * h(n) = \begin{bmatrix} c_{11} * h_{11} + c_{12} * h_{21} & c_{11} * h_{12} + c_{12} * h_{22} \\ c_{21} * h_{11} + c_{22} * h_{21} & c_{21} * h_{12} + c_{22} * h_{22} \end{bmatrix} \tag{Εξίσωση 3.17}$$

Άρα ένα σύστημα αναπαραγωγής εικονικών ηχητικών πηγών εξαρτάται από την υλοποίηση του ανάστροφου φίλτρου, δηλαδή του πίνακα H .

3.2.3 Κανονικοποίηση (Regularization)

Η παράμετρος του παράγοντα κανονικοποίησης εισήχθη στα προβλήματα δημιουργίας ανάστροφου φίλτρου, για τη δημιουργία συστήματος αναπαραγωγής

εικονικών ηχητικών πηγών, από τον Kirkeby (Kirkeby, 1998a). Ο Kirkeby στην προσπάθειά του, για την αντιμετώπιση του προβλήματος του circular convolution effect, το οποίο προκαλείται κατά τη διάρκεια της αναστροφής στο πεδίο των συχνοτήτων (frequency domain) εισήγαγε την κανονικοποίηση. Με την εισαγωγή της κανονικοποίησης μειώνεται η διάρκεια της επίδρασης του ιδανικού φίλτρου σε $N_h/2$ (Kirkeby, 1998a). Η αντιμετώπιση του circular convolution effect μπορεί να επιτευχθεί εισάγοντας μια «λειτουργία κόστους (cost function)», $J(e^{j\omega})$ όπως αναφέρει ο Kirkeby (Kirkeby, 1998a).

$$J(e^{j\omega}) = e^H(e^{j\omega})e(e^{j\omega}) + \beta y^H(e^{j\omega})y(e^{j\omega}) \quad \text{Εξίσωση 3.18}$$

Όπου β , ο παράγοντας κανονικοποίησης (regularization parameter). Ο όρος $\beta y^H(e^{j\omega})y(e^{j\omega})$ είναι ένα μέτρο της συνολικής ισχύος εισόδου στις πηγές και ο όρος $e^H(e^{j\omega})e(e^{j\omega})$ είναι το σφάλμα απόδοσης, το οποίο είναι ένα μέτρο του πόσο καλά αναπαράγονται τα επιθυμητά σήματα στα ηχεία (Kirkeby, 1998).

Ο Nelson (Nelson, 1994) έδειξε ότι για να επιτευχθεί η ελάχιστη δυνατή τιμή του $J(e^{j\omega})$, πρέπει η συχνοτική απόκριση του ανάστροφου φίλτρου να είναι:

$$H(e^{j\omega}) = \frac{C^H(e^{j\omega})}{C^H(e^{j\omega})C(e^{j\omega}) + \beta I} \quad \text{Εξίσωση 3.19}$$

Όταν η παράμετρος κανονικοποίησης έχει την τιμή μηδέν, ο πίνακας του ανάστροφου φίλτρου είναι ακριβώς $H[\mathbf{z}] = \mathbf{C}^{-1}[\mathbf{z}]$. Για θετικές τιμές του β γίνεται μείωση της ακρίβειας της αναστροφής και συνολικής ισχύος εισόδου στις πηγές. Όταν η τιμή του β είναι πολύ μικρή, δημιουργούνται στη συχνοτική απόκριση του φίλτρου πολλά βυθίσματα και κορυφές (Kirkeby, 1998a). Όταν η τιμή του β είναι πολύ μεγάλη, η αποσυνέλιξη υστερεί σε απόδοση (Kirkeby, 1998a). Η τιμή του β , συνεπώς πρέπει να επιλεγεί μετά από δοκιμές διαφόρων τιμών για την τελική επιλογή της τιμής, που θα χρησιμοποιηθεί για τη δημιουργία του φίλτρου.

3.2.4 Ορισμός του πίνακα του ανάστροφου φίλτρου στο πεδίο των συχνοτήτων.

Η μέθοδος του Kirkeby για τον ορισμό του πίνακα ανάστροφου φίλτρου:

$$H_m(z) \square \frac{1}{C(z)} z^{-\Delta} \quad \text{Εξίσωση 3.20}$$

Όπου $z^{-\Delta}$ είναι η καθυστέρηση (modeling delay) μήκους Δ δειγμάτων, για να ξεπεραστεί η περίπτωση στην οποία ο $C(z)$ δεν είναι ελάχιστης φάσης. Η μέθοδος ακολουθεί τα παρακάτω βήματα:

- Επιλογή του μήκους N_h του ανάστροφου φίλτρου $H_m(z)$. Μία πρώτη προσέγγιση του μήκους του ανάστροφου φίλτρου, είναι σύμφωνα με τον Kirkeby η εξής: $4SN_c$ ο όρος S αντιπροσωπεύει τον αριθμό των καναλιών του συστήματος, στην προκειμένη περίπτωση είναι δύο. Το N_c είναι το μήκος του $\alpha(n)$.
- Εφαρμογή γρήγορου μετασχηματισμού Fourier (Fast Fourier Transform) αριθμού σημείων N_h στο $\alpha(n)$
- Στη συνέχεια πραγματοποιείται αναστροφή κάθε δείγματος του αποτελέσματος του FFT, δηλαδή κάθε συχνότητας ξεχωριστά. Η υπολογιστική αποτελεσματικότητα του γρήγορου μετασχηματισμού Fourier αυξάνεται επιλέγοντας το N_h να είναι μία δύναμη του 2 (Papadopoulos, 2006). Το N_h πρέπει να είναι αρκετά μεγάλο ώστε να μην χαθεί σημαντική πληροφορία κατά την δειγματοληψία στο πεδίο των συχνοτήτων (Kirkeby1998).
- Υπολογισμός του $H(k) = \frac{1}{C(k)}$, $0 \leq k \leq N_h - 1$
- Εφαρμογή του ανάστροφου γρήγορου μετασχηματισμού Fourier (IFFT), $h_m(n) = IFFT(H(k))$
- Εφαρμογή καθυστέρησης (modeling delay), μήκους $N_h/2$ δειγμάτων, εφαρμόζοντας cyclic shift στο $h_m(n)$ για να πάρουμε τελικώς το $h(n)$.

3.2.5 Επίλογος.

Στο κεφάλαιο αυτό αναλύθηκαν οι αρχές του συστήματος ανάστροφου φιλτραρίσματος ακύρωσης crosstalk με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση (Fast Deconvolution Using Regularization) με σκοπό την αναπαραγωγή εικονικής ηχητικής πηγής. Αναλύθηκαν οι παράμετροι οι οποίοι επηρεάζουν την αναστροφή καθώς και τα βήματα που πρέπει να ακολουθηθούν προκειμένου να επιτευχθεί η ακύρωση του crosstalk.

3.3 Ανάλυση συστήματος ακύρωσης του crosstalk με την τεχνική ελαχίστων μέσων τετραγώνων (Least mean squares).

3.3.1 Εισαγωγή

Στο παρακάτω κεφάλαιο θα αναλύσουμε το θεωρητικό υπόβαθρο και τις ενέργειες που πρέπει να γίνουν έτσι ώστε, να μπορεί να αναπαραχθεί με επιτυχία binaural αρχείο ήχου από stereo ζεύγος ηχείων με την τεχνική των ελαχίστων μέσων τετραγώνων. Όπως αναφέραμε στο κεφάλαιο 1.7 πρέπει να δημιουργηθεί ένα ανάστροφο φίλτρο το οποίο θα ακυρώνει το crosstalk και την θέση των ηχείων. Η θεωρητική ανάλυση αυτού θα γίνει συγκρίνοντας τον πίνακα τιμών του φυσικού συστήματος με τον πίνακα τιμών της stereo αναπαραγωγής.

3.3.2 Θεωρητική ανάλυση

Το ψηφιακό φίλτρο το οποίο πρέπει να δημιουργήσουμε για την επίτευξη του φαινομένου αποτελείται από τρία επιμέρους φίλτρα. Το πρώτο για την ακύρωση της θέσης των ηχείων και το αποτέλεσμα της συνέλιξης του δεύτερου και του τρίτου για την ακύρωση του crosstalk των ηχείων.

Στην φύση ο άνθρωπος προσδιορίζει μία ηχητική πηγή με βάση την διαφορά των δύο σημάτων που φτάνουν στα αυτιά του. Όπως αναφέρεται από τους Mouchtaris, Reveliotis και Kyriakakis, χρησιμοποιώντας πίνακα για την αναπαράσταση του φυσικού συστήματος (το οποίο είναι όμοιο με την binaural αναπαραγωγή από ακουστικά) έχουμε:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_L & 0 \\ 0 & H_R \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} \quad \text{Εξίσωση 3.21}$$

Όπου E_L και E_R είναι το αριστερό και το δεξί αυτί αντίστοιχα, H_L και H_R τα HRTF της γωνίας της πηγής (ή της εικονικής πηγής) για το αριστερό και το δεξί αυτί αντίστοιχα και S η μονοφωνική πηγή (Mouchtaris, et al 2000). Κατά την στερεοφωνική αναπαραγωγή από ηχεία, τα σήματα τα οποία φτάνουν στα αυτιά όπως έχουμε αναφέρει είναι τέσσερα, οπότε ο πίνακας παίρνει την μορφή:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i & H_c \\ H_c & H_i \end{bmatrix} \begin{bmatrix} H_L & 0 \\ 0 & H_R \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} \quad \text{Εξίσωση 3.22}$$

Οι όροι που προστέθηκαν αντιπροσωπεύουν τα HRTF'S των σημάτων που φτάνουν στα αυτιά από την γωνία των ηχείων. Η_i είναι ο ipsilateral όρος και αντιπροσωπεύει το σήμα προέρχεται από το ηχείο το οποίο βρίσκετε στην ίδια πλευρά με το αυτί (π.χ. δεξί ηχείο στο δεξί αυτί), Η_c είναι ο contralateral όρος και αντιπροσωπεύει το σήμα προέρχεται από το ηχείο το οποίο βρίσκετε στην αντίθετη πλευρά από το αυτί (π.χ. δεξί ηχείο στο αριστερό αυτί). Για να προσομοιώσουμε το φυσικό σύστημα πρέπει να γίνει επεξεργασία στο σήμα εισόδου. Συγκεκριμένα το σήμα εισόδου πρέπει να φιλτραριστεί με τον πίνακα που περιέχει τους ipsilateral και contralateral όρους ανεστραμμένο, έτσι ώστε η σχέση να πάρει την μορφή:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i & H_c \\ H_c & H_i \end{bmatrix}^{-1} \begin{bmatrix} H_L & 0 \\ 0 & H_R \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} \quad \text{Εξίσωση 3.23}$$

Λύνοντας την βλέπουμε ότι είναι ίδια με αυτή του φυσικού συστήματος:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i & H_c \\ H_c & H_i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ H_i & -H_c \\ -H_c & H_i \\ 1 - \frac{H_c^2}{H_i^2} \end{bmatrix} \begin{bmatrix} H_L & 0 \\ 0 & H_R \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} \quad \text{Εξίσωση 3.24}$$

Η οποία τελικά γράφεται

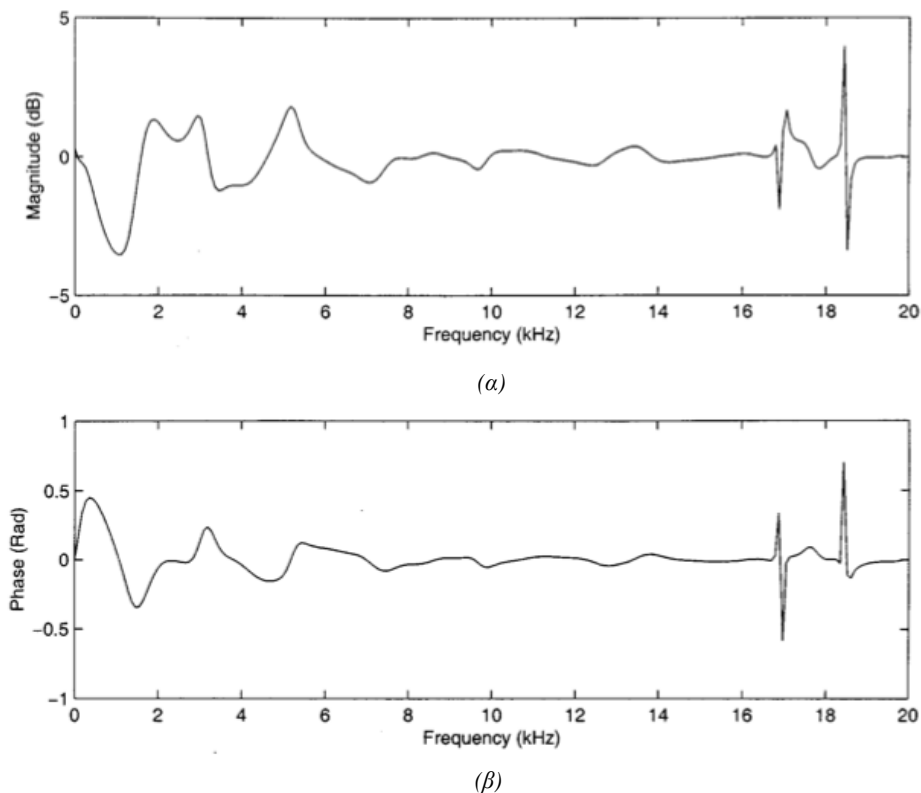
$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i & H_c \\ H_c & H_i \end{bmatrix} \begin{bmatrix} 1 & -\frac{H_c}{H_i} \\ -\frac{H_c}{H_i} & 1 \end{bmatrix} \begin{bmatrix} H_L & 0 \\ 0 & \frac{H_R}{H_i} \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} \quad \text{Εξίσωση 3.25}$$

Υποθέτοντας ότι

$$\frac{1}{1 - \frac{H_c^2}{H_i^2}} \cong 1$$

Εξίσωση 3.26

Η υπόθεση αυτή γίνεται με βάση το γεγονός, ότι η ενέργεια που περιέχεται στον ipsilateral όρο είναι μεγαλύτερη από την ενέργεια που περιέχεται στον contralateral όρο γεγονός το οποίο οφείλετε στην σκιά που προκαλείτε από το κεφάλι. Η παρούσα παραδοχή εξετάστηκε από τους Mouchtaris, Reveliotis και Kyriakakis σε διάγραμμα του πλάτους και της φάσης του όρου (3.26) σε σχέση με το μέγεθος και την φάση ενός ολοπερατού φίλτρου (Mouchtaris, et al 2000). Όπως φαίνεται και στην εικόνα 5.



Εικόνα 5: Διάγραμμα του πλάτους (α) και της φάσης (β), του όρου (3.26) (Mouchtaris, et al., 2000).

Οι όροι H_L / H_i και H_R / H_i αντιστοιχούν στην ακύρωση της θέσης του ηχείου. Ο

πίνακας $\begin{pmatrix} 1 & -\frac{H_c}{H_i} \\ -\frac{H_c}{H_i} & 1 \end{pmatrix}$ αντιστοιχεί στην ακύρωση του crosstalk. Τέλος τα σήματα

τα X_L και X_R τα οποία πρέπει να αναπαραχθούν από τα ηχεία έτσι ώστε η εικονική ηχητική πηγή να βρίσκεται στην επιθυμητή γωνία είναι:

$$\begin{bmatrix} X_L \\ X_R \end{bmatrix} = \begin{bmatrix} \frac{H_L}{H_i} & -\frac{H_c H_R}{H_i H_i} \\ -\frac{H_c H_L}{H_i H_i} & \frac{H_R}{H_i} \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} \quad \text{Εξίσωση 3.27}$$

Η οποία μπορεί να γραφεί και ως:

$$\begin{aligned} X_L &= \begin{bmatrix} \frac{H_L}{H_i} - \frac{H_c H_R}{H_i H_i} \\ \frac{H_c H_L}{H_i H_i} \end{bmatrix} S \\ X_R &= \begin{bmatrix} \frac{H_R}{H_i} - \frac{H_c H_L}{H_i H_i} \\ \frac{H_c H_R}{H_i H_i} \end{bmatrix} S \end{aligned} \quad \text{Εξίσωση 3.28}$$

Αυτό συνεπάγεται ότι τα φίλτρα F_L και F_R για το δεξί και το αριστερό κανάλι πρέπει να είναι της μορφής:

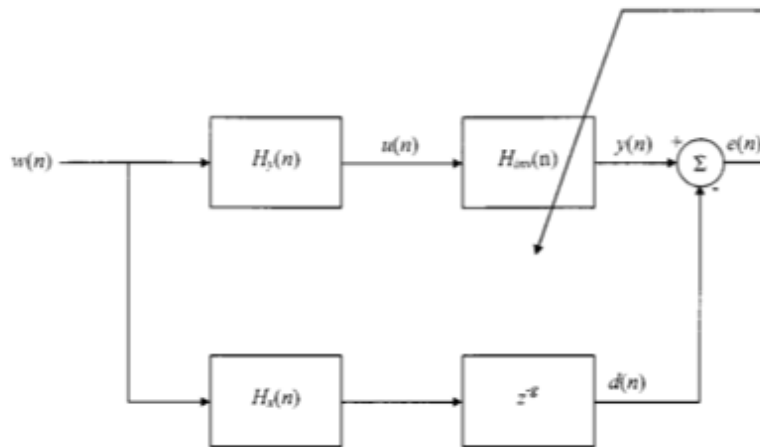
$$\begin{aligned} F_L &= \frac{H_L}{H_i} - \frac{H_c H_R}{H_i H_i} \\ F_R &= \frac{H_R}{H_i} - \frac{H_c H_L}{H_i H_i} \end{aligned} \quad \text{Εξίσωση 3.29}$$

Στην περίπτωση όπου το σήμα στην είσοδο είναι binaural τότε τα HRTF'S της εικονικής πηγής δεν τα περιλαμβάνουμε μέσα στο φίλτρο διότι η πληροφορία αυτή βρίσκεται στο σήμα εισόδου. Έτσι το φίλτρο παίρνει την μορφή:

$$\begin{aligned} F_L &= \frac{1}{H_i} - \frac{H_c}{H_i H_i} \\ F_R &= \frac{1}{H_i} - \frac{H_c}{H_i H_i} \end{aligned} \quad \text{Εξίσωση 3.30}$$

Φίλτρα της μορφής $H_{inv} = H_x / H_y$, όπου, στην συγκεκριμένη περίπτωση το H_x είναι 1 ή H_L ή H_R ή H_c και H_y όπου είναι το H_i μπορούν να κατασκευαστούν με πολλούς τρόπους. Η διαίρεση του κλάσματος στο πεδίο των συχνοτήτων είναι μία από αυτούς. Αυτό όμως θα είχε ως αποτέλεσμα το φίλτρο να είναι ασταθές εξαιτίας του ότι ο όρος H_y δεν είναι ελάχιστης φάσης. Με την μέθοδο των ελαχίστων μέσων τετραγώνων διατηρείται η φυσική πληροφορία των HRTF'S και βρίσκεται μια μη αιτιατή αλλά σταθερή κρουστική απόκριση, η οποία είναι της μορφής H_x / H_y . Προσθέτοντας καθυστέρηση στο φίλτρο τέτοια, ώστε να μην γίνεται αντιληπτή από τον ακροατή και να διατηρεί χαμηλή την υπολογιστική ισχύ, επιτυγχάνεται η

αιτιότητα του. Η σχέση μεταξύ των φίλτρων H_x , H_y και το φίλτρο H_{inv} απεικονίζεται στο παρακάτω μπλοκ διάγραμμα.



Εικόνα 6: Block διάγραμμα του αλγορίθμου (LMS), για τον υπολογισμό του ανάστροφου φίλτρου (Mouchtaris et al., 2000).

Η βασική ιδέα του LMS είναι να προσεγγίσει τα βέλτιστα βάρη των ζητούμενων φίλτρων. Η ενημέρωση των βαρών μέσα από την διαδικασία LMS έχει ως σκοπό τα βάρη να συγκλίνουν σε αυτά του ζητούμενου φίλτρου. Ο αλγόριθμος ξεκινά υποθέτοντας μικρά βάρη στην αρχή (ξεκινώντας από το 0) και σε κάθε βήμα με την εύρεση της κλίσης του μέσου τετραγωνικού σφάλματος τα βάρη ενημερώνονται. Αν η κλίση mse (mean square error) είναι θετική αυτό συνεπάγεται ότι το σφάλμα θα συνεχίσει να αυξάνεται θετικά. Εάν το ίδιο βάρος χρησιμοποιείται για περαιτέρω επαναλήψεις σημαίνει ότι πρέπει να μειώσουμε τα βάρη κατά τον ίδιο τρόπο, εάν η κλίση είναι αρνητική θα πρέπει να αυξηθούν τα βάρη (www.wikipedia.org).

Η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος μεταξύ του $y(n)$ και $d(n)$ έγκειται στον ορισμό του φίλτρου H_{inv} . Αρχικά πρέπει να γίνει αναγνώριση του ιδανικού συστήματος αναπαραγωγής (H_x), έπειτα ακολουθεί αναστροφή του πραγματικού συστήματος αναπαραγωγής (H_y). Η μέθοδος των ελαχίστων μέσων τετραγώνων είναι μία από τις βασικές προσαρμοστικές μεθόδους, όπου πάνω σε αυτή μπορεί να βασιστεί η λύση του παραπάνω προβλήματος. Η επανάληψη προσαρμογής είναι μια διαδικασία η οποία περιλαμβάνει τη χρήση μιας συνάρτησης κόστους η οποία τροφοδοτεί τον αλγόριθμο των ελαχίστων μέσων τετραγώνων προκειμένου να καθοριστεί ο τρόπος τροποποίησης του ανάστροφου φίλτρου.

Συγκεκριμένα το μήκος του φίλτρου H_{inv} για n επαναλήψεις μπορεί να υπολογιστεί με βάση τον τύπο προσαρμογής των βαρών:

$$h_{inv}(n+1) = h_{inv}(n) + \mu u(n) \epsilon(n) \quad \text{Εξίσωση 3.31}$$

Όπου:

$$\epsilon(n) = d(n) - h_{inv}^H(n)u(n) \quad \text{Εξίσωση 3.32}$$

Στην εξίσωση 3.32 το H δηλώνει τον ερμιτιανό συζυγή φορέα του h_{inv} . Η επιθυμητή απόκριση $d(n)$ μπορεί να βρεθεί μέσω του block διαγράμματος να είναι ίση με:

$$d(n) = h_x^H(n)w(n-g) \quad \text{Εξίσωση 3.33}$$

Το $u(n)$ είναι ο πίνακας με μήκος M στον οποίο θα καταχωρηθούν οι τιμές των δειγμάτων του φίλτρου:

$$u(n) = [u(n)u(n-1)\dots u(n-M+1)]^T \quad \text{Εξίσωση 3.34}$$

Όπου M είναι η τάξη του φίλτρου h_{inv} και το μήκος του σήματος εισόδου $w(n)$. Για την είσοδο του συστήματος μπορεί να γίνει τυχαία επιλογή, η ποιο συνηθισμένη διαδικασία όμως, για προβλήματα αναγνώρισης συστήματος, είναι η επιλογή λευκού θορύβου στην είσοδο. Ο λευκός θόρυβος εξαιτίας της ιδιότητας του να διαθέτει ίση ενέργεια ανά οκτάβα έχει ως αποτέλεσμα όλες οι συχνότητες του φάσματος κατά την διαδικασία προσαρμογής του φίλτρου να υπολογίζονται ισόποσα. Το μήκος του φίλτρου M καθώς και η καθυστέρηση g , μπορούν να επιλεχθούν με βάση την ελαχιστοποίηση του τετραγωνικού σφάλματος. Στον αλγόριθμο γίνεται χρήση μίας παραλλαγής του LMS, ο αλγόριθμος normalized LMS, όπου γίνεται σταδιακή μείωση του μεγέθους του βήματος όπου οδηγεί σε πιο γρήγορα αποτελέσματα με μικρότερες αποκλίσεις. Σε κάθε επανάληψη το μέγεθος του βήματος μ αλλάζει σύμφωνα με την σχέση ενημέρωσης:

$$\mu(n) = \frac{\beta}{\alpha + \|u(n)\|^2} \quad \text{Εξίσωση 3.35}$$

Το β είναι μία θετική σταθερά συνήθως μικρότερη από δύο και το α είναι μία μικρή θετική σταθερά.

Το αποτέλεσμα αυτής της μεθόδου θα είναι το φίλτρο της μορφής h_{inv} όπου στο πεδίο των συχνοτήτων είναι ίσο με H_x/H_y . Στην περίπτωση που το φίλτρο κατασκευαστεί με προϋπόθεση να φιλτράρει binaural πληροφορία το h_{inv} θα είναι της μορφής $1/H_y$. Το αποτέλεσμα και στις δύο περιπτώσεις είναι ένα FIR φίλτρο.

Στο φίλτρο της μορφής $1/H_y$ το οποίο κατασκευάστηκε την μέθοδο του LMS όπως αναφέραμε παραπάνω χρησιμοποιήσαμε μία καθυστέρηση δειγμάτων με σκοπό να ικανοποιήσουμε την αιτιότητα του φίλτρου. Οι συντελεστές αυτών των φίλτρων FIR σχεδιάστηκαν χρησιμοποιώντας το Matlab. Οι καθυστερήσεις και τα μήκη για τα φίλτρα που χρησιμοποιήθηκαν για να βελτιστοποιηθεί και να επιτευχθεί η μέγιστη αναλογία ισχύος σήματος προς το σφάλμα SER (signal to error ratio) έγινε στον πεδίο του χρόνου ανάμεσα στα φίλτρα H_{inv} , H_y τα οποία συνθέτουν ένα φίλτρο καταρράκτη, και το H_x . Σε αυτή την περίπτωση το SER ορίζεται από την σχέση:

$$\frac{\sum_{k=1}^N h_x^2(k)}{\sum_{k=1}^N (h_x(k) - h_{ca}(k))^2} \quad \text{Εξίσωση 3.36}$$

Όπου h_{ca} είναι η κρουστική απόκριση του φίλτρου καταρράκτη. Είναι σημαντικό να υπολογιστεί το σφάλμα στο πεδίου χρόνου, ούτως ώστε να πραγματοποιηθεί όσο το δυνατό καλύτερη προσέγγιση τόσο στο μέγεθος όσο και στην απόκριση της φάσης.

Η παραπάνω μέθοδος λειτουργεί με επιτυχία χρησιμοποιώντας HRTF (τις γωνίας του ηχείου για τις 30 μοίρες) μήκους 128 δειγμάτων. Η λήψη των HRTF αυτών προέρχεται από dummy head KEMAR με διαστάσεις κεφαλής και σώματος ίσες με τον μέσο ανθρώπινο όρο.

3.3.3 Ακύρωση crosstalk.

Ξαναγράφοντας εξίσωση του φίλτρου 3.30 και έχοντας προσθέσει τις καθυστερήσεις στα επιμέρους φίλτρα έχουμε

$$\begin{aligned} f_l &= h_{li} - h_{ci} * h_{ri} \\ f_r &= h_{ri} - h_{ci} * h_{li} \end{aligned} \quad \text{Εξίσωση 3.37}$$

Όπου * δηλώνει συνέλιξη.

Το αρχείο ήχου το οποίο θα εφαρμοστεί στην είσοδο του φίλτρου είναι binaural οπότε τα επιμέρους φίλτρα θα πρέπει να είναι της μορφής $h_{ii} = h_{ri} = 1/H_i$ και $h_{ci} = H_c/H_i$. Η καθυστέρηση $d1$ του πρώτου όρου πρέπει να είναι ίση σε δείγματα με το άθροισμα των άλλων δυο ($d2, d3$), δηλαδή $d1=d2+d3$. Το άθροισμα $n2+n3$ των συνολικών δειγμάτων του δεύτερου και του τρίτου όρου, πρέπει να είναι μεγαλύτερο κατά ένα δείγμα από του πρώτου έτσι ώστε όταν γίνει η μεταξύ τους συνέλιξη με βάση την θεωρία:

$n2 * n3 = (n2 + n3) - 1 = n1$. Με αυτό τον τρόπο έχουμε πετύχει το μέγεθος των πινάκων να είναι ίσο έτσι ώστε να μπορούν να γίνουν οι μεταξύ τους πράξεις και η καθυστέρηση των σημάτων να είναι ίση.

Εφόσον το σήμα στην είσοδο όπως είπαμε παραπάνω είναι binaural η (3.25) ξαναγράφεται:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i & H_c \\ H_c & H_i \end{bmatrix} \begin{bmatrix} 1 & -\frac{H_c}{H_i} \\ -\frac{H_c}{H_i} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{H_i} & 0 \\ 0 & \frac{1}{H_i} \end{bmatrix} \begin{bmatrix} S_{bleft} \\ S_{bright} \end{bmatrix} \quad \text{Εξίσωση 3.38}$$

Αν ορίσουμε τα φίλτρα $F_{ii} = 1/H_i$ και $F_{ci} = -H_c/H_i^2$ τότε η (3.38) γράφεται:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i & H_c \\ H_c & H_i \end{bmatrix} \begin{bmatrix} F_{ii} & F_{ci} \\ F_{ci} & F_{ii} \end{bmatrix} \begin{bmatrix} S_{bleft} \\ S_{bright} \end{bmatrix} \quad \text{Εξίσωση 3.39}$$

Η οποία εν τέλει γίνεται:

$$\begin{bmatrix} E_L \\ E_R \end{bmatrix} = \begin{bmatrix} H_i F_{ii} + H_c F_{ci} & H_i F_{ci} + H_c F_{ii} \\ H_i F_{ci} + H_c F_{ii} & H_i F_{ii} + H_c F_{ci} \end{bmatrix} \begin{bmatrix} S_{bleft} \\ S_{bright} \end{bmatrix} \quad \text{Εξίσωση 3.40}$$

Από την στιγμή που ο σκοπός μας είναι η προσομοίωση του φυσικού συστήματος, η εξίσωση 3.21 για να είναι ίδια με την 3.40, οι όροι $H_i F_{ii} + H_c F_{ci}$, τους οποίους από εδώ και στο εξής θα τους ονομάζουμε diagonal όρους, πρέπει να είναι ίσοι με 1 και οι όροι $H_i F_{ci} + H_c F_{ii}$, τους οποίους από εδώ και στο εξής θα τους ονομάζουμε off diagonal όρους, πρέπει να είναι ίσοι με 0. Σε αυτή την περίπτωση η ακύρωση του crosstalk έχει επιτευχθεί. Αυτό θα το παρατηρήσουμε στο επόμενο κεφάλαιο στο οποίο θα παρατεθούν σχετικά διαγράμματα.

3.3.4 Επίλογος.

Σε αυτό το κεφάλαιο αναλύσαμε την δομή την οποία θα πρέπει να έχει το τελικό φίλτρο και τα επιμέρους φίλτρα αυτού, έτσι ώστε να είμαστε σε θέση να αναπαράγουμε με επιτυχία binaural αρχείο ήχου από stereo ζεύγος ηχείων.

4. Εφαρμογή των τεχνικών δημιουργίας ανάστροφου φίλτρου για την ακύρωση του crosstalk και της θέσης των ηχείων.

4.1 Εισαγωγή.

Στο τέταρτο κεφάλαιο παρουσιάζεται η εφαρμογή των δύο μεθόδων που αναλύθηκαν στο κεφάλαιο 3. Παρουσιάζονται οι τιμές οι οποίες χρησιμοποιήθηκαν κατά τις δοκιμές δημιουργίας των ανάστροφων φίλτρων για να καταλήξουμε τελικά στις καλύτερες δυνατές. Επιπλέον παρατίθενται σχήματα τα οποία απεικονίζουν την απόδοση των φίλτρων. Τέλος απαριθμούνται τα απαραίτητα βήματα που πρέπει να πραγματοποιηθούν για την δημιουργία των φίλτρων

4.2 Υλοποίηση ανάστροφου φίλτρου με την τεχνική της γρήγορης αποσυνέλιξης με κανονικοποίηση.

4.2.1 Εισαγωγή.

Σε αυτό το κεφάλαιο θα αναλυθεί η εφαρμογή της τεχνικής της γρήγορης αποσυνέλιξης με κανονικοποίηση για τη δημιουργία ανάστροφου φίλτρου. Θα παρατεθούν γραφικές παραστάσεις των αποτελεσμάτων που προκύπτουν από την εναλλαγή των τιμών των παραμέτρων, οι οποίες επηρεάζουν τις τιμές εξόδου του ανάστροφου φίλτρου και κατ' επέκταση το ηχητικό αποτέλεσμα.

4.2.2 Εφαρμογή των βημάτων για τη δημιουργία πίνακα ανάστροφου φίλτρου στο πεδίο των συχνοτήτων.

Για την δημιουργία του ανάστροφου φίλτρου χρησιμοποιήθηκε η έκδοση Beta του Matlab R2009a.

Όπως αναφέρθηκε στο κεφάλαιο 3.2.4, η επιλογή του μήκους N_h του ανάστροφου φίλτρου $H_m(z)$ γίνεται κατά προσέγγιση. Η επιλογή του μήκους N_h στην περίπτωση του συγκεκριμένου ανάστροφου φίλτρου έγινε έπειτα από δοκιμές διαφόρων τιμών. Σύμφωνα με τον Kirkeby, όπως αναφέρθηκε και στο κεφάλαιο 3.2.4, μία πρώτη προσέγγιση του μήκους του ανάστροφου φίλτρου $H_m(z)$, N_h είναι $4SN_c$, όπου ο όρος S αντιπροσωπεύει τον αριθμό των καναλιών του συστήματος,

στην προκειμένη περίπτωση είναι δύο. Το N_c είναι το μήκος του $c(n)$. Οι διαστάσεις του $c(n)$ είναι $2 \times 2 \times 512$ στο πεδίο του χρόνου.

Το μήκος του φίλτρου εξαρτάται άμεσα από τον γρήγορο μετασχηματισμό Fourier που θα εφαρμοστεί στην κρουστική απόκριση των HRTFs της θέσης των ηχείων. Δηλαδή το μήκος του φίλτρου θα είναι ίσο με τον αριθμό των δειγμάτων, για τα οποία θα πραγματοποιηθεί FFT στην κρουστική απόκριση των HRTFs της θέσης των ηχείων. Στην προκειμένη περίπτωση εφαρμόστηκε στον πίνακα $c(n)$, Fast Fourier Transform (FFT) μέσω της εντολής `fft` του Matlab για N_h σημεία.

Για τον υπολογισμό του $H(k) = \frac{1}{C(k)}$ $0 \leq k \leq N_h - 1$, πρέπει για N_h

δείγματα να πραγματοποιηθεί αναστροφή στο κάθε δείγμα ξεχωριστά. Άρα πρέπει να εφαρμοστεί η αναστροφή, $H(k) = \frac{1}{C(k)}$, N_h φορές. Αυτό που έχει μεγάλη σημασία

σε αυτή τη διαδικασία είναι ότι η αναστροφή δεν μπορεί να είναι η ίδια σε όλο το συχνοτικό φάσμα. Το ποσοστό της αναστροφής ελέγχεται από τον παράγοντα β , τον παράγοντα κανονικοποίησης (regularization parameter). Στις χαμηλές συχνότητες, των οποίων το μήκος κύματος είναι πολύ μεγαλύτερο από τις διαστάσεις του ανθρώπινου κεφαλιού, ο παράγοντας κανονικοποίησης πρέπει να είναι μεγάλος, ούτως ώστε να μην πραγματοποιείται μεγάλο ποσοστό αναστροφής. Εξαιτίας του μεγάλου μήκους κύματος η αναστροφή δεν θα επιφέρει το επιθυμητό αποτέλεσμα, αφού τα ILD και ITD παύουν να λειτουργούν στη χαμηλή συχνοτική περιοχή. Άρα το να αναστρέψουμε τις συχνότητες της χαμηλής συχνοτικής περιοχής, δηλαδή το να τροποποιηθεί το σήμα σε αυτή τη συχνοτική περιοχή, δεν θα επιφέρει κάποιο ουσιαστικό αποτέλεσμα παρά μόνο παραμόρφωση του σήματος χωρίς ουσιαστικό λόγο.

Στη συχνοτική περιοχή από τα 9 kHz και πάνω είναι επίσης προβληματική η αναστροφή αφού λόγω του πολύ μικρού μήκους κύματος, δηλαδή της μεγάλης κατευθυντικότητας του ηχητικού σήματος, δεν είναι εφικτό να ξεγελαστεί το ανθρώπινο ψυχοακουστικό σύστημα. Σε αυτή τη συχνοτική περιοχή η φιλοσοφία είναι ανάλογη με την παραπάνω, στις χαμηλές συχνότητες. Εισάγεται δηλαδή μεγάλος παράγοντας κανονικοποίησης για την μικρότερη δυνατή αναστροφή.

Το συμπέρασμα, όσον αφορά την συχνοτική περιοχή λειτουργίας του φίλτρου, είναι ότι το φίλτρο λειτουργεί ουσιαστικά από τα 400 Hz περίπου, ως τα 9000 Hz περίπου.

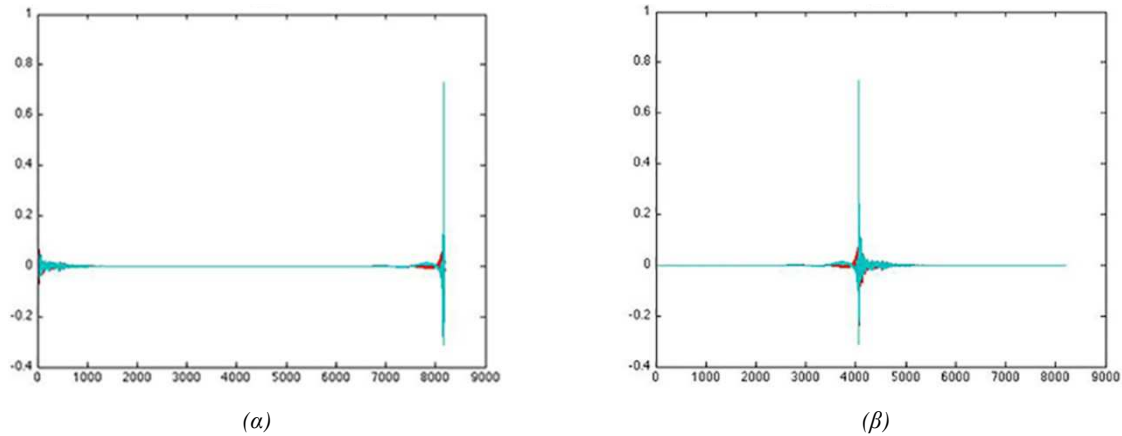
Έπειτα από τον υπολογισμό του $H(k) = \frac{1}{C(k)}$, $0 \leq k \leq N_h - 1$, αφού δηλαδή έχει γίνει αναστροφή όλων των δειγμάτων του αποτελέσματος του Fast Fourier Transform με τον κατάλληλο παράγοντα κανονικοποίησης για κάθε συχνοτική περιοχή, πρέπει το σήμα να μεταβεί στο πεδίο του χρόνου. Για να επιτευχθεί αυτό, χρησιμοποιήθηκε η εντολή του Matlab, η οποία πραγματοποιεί ανάστροφο γρήγορο μετασχηματισμό Fourier, IFFT.

Το επόμενο βήμα και τελευταίο για τη δημιουργία του ανάστροφου φίλτρου είναι η εφαρμογή του modeling delay, μήκους $N_h/2$ δειγμάτων, έτσι ώστε να ικανοποιηθεί η αιτιότητα του φίλτρου. Η αιτιότητα του φίλτρου εξασφαλίζει ότι η εκάστοτε έξοδος του φίλτρου, έπεται χρονικά της αντίστοιχης εισόδου. Αυτή η διαδικασία πραγματοποιήθηκε με την τεχνική cyclic shift σε κάθε δείγμα του $h_m(n)$, χρησιμοποιώντας την εντολή FFTSHIFT του Matlab, για να πάρουμε το $h(n)$. Το $h(n)$ είναι ένας πίνακας διαστάσεων $2X2XN_h$. Επιπλέον με την εισαγωγή του modeling delay μήκους $N_h/2$ συγκεντρώνεται όλη η ενέργεια στο κέντρο του πίνακα τιμών, που ουσιαστικά αντιπροσωπεύει το φίλτρο, δημιουργώντας έτσι την κρουστική απόκριση του φίλτρου, εικόνα 7. Συνοψίζοντας τα βήματα της δημιουργίας του ανάστροφου φίλτρου, τα οποία αναφέρθηκαν σε αυτό το κεφάλαιο, αυτά είναι τα παρακάτω.

- Εφαρμογή της εντολής FFT του Matlab στον πίνακα $C(z)$ για N_h δείγματα.
- Διαχωρισμός του σήματος σε συχνοτικές περιοχές και εφαρμογή του κατάλληλου παράγοντα κανονικοποίησης β σε κάθε συχνοτική μπάντα.
- Αναστροφή κάθε δείγματος του $H(k) = \frac{1}{C(k)}$, $0 \leq k \leq N_h - 1$, με τον

$$\text{αντίστοιχο παράγοντα κανονικοποίησης } \beta. \quad H(k) = \frac{C^H(k)}{C^H(k) * C(k) + \beta}$$

- Εφαρμογή της εντολής IFFT του Matlab στο $H(k) = \frac{1}{C(k)}$, με σκοπό την επαναφορά του σήματος στο πεδίο του χρόνου, $H_m(n) = IFFT(H(k))$.
- Εφαρμογή της εντολής FFTSHIFT του Matlab στο $H_m(n)$. Εφαρμόζοντας FFTSHIFT σε κάθε δείγμα του $H_m(n)$ εφαρμόζεται modeling delay μήκους $N_h / 2$ δειγμάτων, παίρνοντας το τελικό φίλτρο $H(n)$.



Εικόνα 7: Κρουστική απόκριση ανάστροφου φίλτρου ακύρωσης crosstalk. (α) πριν την εφαρμογή του modelling delay, (β) μετά την εφαρμογή του modelling delay.

4.2.3 Εφαρμογή τιμών στις παραμέτρους, που επηρεάζουν την έξοδο του ανάστροφου φίλτρου.

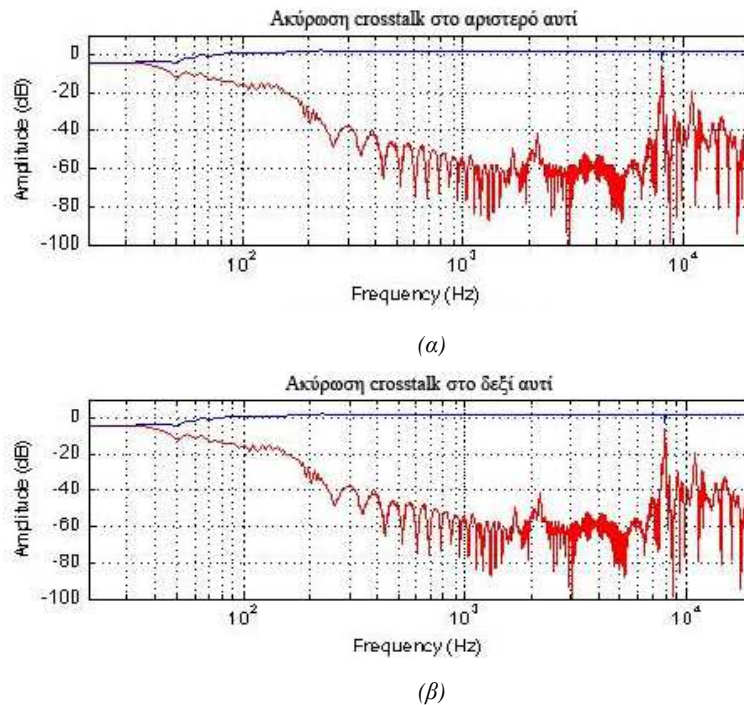
Στον πίνακα 1 παρουσιάζονται οι τιμές του παράγοντα κανονικοποίησης β ανά συχνοτική περιοχή. Τα όρια χωρισμού του συχνοτικού φάσματος σε μπάντες (reg limits) και η εκάστοτε επιλογή του μήκους (length) του φίλτρου. Άρα παίρνοντας ως παράδειγμα την πρώτη σειρά του πίνακα, ο παράγοντας κανονικοποίησης β είναι 10^{-2} από 0 Hz ως 200 Hz, 10^{-3} από 200 Hz ως 1000 Hz, 10^{-4} από 1000 Hz ως 2000 Hz, 10^{-3} από 2000 Hz ως 8000 Hz και 10^{-2} από 8000 Hz ως 20000 Hz. Το μήκος του φίλτρου είναι στα 8192 δείγματα. Αυτό σημαίνει ότι εφαρμόστηκε γρήγορος μετασχηματισμός Fourier για 8192 σημεία στην κρουστική απόκριση των HRTFs της γωνίας που είναι τοποθετημένα τα ηχεία. Η στήλη A/A παραθέτει τους αύξοντες αριθμούς των δοκιμαστικών φίλτρων που κατασκευάστηκαν.

A/A	β	reg limits (Hz)	Length N_h (Taps)	A/A	β	reg limits (Hz)	Length N_h (Taps)
1	10^{-2}	0-200	8192	6	10^{-3}	0-200	8192
	10^{-3}	201-1000			10^{-4}	201-1000	
	10^{-4}	1001-2000			10^{-5}	1001-2000	
	10^{-3}	2001-8000			10^{-4}	2001-8000	
	10^{-2}	8001-20000			10^{-3}	8001-20000	
2	10^{-2}	0-200	4096	7	10^{-4}	0-200	8192
	10^{-3}	201-1000			10^{-5}	201-1000	
	10^{-4}	1001-2000			10^{-6}	1001-2000	
	10^{-3}	2001-8000			10^{-5}	2001-8000	
	10^{-2}	8001-20000			10^{-4}	8001-20000	
3	10^{-2}	0-200	2048	8	10^{-4}	0-200	4096
	10^{-3}	201-1000			10^{-5}	201-1000	
	10^{-4}	1001-2000			10^{-6}	1001-2000	
	10^{-3}	2001-8000			10^{-5}	2001-8000	
	10^{-2}	8001-20000			10^{-4}	8001-20000	
4	10^{-2}	0-200	1024	9	10^{-4}	0-200	2048
	10^{-3}	201-1000			10^{-5}	201-1000	
	10^{-4}	1001-2000			10^{-6}	1001-2000	
	10^{-3}	2001-8000			10^{-5}	2001-8000	
	10^{-2}	8001-20000			10^{-4}	8001-20000	
5	10^{-1}	0-200	8192	10	10^{-2}	0-200	16389
	10^{-2}	201-1000			10^{-3}	201-1000	
	10^{-3}	1001-2000			10^{-4}	1001-2000	
	10^{-2}	2001-8000			10^{-3}	2001-8000	
	10^{-1}	8001-20000			10^{-2}	8001-20000	

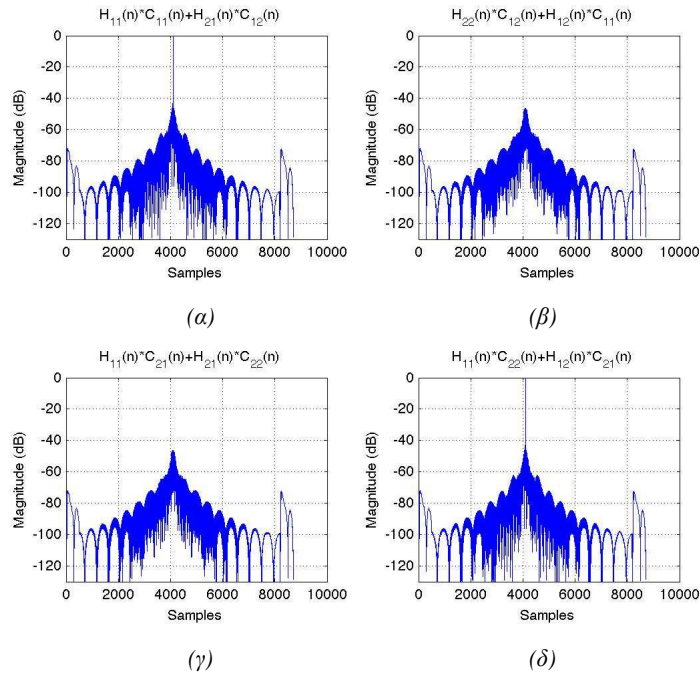
Πίνακας 1: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου.

Δημιουργήθηκαν ανάστροφα φίλτρα με τις τιμές του πίνακα 1 και τα αποτελέσματά τους όπως ήταν εξ' αρχής αναμενόμενο παρουσίασαν διαφορές μεταξύ τους. Στις γραφικές παραστάσεις που ακολουθούν παρατίθενται η ακύρωση του crosstalk και οι

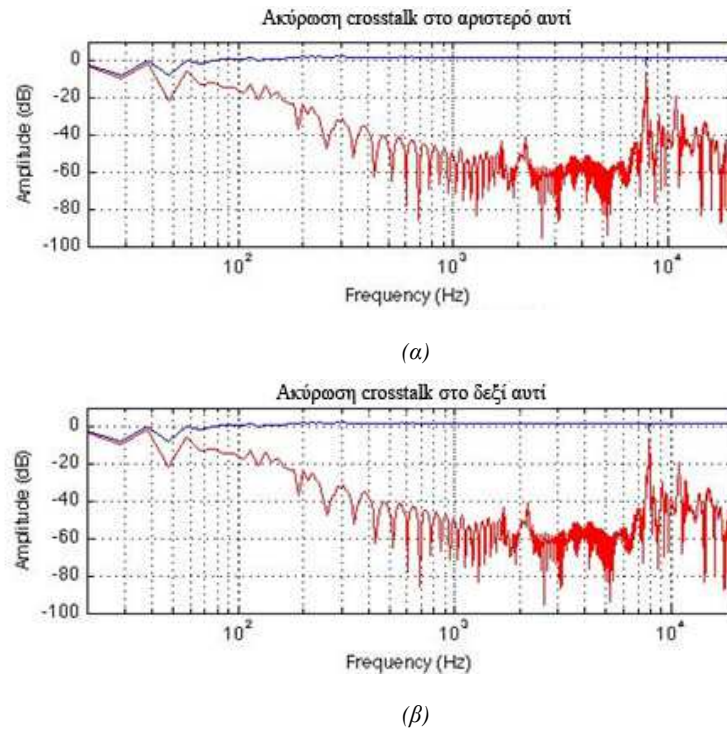
όροι diagonal και off-diagonal. Για να μεταφερθεί στο κάθε αυτί του ακροατή το επιθυμητό σήμα, οι όροι diagonal $H_{11}(n)*C_{11}(n)+H_{21}(n)*C_{12}(n)$ και $H_{11}(n)*C_{22}+H_{12}(n)*C_{21}(n)$ πρέπει να είναι 1 και οι όροι off-diagonal $H_{22}(n)*C_{12}(n)+H_{12}(n)*C_{11}(n)$ και $H_{11}(n)*C_{21}(n)+H_{21}(n)*C_{22}(n)$ πρέπει να είναι 0, να είναι δηλαδή πολύ χαμηλότερης στάθμης σε σχέση με τον diagonal όρο. Αυτό σημαίνει ότι έχει επιτευχθεί η ακύρωση της θέσης των ηχείων και η ακύρωση του crosstalk. Στις εικόνες 8 ως 27 παρουσιάζονται τα αποτελέσματα της ακύρωσης του crosstalk και οι όροι, diagonal και off-diagonal. Επιπλέον η ακύρωση του crosstalk πρέπει να είναι τουλάχιστον 15 dB για να καταφέρει το ανθρώπινο ακουστικό αισθητήριο να παραβλέψει την παρουσία του.



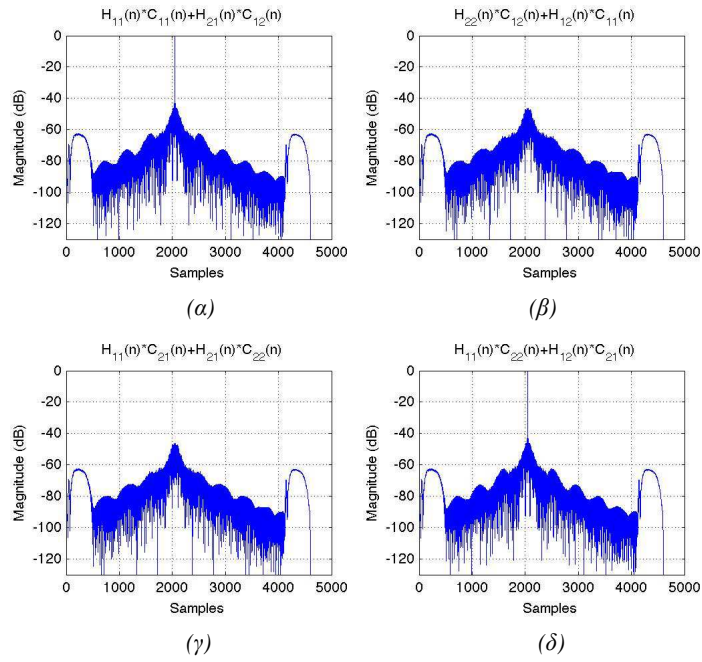
Εικόνα 8: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 1). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



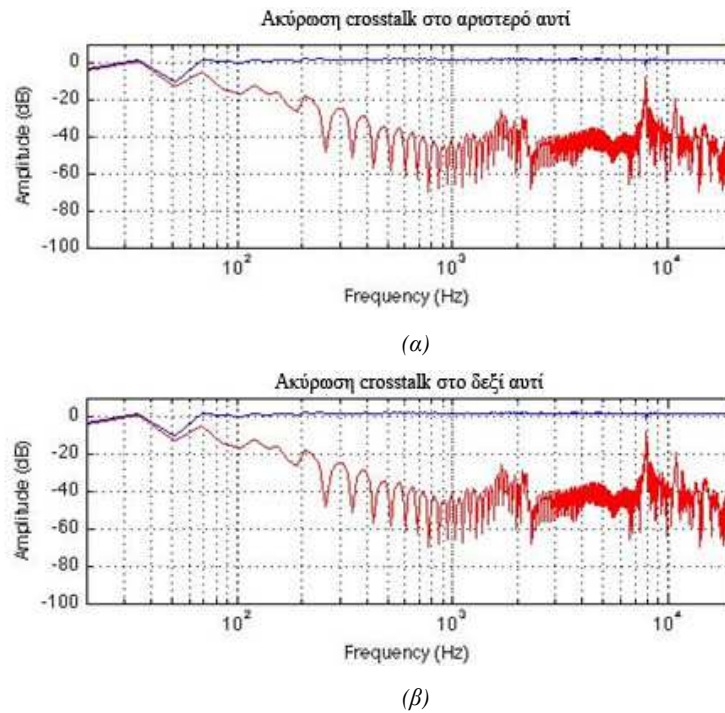
Εικόνα 9: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 1) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.



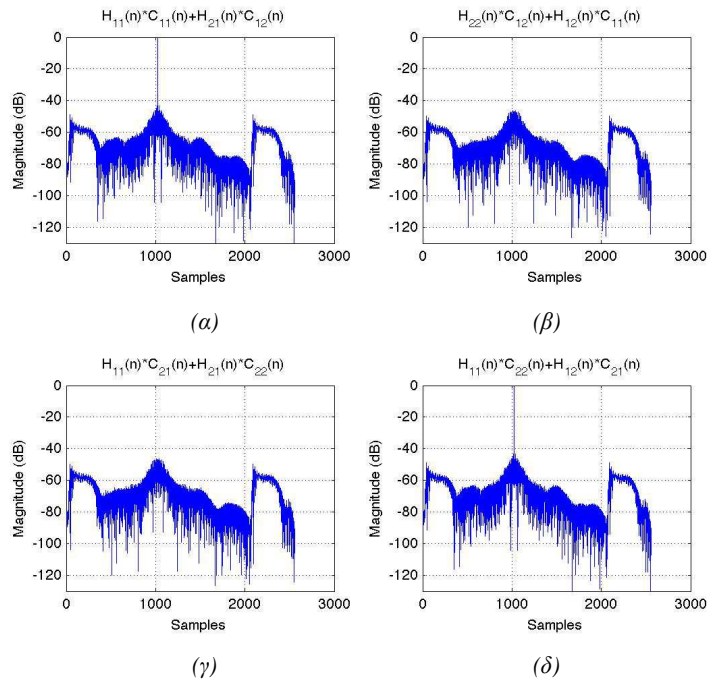
Εικόνα 10: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 2). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



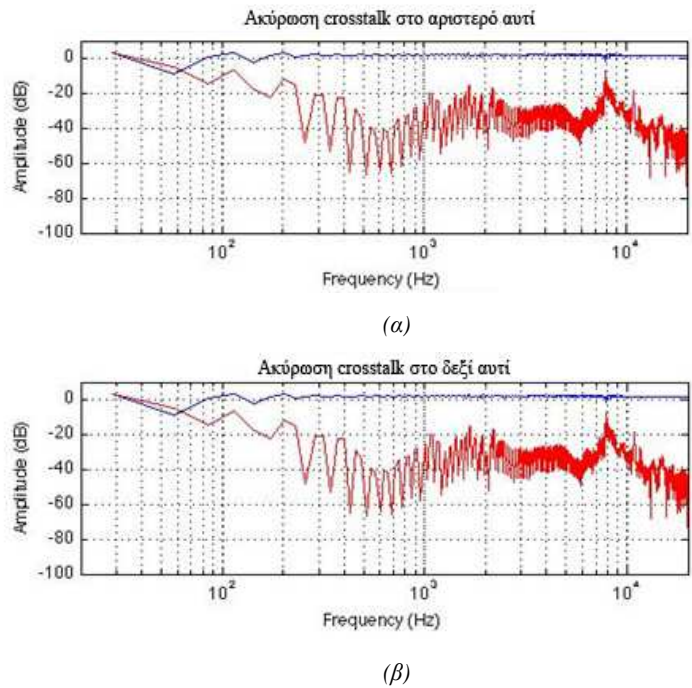
Εικόνα 11: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 2) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



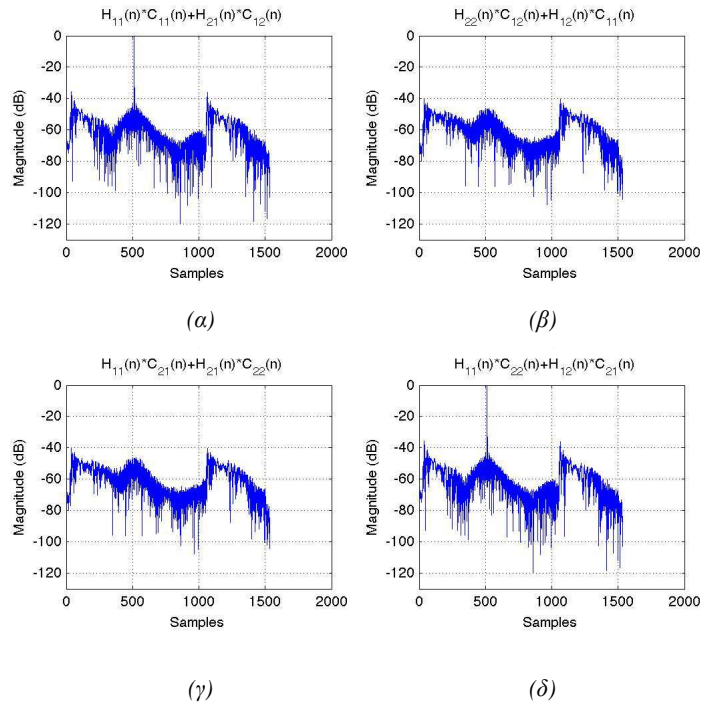
Εικόνα 12: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 3). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



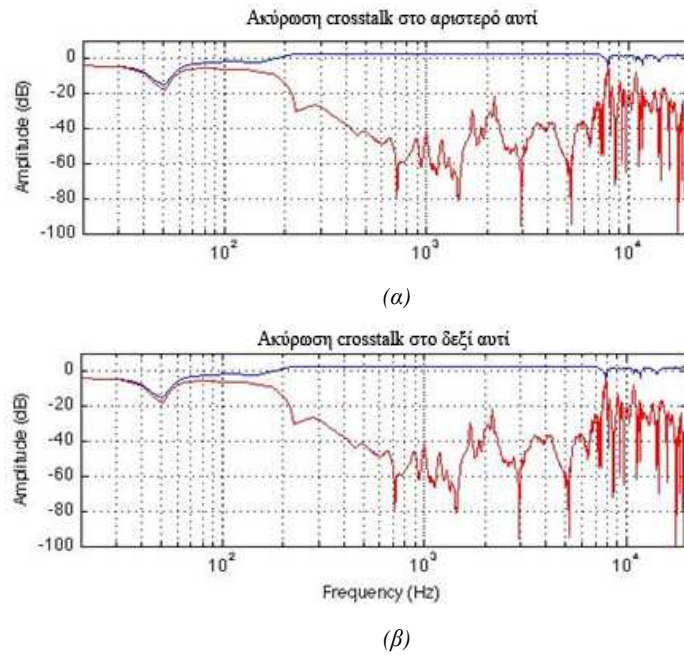
Εικόνα 13: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 3) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



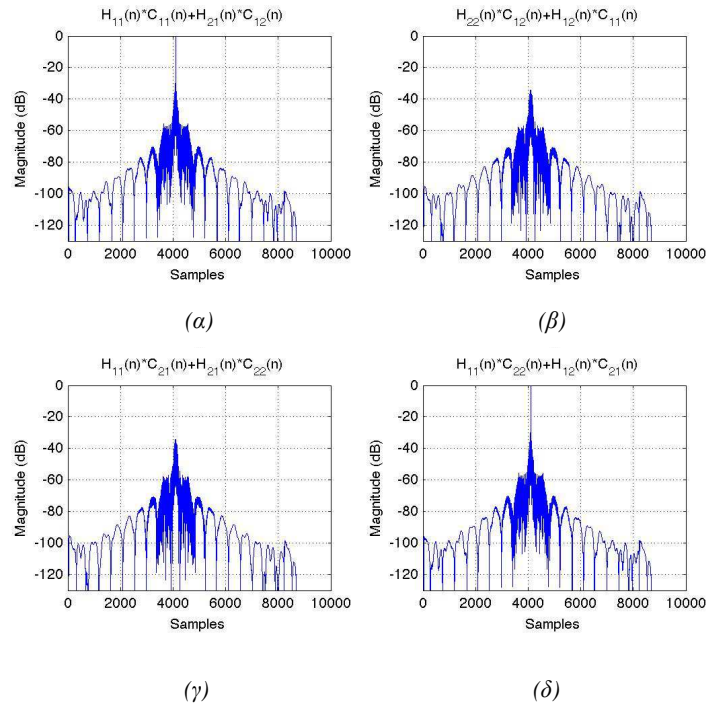
Εικόνα 14: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 4). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



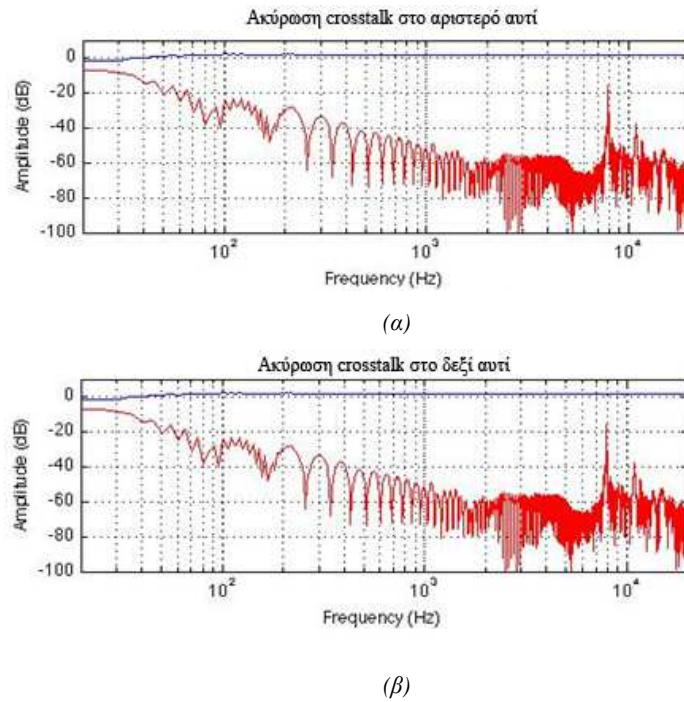
Εικόνα 15: Οι όροι *diagonal* (α, δ) και *off-diagonal* (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 4) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



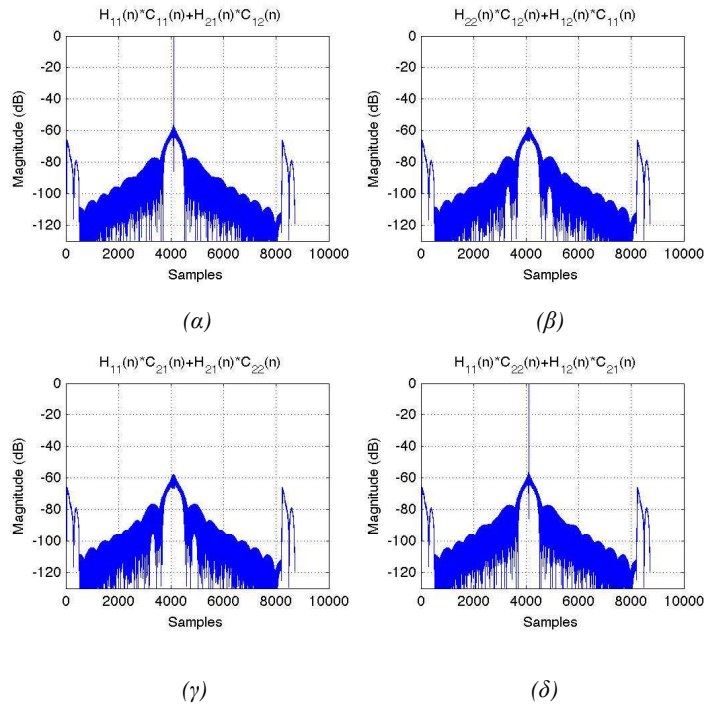
Εικόνα 16: Ακύρωση του *crosstalk* για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 5). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (*ipsilateral*) και η κόκκινη καμπύλη το *crosstalk*.



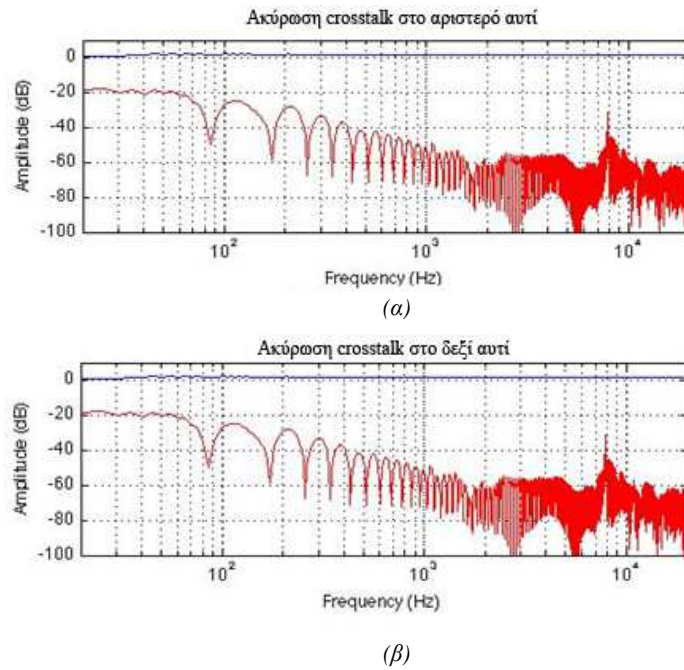
Εικόνα 17: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 5) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



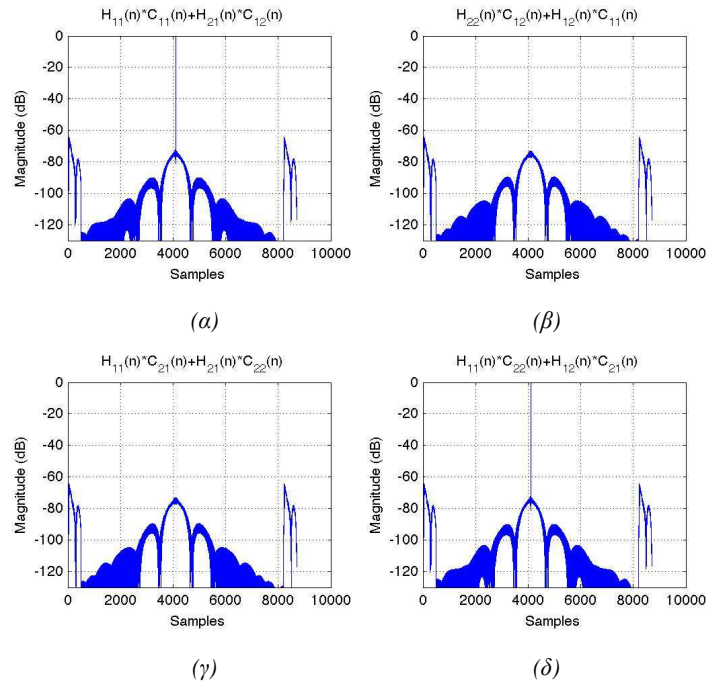
Εικόνα 18: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 6). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



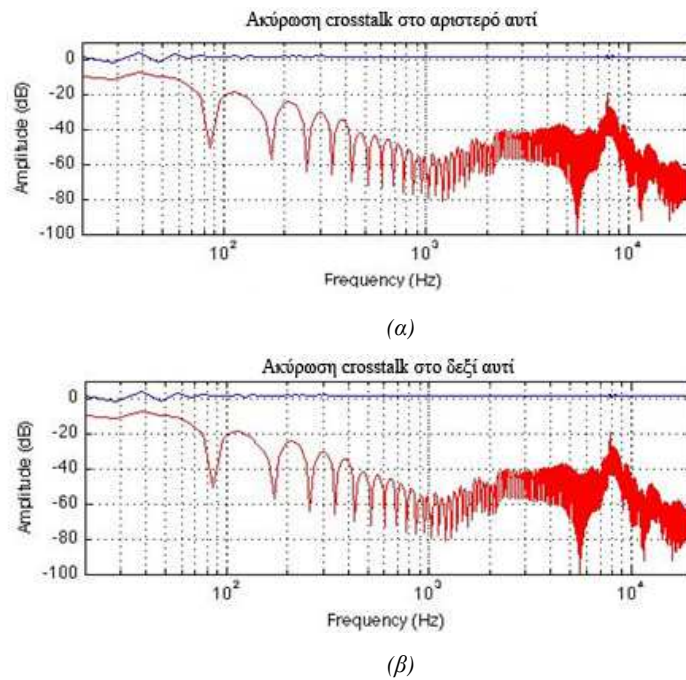
Εικόνα 19: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 6) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



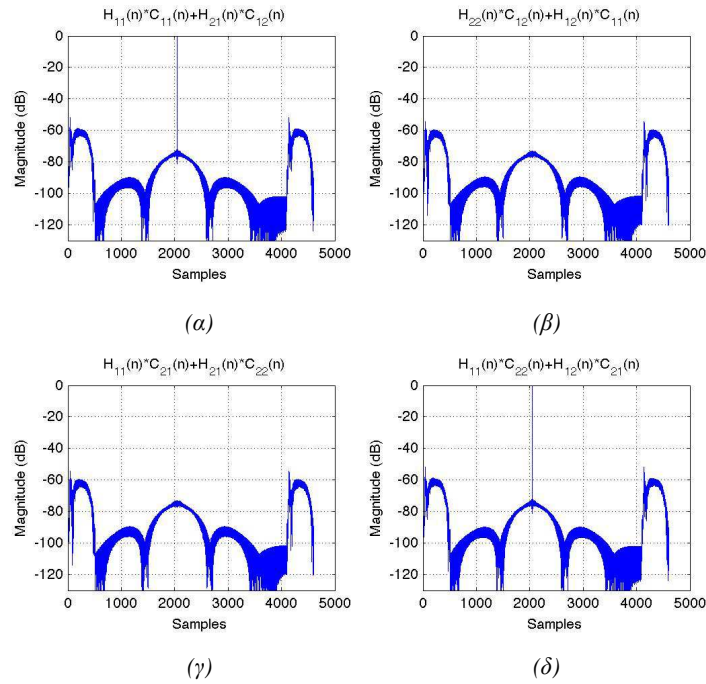
Εικόνα 20: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 7). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



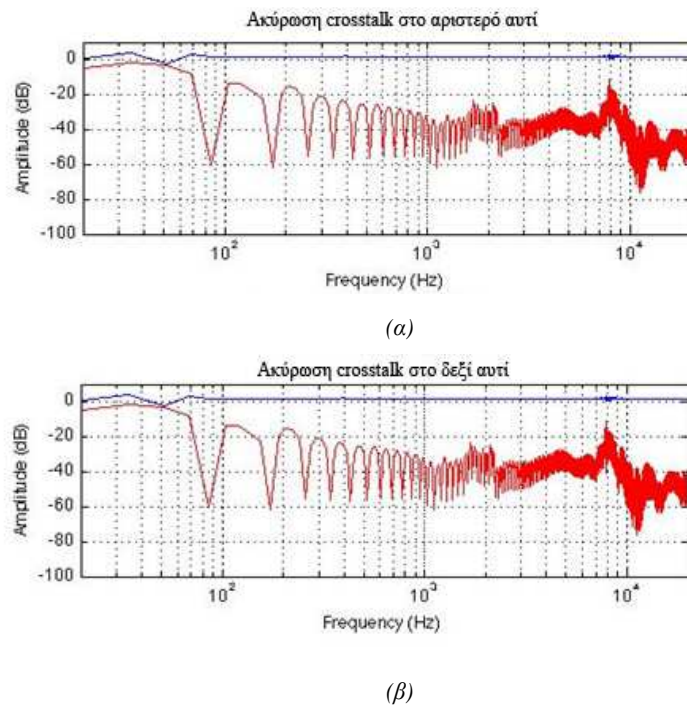
Εικόνα 21: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 7) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



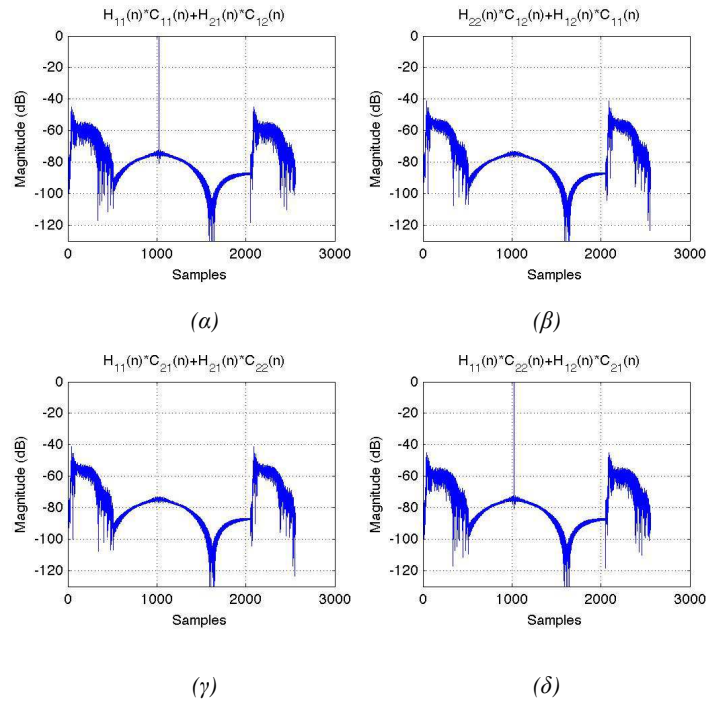
Εικόνα 22: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 8). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



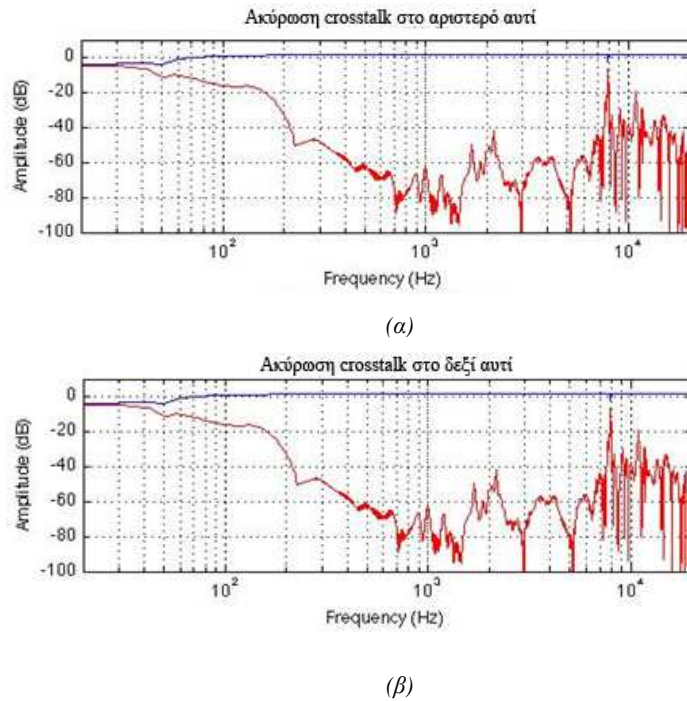
Εικόνα 23: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 8) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.



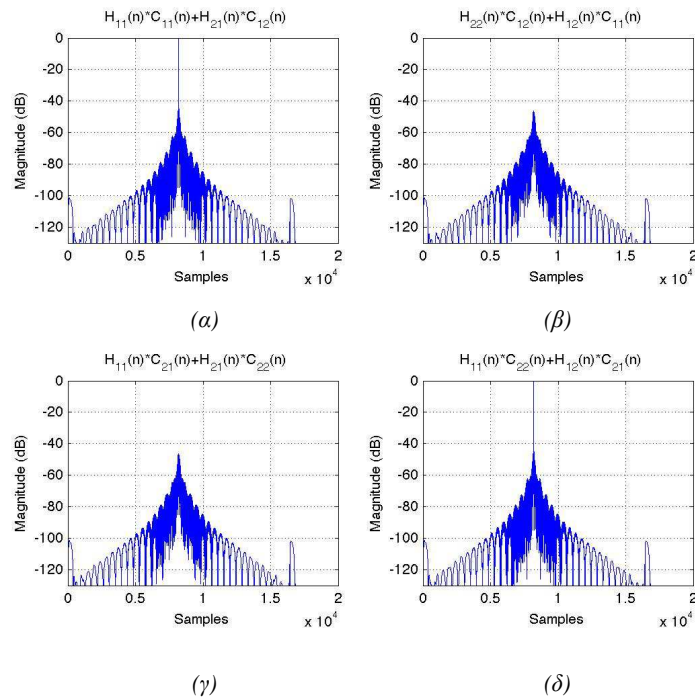
Εικόνα 24: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 9). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.



Εικόνα 25: Οι όροι diagonal (α, δ) και off-diagonal (β, γ) που προκύπτουν με τη χρήση του φίλτρου (A/A 9) για το αριστερό (α, β) και το δεξί (γ, δ) κανάλι.

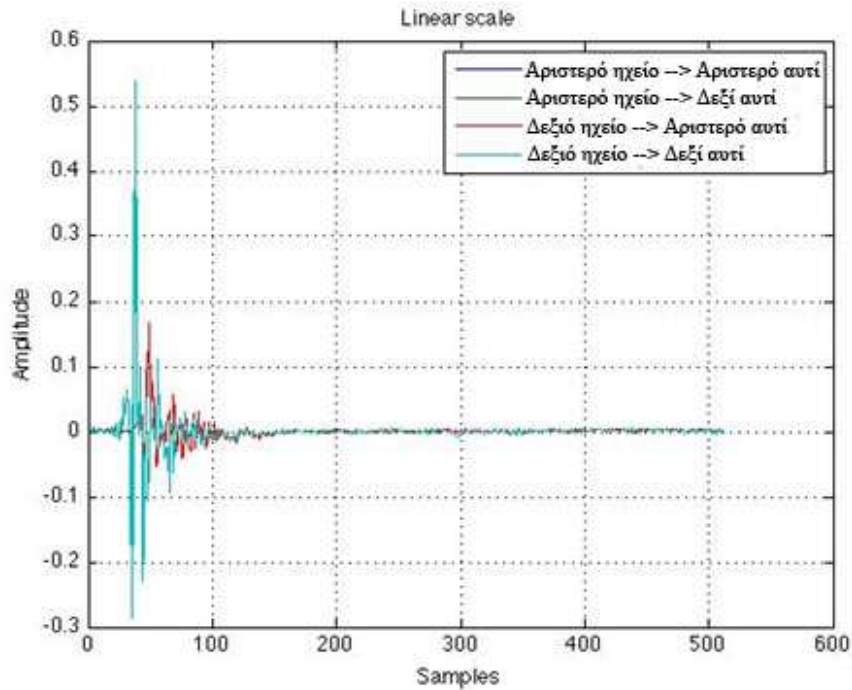


Εικόνα 26: Ακύρωση του crosstalk για το αριστερό αυτί (α) και για το δεξί αυτί (β) με τη χρήση του φίλτρου (A/A 10). Η μπλε καμπύλη αναπαριστά τον απευθείας ήχο (ipsilateral) και η κόκκινη καμπύλη το crosstalk.

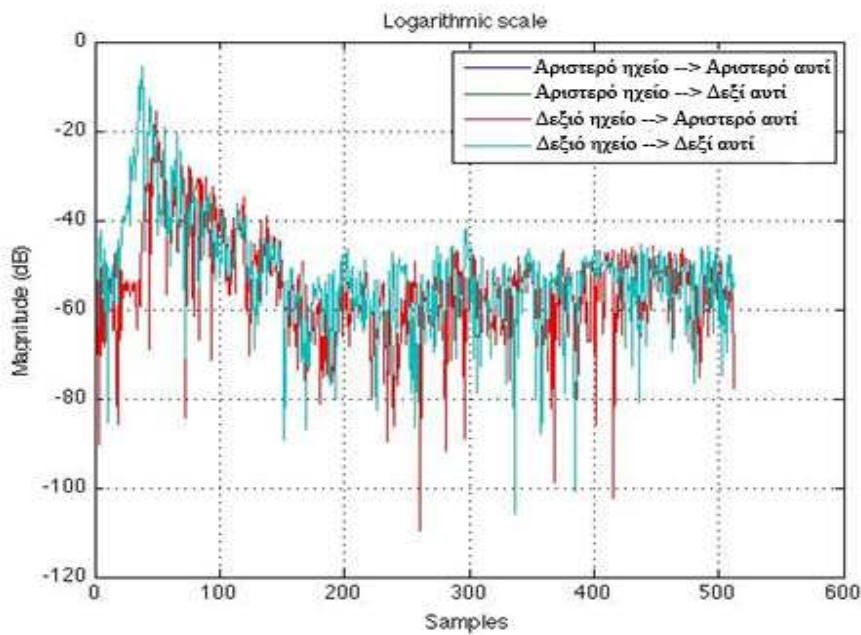


Εικόνα 27: Οι όροι diagonal (α , δ) και off-diagonal (β , γ) που προκύπτουν με τη χρήση του φίλτρου (A/A_{10}) για το αριστερό (α , β) και το δεξί (γ , δ) κανάλι.

Στην εικόνα 28 απεικονίζεται η κρουστική απόκριση, στο πεδίο του χρόνου, των HRTFs της γωνίας των 30° όπου είναι τοποθετημένα τα ηχεία στο σύστημα ακρόασης. Στην εικόνα 29 απεικονίζεται το φάσμα της ίδιας κρουστικής, μετά την εφαρμογή του FFT, δηλαδή στο πεδίο των συχνοτήτων. Αυτή είναι η κρουστική απόκριση που χρησιμοποιήθηκε.



Εικόνα 28: Η κρουστική απόκριση των HRTFs των 30° στο πεδίο του χρόνου (IRCAM database).



Εικόνα 29: Η κρουστική απόκριση των HRTFs των 30° στο πεδίο των συχνότητων (IRCAM database).

4.2.4 Ανάλυση των αποτελεσμάτων των φίλτρων.

Παρατηρώντας τις εικόνες 8 μέχρι και 27 διακρίνονται οι διαφορές που υπάρχουν στα αποτελέσματα της ακύρωσης του crosstalk και των όρων diagonal και off-diagonal σε κάθε απόπειρα δημιουργίας του επιθυμητού φίλτρου.

Δημιουργώντας φίλτρα, χρησιμοποιώντας τις τιμές του πίνακα 1 και λαμβάνοντας τα αποτελέσματα τους, παρατηρείται ότι το καλύτερο αποτέλεσμα το δίνει το φίλτρο με αύξοντα αριθμό 1 (A/A 1). Επίσης ικανοποιητικό αποτέλεσμα έχει το φίλτρο A/A 2, με τη διαφορά ότι παρατηρούνται στην εικόνα 26 απότομες κορυφές και απότομα βυθίσματα σε αντίθεση με την εικόνα 1 όπου η ακύρωση του crosstalk είναι ομαλότερη. Αυτό πρακτικά σημαίνει μεγαλύτερη αλλοίωση του σήματος στη δεύτερη περίπτωση. Μεγαλύτερη από αυτή που επιφέρει στο σήμα το φίλτρο A/A 1. Ένα επιπλέον χαρακτηριστικό που παρατηρείται είναι το γεγονός ότι στις περισσότερες από τις απόπειρες δημιουργίας του φίλτρου, τα αποτελέσματα, τα οποία φαίνονται στις εικόνες 1 μέχρι και 27, ικανοποιούν θεωρητικά τους όρους επιτυχής λειτουργίας του φίλτρου. Σε αυτό το σημείο επιβεβαιώνεται και αυτό που αναφέρθηκε παραπάνω σχετικά με τον τρόπο που επιλέγονται οι τιμές των παραμέτρων του φίλτρου, δηλαδή κατά προσέγγιση. Ουσιαστικά επιλέχθηκαν τιμές με σχετικά μεγάλη διαφορά μεταξύ τους και παρά το γεγονός αυτό, τα αποτελέσματα είναι ικανοποιητικά. Παρατηρείται δηλαδή ότι η ακύρωση του crosstalk είναι πάνω από 15 dB σε όλες τις περιπτώσεις, οι όροι diagonal είναι στα 0 dB και οι όροι off-diagonal είναι στα -40 dB περίπου, πράγμα που σημαίνει ότι ουσιαστικά δεν επηρεάζουν την ακρόαση.

Παρατηρώντας τα αποτελέσματα των φίλτρων από A/A 1 ως και A/A 4, στα οποία το μόνο που αλλάζει είναι το μήκος του φίλτρου, είναι εμφανές ότι όσο μειώνεται το μήκος του φίλτρου, τόσο μειώνεται η απόδοση του.

Στην περίπτωση του φίλτρου A/A 5, στο οποίο είναι αυξημένος ο παράγοντας κανονικοποίησης, αλλά όχι το μήκος του, παρατηρείται μείωση στην απόδοση σε σχέση με το A/A 1 το οποίο έχει το ίδιο μήκος.

Το φίλτρο A/A 6, το οποίο έχει ίσο μήκος με το A/A 1, με μία αλλαγή στον παράγοντα κανονικοποίησης παρουσιάζει αποτελέσματα πολύ ικανοποιητικά, όσον αφορά στην ακύρωση του crosstalk και στους όρους diagonal, off-diagonal, παρουσιάζει όμως κατά την ακρόαση πολύ υψηλής στάθμης συχνότητα κουδουνίσματος, σε σχέση με το φίλτρο A/A 1.

Τα φίλτρα A/A 7, A/A 8 και A/A 9, διαφέρουν μόνο ως προς το μήκος τους, Το φίλτρο A/A 7 παρουσιάζει ικανοποιητικά αποτελέσματα όσον αφορά στην ακύρωση του crosstalk και στους όρους diagonal, off-diagonal, όμως κατά την ακρόαση, παρουσιάζει υψηλή στάθμης, συχνότητα κουδουνίσματος σε σχέση με το

φίλτρο A/A 1. Τα φίλτρα A/A 8 και A/A 9 παρουσιάζουν διαδοχικά μείωση στην απόδοσή τους σε σχέση με το φίλτρο A/A 7, λόγω του μειωμένου τους μήκους.

Όπως αναφέρθηκε και παραπάνω το καλύτερο αποτέλεσμα το επιφέρει το φίλτρο A/A 1. Παρόλο που όλα τα φίλτρα τηρούν τις θεωρητικές προδιαγραφές, το αποτέλεσμα του A/A 1 είναι καλύτερο όσον αφορά την παραμόρφωση στο σήμα, το ringing frequency και την διαύγεια ως προς τη θέση που αντιλαμβάνεται ο ακροατής την εικονική ηχητική πηγή. Οι τιμές που χρησιμοποιήθηκαν για τη δημιουργία του φίλτρου, με το οποίο επεξεργάστηκαν οι ήχοι που χρησιμοποιήθηκαν στο πείραμα, είναι οι τιμές του φίλτρου A/A 1.

Επιπλέον παρατηρήθηκε ότι όσον αφορά το μήκος του φίλτρου, N_h , τα 8192 δείγματα είναι το μεγαλύτερο μήκος στο οποίο επέρχεται βελτίωση της απόδοσης του φίλτρου. Με άλλα λόγια, όσο και να αυξηθεί το μήκος του φίλτρου δεν υπάρχει κάποια βελτίωση στο αποτέλεσμα πάνω από τα 8192 δείγματα.

4.2.5 Η εφαρμογή του φίλτρου στο binaural σήμα.

Η εφαρμογή του φίλτρου στο binaural σήμα πραγματοποιείται σύμφωνα με τη σχέση 3.16 Ο πολλαπλασιασμός των όρων των $C(z)$ και $H(z)$ πραγματοποιείται με τη χρήση της εντολής FFTFILT του Matlab. Έστω ότι $Y = \text{FFTFILT}(B, X)$. Η εντολή FFTFILT φιλτράρει τα δεδομένα χρησιμοποιώντας τη μέθοδο overlap-add, η οποία βασίζεται στην διαδικασία FFT (Fast Fourier Transform). Μία τεχνική φιλτραρίσματος στο πεδίο των συχνοτήτων η οποία λειτουργεί μόνο για FIR φίλτρα. Παίρνοντας ως παράδειγμα την ισότητα $Y = \text{FFTFILT}(B, X)$, σε αυτήν την περίπτωση η FFTFILT φιλτράρει τα δεδομένα στο διάνυσμα X με το φίλτρο που περιγράφεται από τον συντελεστή B . Επιστρέφει τον πίνακα τιμών Y . Η λειτουργία, η οποία εκτελείται από την FFTFILT, περιγράφεται στο πεδίο του χρόνου από την εξίσωση διαφοράς:

$$Y(n) = B(1)X(n) + B(2)X(n-1) + \dots + B(nB+1)X(n-nB) \quad \text{Εξίσωση 4.1}$$

Μία ισοδύναμη αναπαράσταση είναι ο μετασχηματισμός $-z$ ή η αναπαράσταση στο πεδίο των συχνοτήτων.

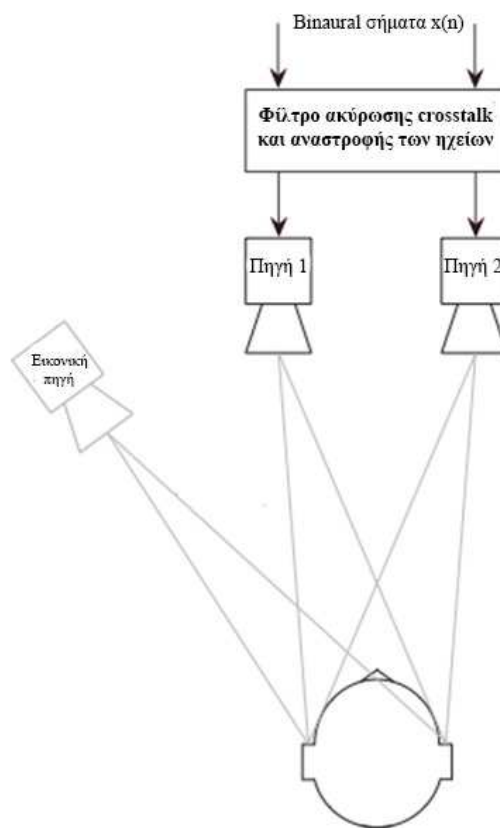
$$Y(z) = [B(1) + B(2)z^{-1} + \dots + B(nB+1)z^{-nB}]Xz \quad \text{Εξίσωση 4.2}$$

Η FFTFILT επιλέγει αυτόματα το μήκος, σε δείγματα, του FFT και το μήκος των δεδομένων που εγγυώνται αποτελεσματικό χρόνο εκτέλεσης.

Μετά την εφαρμογή του φίλτρου, το σήμα είναι έτοιμο προς αναπαραγωγή από σύστημα δύο καναλιών, με τα ηχεία τοποθετημένα στις μοίρες των HRTFs που χρησιμοποιήθηκαν για τη δημιουργία του φίλτρου. Στην εικόνα 30 παρουσιάζεται το σύστημα αναπαραγωγής binaural σήματος από δύο ηχεία.

4.2.6 Επίλογος.

Στο κεφάλαιο αυτό αναλύθηκαν τα βήματα, τα οποία πραγματοποιήθηκαν, για τη δημιουργία ανάστροφου φίλτρου με την τεχνική της γρήγορης αποσυνέλιξης με κανονικοποίηση. Παρατέθηκαν και αναλύθηκαν τα αποτελέσματα των φίλτρων για ένα πλήθος τιμών ανά παράμετρο του φίλτρου. Τέλος έγινε αναφορά στην εφαρμογή του φίλτρου στο, προς αναπαραγωγή binaural σήμα.



Εικόνα 30: Δικάναλο σύστημα ηχείων για την αναπαραγωγή binaural σήματος (Manias, 2008).

4.3 Υλοποίηση ανάστροφου φίλτρου με την τεχνική των ελαχίστων μέσων τετραγώνων.

4.3.1 Εισαγωγή.

Σε αυτό το κεφάλαιο θα αναλύσουμε την μεθοδολογία που ακολουθήσαμε για την κατασκευή του φίλτρου στο Matlab. Θα γίνει ανάλυση των αποτελεσμάτων μέσω γραφημάτων των όρων diagonal – off diagonal και της διαφοράς έντασης ανά συχνότητα των ipsilateral και contralateral όρων, για διαφορετικές τιμές δειγμάτων και καθυστέρησης των επιμέρους φίλτρων.

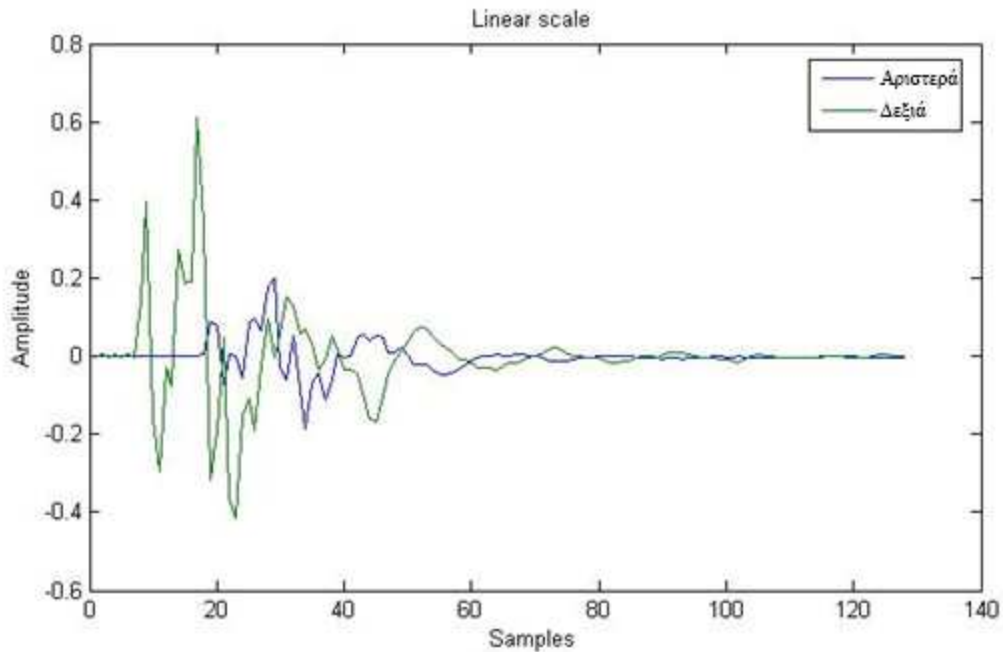
4.3.2 Ανάλυση αλγορίθμου.

Ο αλγόριθμος που εξετάζουμε για την δημιουργία φίλτρου ακύρωσης του crosstalk των ηχείων σε θέση 30 μοιρών αποτελείται από δύο mfiles. Στο πρώτο ορίζουμε τα δείγματα, την καθυστέρηση για επιμέρους φίλτρα f_{ii} και f_{ci} και το SER. Στο δεύτερο εμπεριέχετε αλγόριθμος LMS ο οποίος με βάση τα στοιχεία που έχουμε δώσει κατασκευάζει το φίλτρο, με συνθήκη να ολοκληρωθεί όταν το SER φτάσει την επιθυμητή τιμή.

Αρχικά στο πρώτο mfile που ορίζουμε τις τιμές:

- Η συνθήκη που πρέπει να ικανοποιείται για να δουλέψει ο αλγόριθμος είναι, ο αριθμός δειγμάτων των φίλτρων να είναι μεγαλύτερος από αυτών της κρουστικής. Η κρουστική που χρησιμοποιήσαμε είναι 128 δειγμάτων. Αυτό μας δίνει τη δυνατότητα να κατασκευάσουμε φίλτρα με αρκετά μικρό μήκος και να διατηρήσουμε τον υπολογιστικό χρόνο σε χαμηλά επίπεδα.

Στην εικόνα 31 μπορούμε να παρατηρήσουμε την κρουστική των 30 μοιρών δεξιά από τον ακροατή στο πεδίο του χρόνου.



Εικόνα 31: Η κρουστική απόκριση των 30° στο πεδίο του χρόνου (KEMAR).

- Χωρίζουμε το αριστερό και το δεξί κανάλι, δημιουργώντας τους ipsilateral και contralateral όρους. Από την στιγμή που η κρουστική είναι 30 μοίρες δεξιά ορίζουμε ipsilateral το δεξί κανάλι και contralateral το αριστερό.

Στην συνέχεια δημιουργούμε το φίλτρο f_{ci} .

- Ορίζουμε ένα πίνακα με τόσα μηδενικά όσα και τα δείγματα του φίλτρου.
- Ορίζουμε αρχικά το SER ίσο με το μηδέν.
- Με την χρήση της εντολής while ξεκινάμε διαδικασία επανάληψης η οποία θα ολοκληρωθεί όταν η τιμή του SER περάσει τα 30dB.
- Θέτουμε σε λειτουργία το δεύτερο mfile που περιέχει τον αλγόριθμο του LMS

Για την δημιουργία του φίλτρου f_{ii} όπως αναφέραμε και στο κεφάλαιο 3.3.3 θα εφαρμόσουμε συνέλιξη μεταξύ των δυο επιμέρους φίλτρων με συχνοτική απόκριση H_c/H_i και $1/H_i$. Οπότε επαναλαμβάνουμε την ίδια διαδικασία όπως στο πρώτο φίλτρο, αλλάζοντας μόνο τον αριθμό των δειγμάτων και της καθυστέρησης, με βάση την θεωρία που αναλύσαμε στο κεφάλαιο 3.3.3 και ομοίως για το δεύτερο, μόνο που αυτή την φορά εκτός από το μήκος του φίλτρου και την καθυστέρηση αλλάζουμε και τον αριθμητή, αντικαθιστώντας τον με τον αριθμό 1.

Στο δεύτερο mfile:

- Πραγματοποιούμε αρχικοποίηση του φίλτρου. ορίζουμε h_x τον αριθμητή του φίλτρου, h_y ο παρονομαστής του φίλτρου, len το μήκος σε δείγματα του φίλτρου, $delay$ η καθυστέρηση σε δείγματα του φίλτρου και lms_loops ο αριθμός επαναλήψεων.
- Ορίζουμε το μέγεθος λευκού θορύβου που θα χρησιμοποιήσουμε ως είσοδο στο σύστημα ($w(n)$).
- Ορίζουμε το a ίσο με ένα πολύ μικρό αριθμό (10^{-4}).
- Ορίζουμε το $e(n)$ ίσο με μηδέν αρχικά.
- Έπειτα ξεκινάμε την διαδικασία επανάληψης για στην οποία θα κατασκευάσουμε λευκό θόρυβο και θα τον φιλτράρουμε με το h_x και το h_y .
- Δημιουργούμε λευκό θόρυβο για είσοδο.
- Υπολογίζουμε τις εισόδους h_x και h_y με την χρήση της εντολής `filter`. Με αυτό τον τρόπο φιλτράρουμε τον αριθμητή και τον παρονομαστή με τον λευκό θόρυβο που έχουμε δημιουργήσει
- Σε αυτό το σημείο ξεκινά η διαδικασία του `normalized lms`. Ξεκινάει μια ακόμη επανάληψη με αριθμό επαναλήψεων ίσο με αυτό των αριθμών δειγμάτων του φίλτρου.
- Ορίζουμε το $u(n)$
- Το $e(n)$ παίρνει τιμή με βάση τα κάποια από τα παραπάνω στοιχεία σύμφωνα με την εξίσωση που το περιγράφει
- Έπειτα υψώνουμε στο τετράγωνο το άθροισμα των στοιχείων του $u(n)$
- Και δημιουργούμε την εξίσωση $h_{inv}(n+1)$
- Τέλος των δύο επαναλήψεων.
- Σε αυτό το σημείο υπολογίζουμε το `SER` και όταν φτάσει το επιθυμητό αριθμό σε dB το φίλτρο είναι έτοιμο.

4.3.3 Ανάλυση αποτελεσμάτων για διαφορετικές τάξεις φίλτρων.

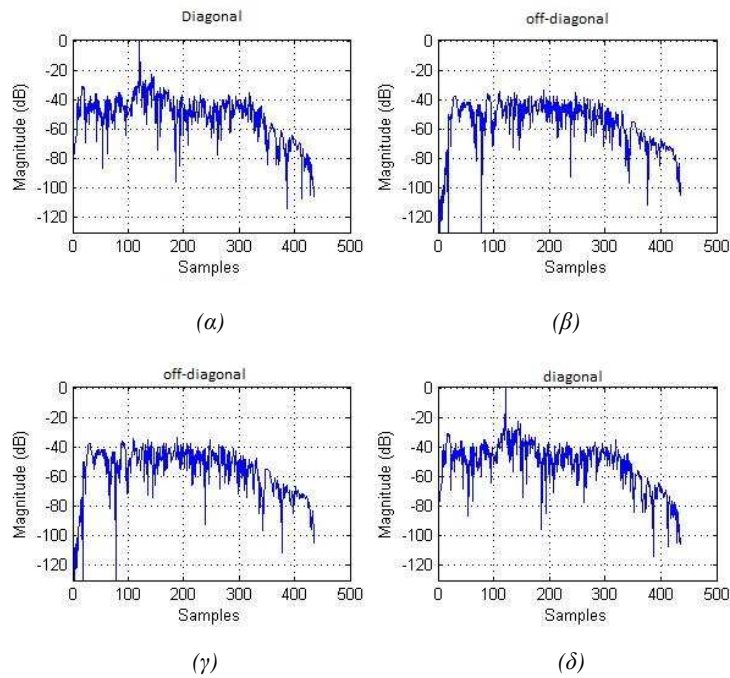
Η τάξη και η καθυστέρηση του φίλτρου μπορούν να επιλεγθούν εμπειρικά, εμείς δοκιμάσαμε διαφορετικές τιμές έως ότου πετύχουμε SER 30dB και στα τρία φίλτρα, με όσο το δυνατό ομαλότερη εξασθένιση του contralateral όρου στο συχνοτικό φάσμα. Έπειτα δοκιμάσαμε να αυξήσουμε τις τιμές δύο ακόμη φορές διατηρώντας το SER στα 30dB για να παρατηρήσουμε τυχόν αλλαγές σε σχέση με τα δείγματα και την καθυστέρηση. Στους παρακάτω πίνακες παρουσιάζονται οι τιμές δειγμάτων, καθυστέρησης και το SER που επιτεύχθηκε για τις εκάστοτε τιμές. Σε αυτό το σημείο θα επαναλάβουμε ότι το φίλτρο $f_{ii} = 1/H_i$ και το $f_{ci} = f_{cia} * f_{cib}$, όπου $f_{cia} = H_c/H_i$ και $f_{cib} = 1/H_i$.

Οι πρώτες τιμές που δοκιμάσαμε, πάρηκαν με γνώμονα ότι το φίλτρο με τα λιγότερα taps (f_{cia}) πρέπει να έχει περισσότερα δείγματα από την κρουστική.

	f_{cia}	f_{cib}	f_{ii}
Taps	130	180	309
Delay	60	60	120
Ser (dB)	24	15	15

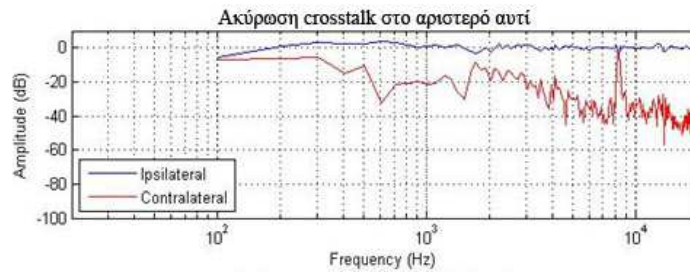
Πίνακας 2: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την πρώτη δοκιμή.

Ο χρόνος που χρειάστηκε για να ολοκληρώσει το Matlab την δημιουργία των φίλτρων ήταν 36 sec, ο χρόνος αυτός όμως, όπως και οι παρακάτω που θα αναφέρουμε δεν ανταποκρίνεται με ακρίβεια στο χρονικό διάστημα που χρειάζεται το Matlab για να ολοκληρώσει την διεργασία του αλγορίθμου διότι ο επεξεργαστής παράλληλα εκτελούσε και άλλες ενέργειες. Στην εικόνα 32 παρατηρούμε τους diagonal και off-diagonal όρους για το αριστερό κανάλι και τους off-diagonal και diagonal για το δεξί κανάλι. Όπως είχαμε αναφέρει και στο κεφάλαιο 3.3.3 ο diagonal όρος πρέπει να έχει την τιμή 1 και ο off-diagonal όρος πρέπει να έχει την τιμή 0 προκειμένου να γίνει επιτυχημένη ακύρωση crosstalk. Στην εικόνα 33 παρατηρούμε την μείωση του contralateral όρου σε σύγκριση με τον ipsilateral για το αριστερό και το δεξί αντί αντίστοιχα.

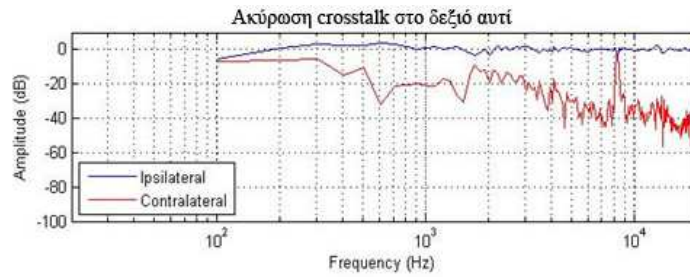


Εικόνα 32: Οι όροι diagonal και off-diagonal για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – πρώτη δοκιμή.

Όπως μπορούμε να παρατηρήσουμε στην εικόνα 32, οι diagonal όροι παίρνουν μία μεγάλη τιμή περίπου στο 110^ο δείγμα τους σε αντίθεση με τους off-diagonal. Άρα ικανοποιούν την συνθήκη, έτσι ξέρουμε ότι η ακύρωση του crosstalk θα είναι επιτυχημένη. Αυτό που δεν είναι ικανοποιητικό με τις τιμές αυτές είναι η εξασθένιση του contralateral όρου. Η λειτουργία του φίλτρου ξεκινά από τα 100Hz, έως και τα 500 Hz παρατηρούμε εξασθένιση 5dB - 15dB σε σχέση με τον ipsilateral. Από τα 500Hz και πάνω παρατηρούμε έντονες μεταβολές κατά μήκος του συχνοτικού φάσματος με μία έντονη κορυφή στα 8KHz.



(α)



(β)

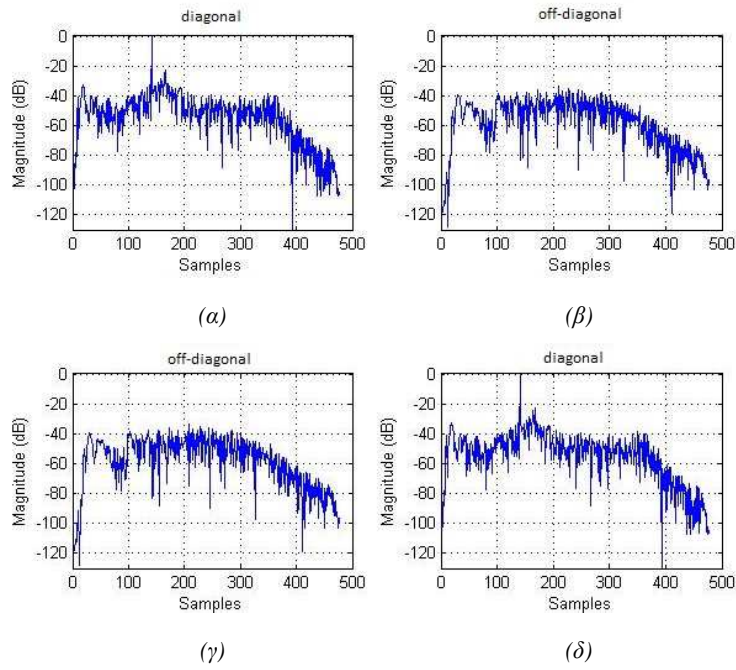
Εικόνα 33: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της πρώτης δοκιμής.

Στην δεύτερη δοκιμή οι τιμές που χρησιμοποιήσαμε ήταν οι παρακάτω

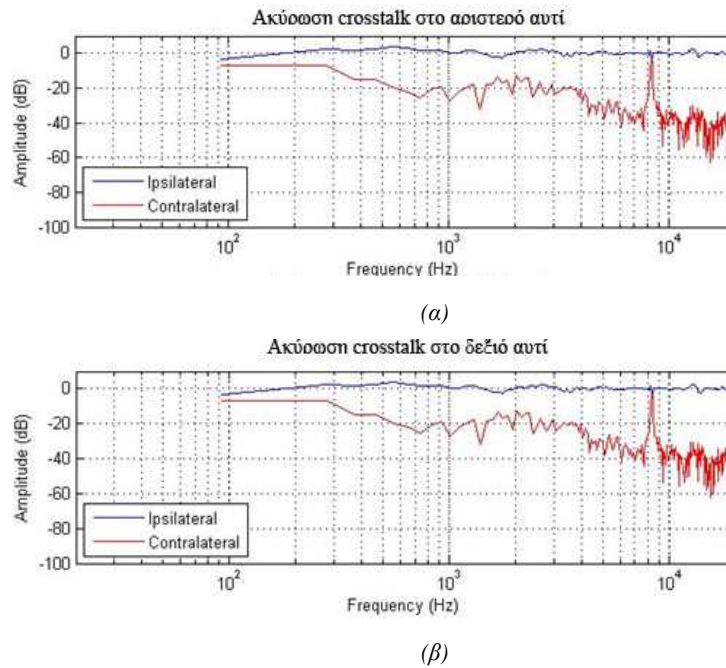
	fcia	fcib	fii
Taps	150	200	349
Delay	70	70	140
Ser (dB)	27	15	15

Πίνακας 3: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την δεύτερη δοκιμή.

Παρατηρούμε στην εικόνα 35 ότι η λειτουργία του φίλτρου ξεκινά από τα 90Hz και έχουμε μείωση του contralateral όρου από τα 500Hz έως και τα 4KHz περίπου ίση με 20dB σε σχέση με τον ipsilateral όρο. Από τα 4KHz και πάνω η μείωση είναι ικανοποιητική παρ' όλα αυτά δεν παρατηρούμε καμία μεταβολή στην κορυφή των 8KHz.



Εικόνα 34: Οι όροι *diagonal* και *off-diagonal* για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – δεύτερη δοκιμή.

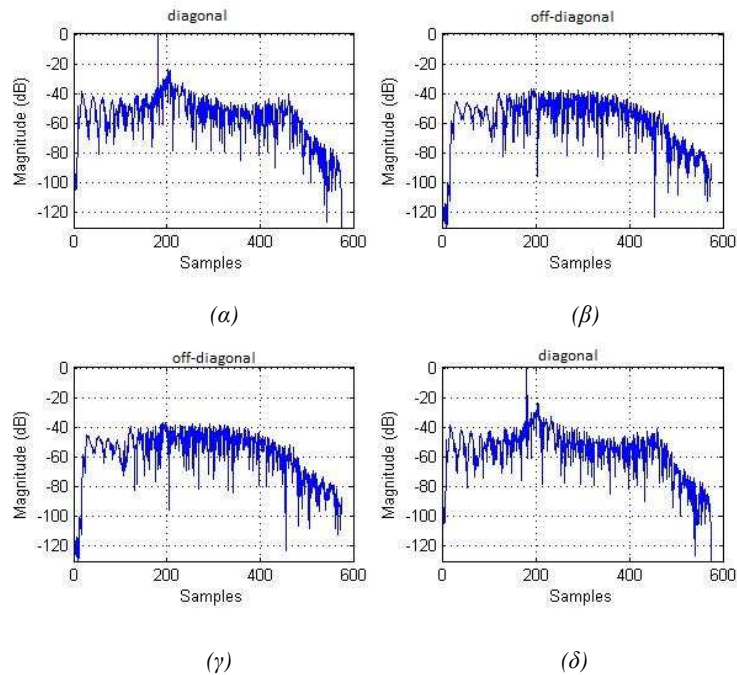


Εικόνα 35: Ακύρωση του *crosstalk* για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της δεύτερης δοκιμής.

Στην τρίτη δοκιμή οι τιμές που χρησιμοποιήσαμε ήταν οι παρακάτω

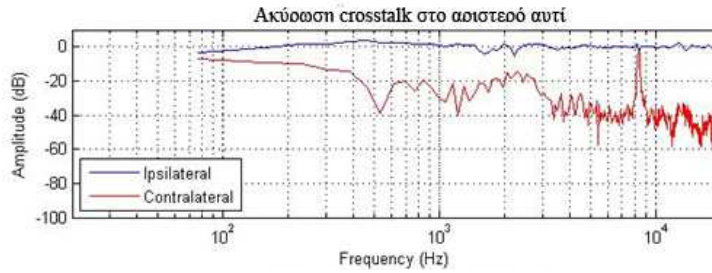
	fcia	fcib	fii
Taps	200	250	449
Delay	90	90	180
Ser (dB)	30	17	23

Πίνακας 4: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την τρίτη δοκιμή.

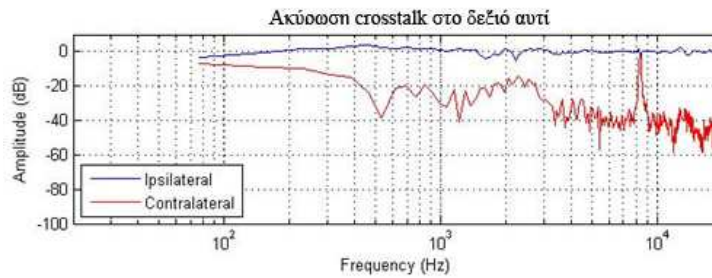


Εικόνα 36: Οι όροι *diagonal* και *off-diagonal* για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) τρίτη δοκιμή.

Όπως μπορούμε να παρατήσουμε στην εικόνα 37 η λειτουργία του φίλτρου με τις συγκεκριμένες τιμές ξεκινάει από τα 75Hz. Η εξασθένιση είναι σχετικά ίδια με τις προηγούμενες τιμές και η κορυφή στα 8KHz συνεχίζει να υπάρχει.



(α)



(β)

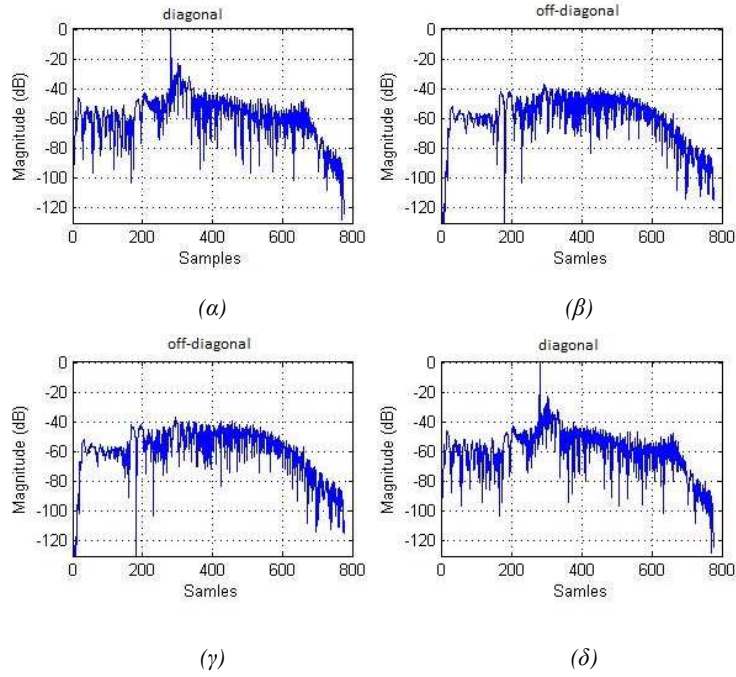
Εικόνα 37: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της τρίτης δοκιμής.

Στην τέταρτη δοκιμή οι τιμές που χρησιμοποιήσαμε ήταν οι παρακάτω

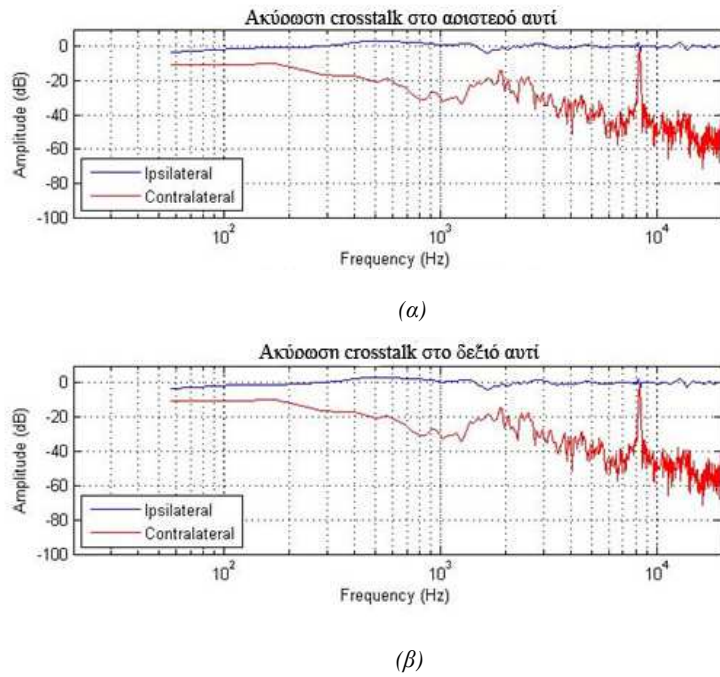
	fcia	Fcib	fii
Taps	300	350	649
Delay	140	140	280
Ser (dB)	36	20	280

Πίνακας 5: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την τέταρτη δοκιμή

Στην εικόνα 39 παρατηρούμε ότι η λειτουργία του φίλτρου με τις συγκεκριμένες τιμές ξεκινά από τα 55Hz ενώ ο contralateral όρος ξεκινάει να ελαττώνεται και στις χαμηλές συχνότητες σε σχέση με τον ipsilateral. Στο υπόλοιπο φάσμα δεν παρατηρούνται ιδιαίτερες αλλαγές σε σχέση με τις προηγούμενες τιμές όσον αφορά την διαφορά έντασης.



Εικόνα 38: Οι όροι diagonal και off-diagonal για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – τέταρτη δοκιμή

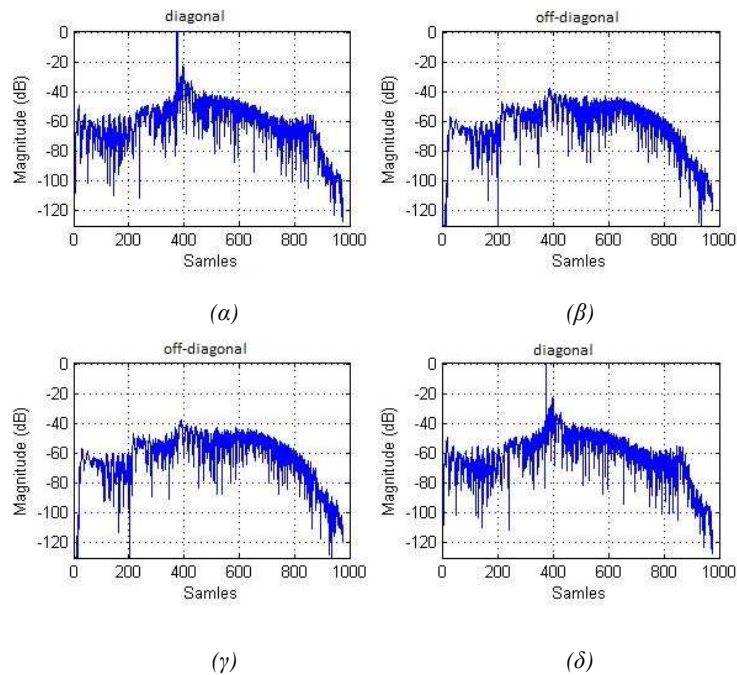


Εικόνα 39: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της τέταρτης δοκιμής.

Στην πέμπτη δοκιμή οι τιμές που χρησιμοποιήσαμε είναι οι παρακάτω

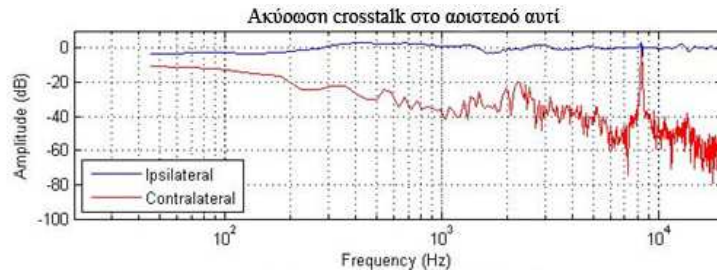
	fcia	fcib	fii
Taps	400	450	849
Delay	187	187	374
Ser (dB)	40	24	30

Πίνακας 6: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την πέμπτη δοκιμή.

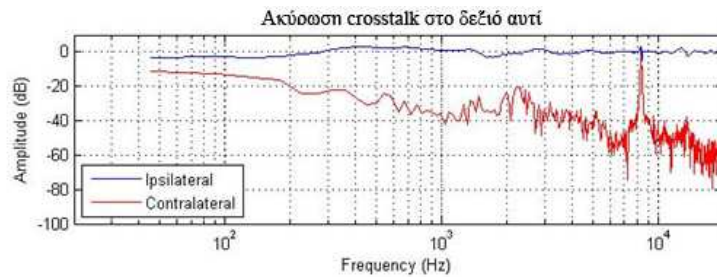


Εικόνα 40: Οι όροι *diagonal* και *off-diagonal* για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – πέμπτη δοκιμή .

Στην εικόνα 41 παρατηρούμε ότι η λειτουργία του φίλτρου ξεκινάει από τα 45Hz. Παρατηρείται βελτίωση διαφοράς έντασης και στις χαμηλές συχνότητες της τάξης των 15dB. Από τα 200Hz και πάνω η διαφορά του *contralateral* όρου σε σχέση με τον *ipsilateral* είναι μεγαλύτερη των 20dB με την κορυφή των 8 KHz να συνεχίζει να υπάρχει με πολύ μικρότερο εύρος.



(α)



(β)

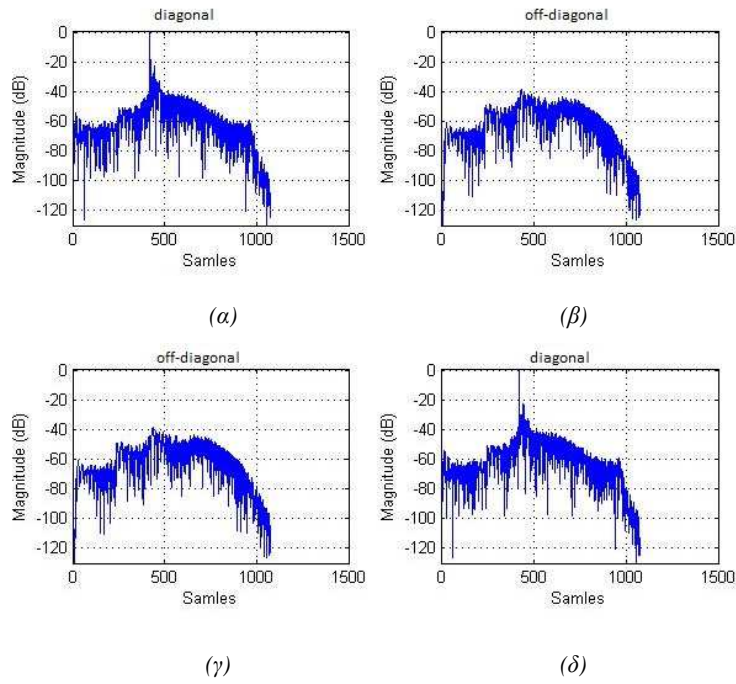
Εικόνα 41: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της πέμπτης δοκιμής.

Στην έκτη δοκιμή οι τιμές που χρησιμοποιήσαμε είναι οι παρακάτω

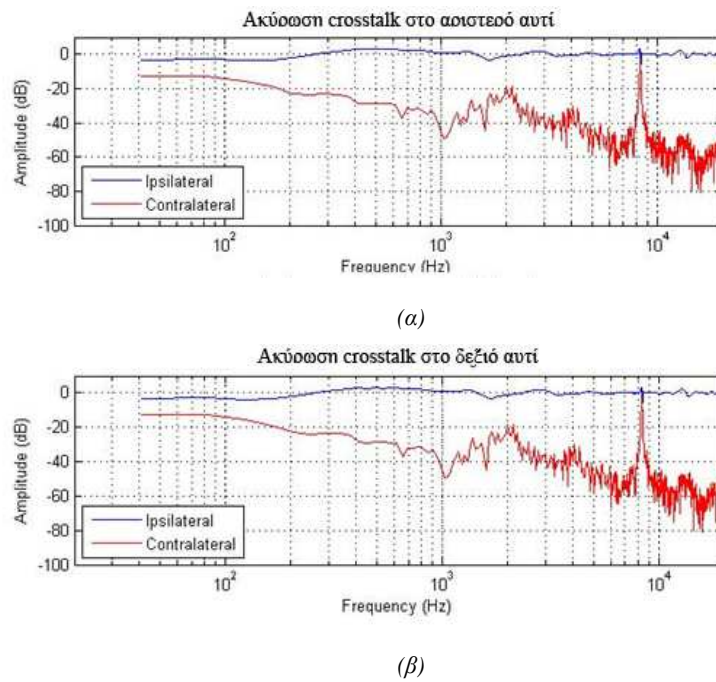
	fcia	fcib	fii
Taps	450	500	949
Delay	210	210	420
Ser (dB)	41	26	26

Πίνακας 7: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την έκτη δοκιμή.

Στην εικόνα 43 παρατηρούμε ότι το αποτέλεσμα μείωσης του contralateral όρου είναι σχετικά ίδιο με τις προηγούμενες τιμές με διαφορά ότι από το 1KHz και κάτω η καμπύλη είναι πιο ομαλή σε σχέση με πριν. Η συχνότητα από την οποία ξεκινά να επιδρά το φίλτρο έχει βελτιωθεί στα 40Hz.



Εικόνα 42: Οι όροι *diagonal* και *off-diagonal* για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – έκτη δοκιμή.

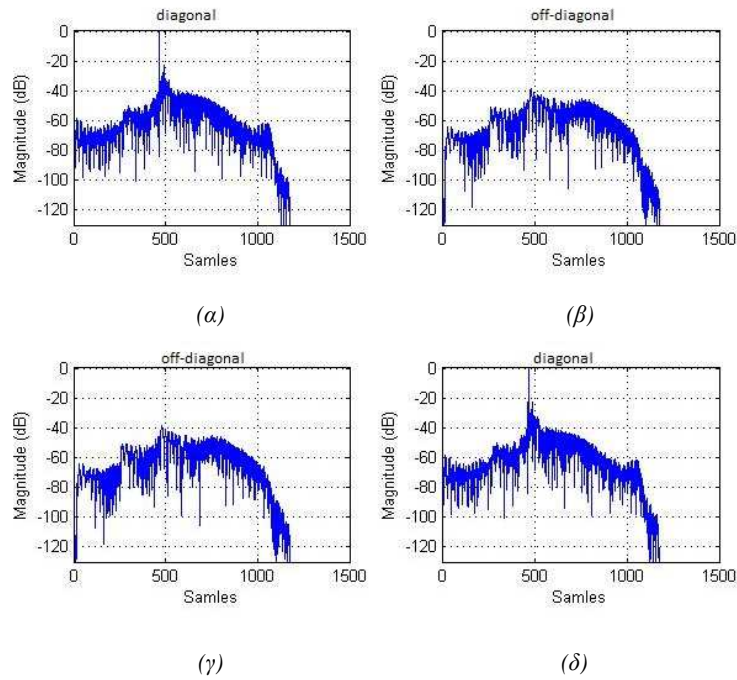


Εικόνα 43: Ακύρωση του *crosstalk* για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της έκτης δοκιμής.

Στην έβδομη δοκιμή οι τιμές που χρησιμοποιήσαμε είναι οι παρακάτω

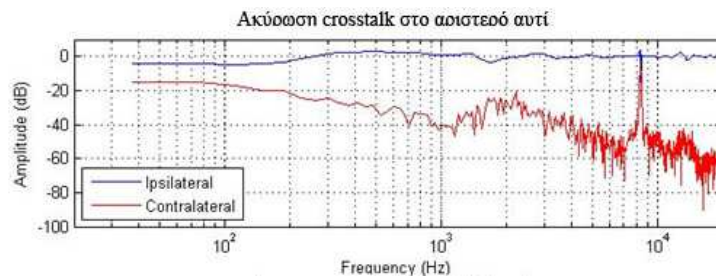
	fcia	fcib	fii
Taps	500	550	1049
Delay	233	233	466
Ser (dB)	42	27	30

Πίνακας 8: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την έβδομη δοκιμή.

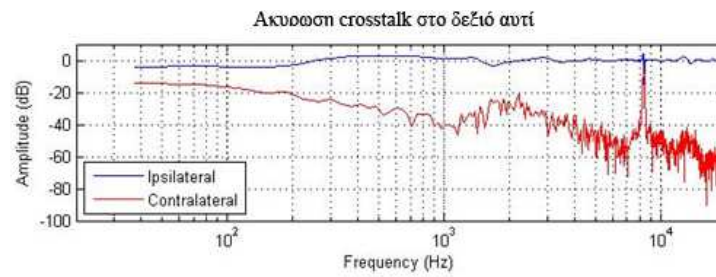


Εικόνα 44: Οι όροι *diagonal* και *off-diagonal* για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – έβδομη δοκιμή.

Στην εικόνα 45 παρατηρούμε ότι το φίλτρο ξεκινά να επιδρά περίπου από τα 35Hz. Σε σχέση με τις προηγούμενες τιμές παρατηρούμε βελτίωση στην διαφορά έντασης στην συχνοτική περιοχή των 1500Hz.



(α)



(β)

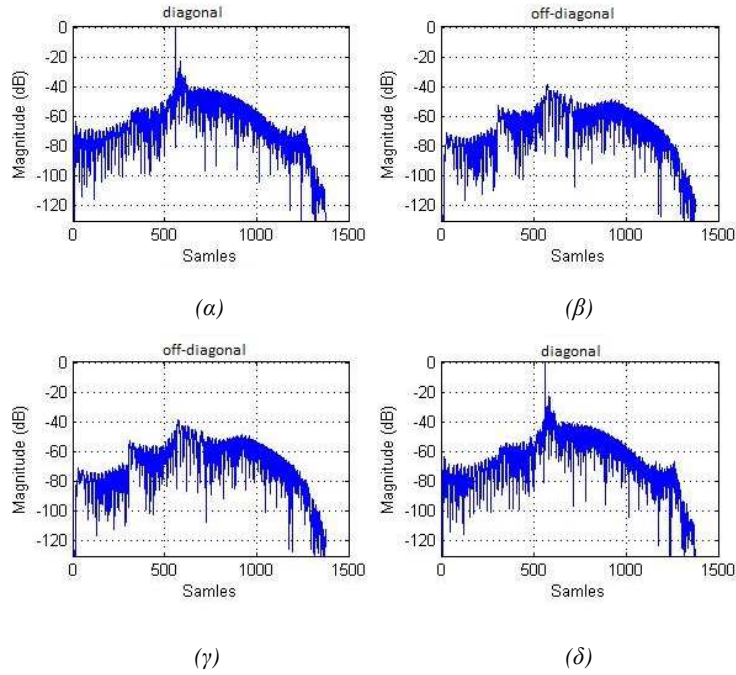
Εικόνα 45: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της έβδομης δοκιμής.

Στην όγδοη δοκιμή οι τιμές που χρησιμοποιήσαμε είναι οι παρακάτω

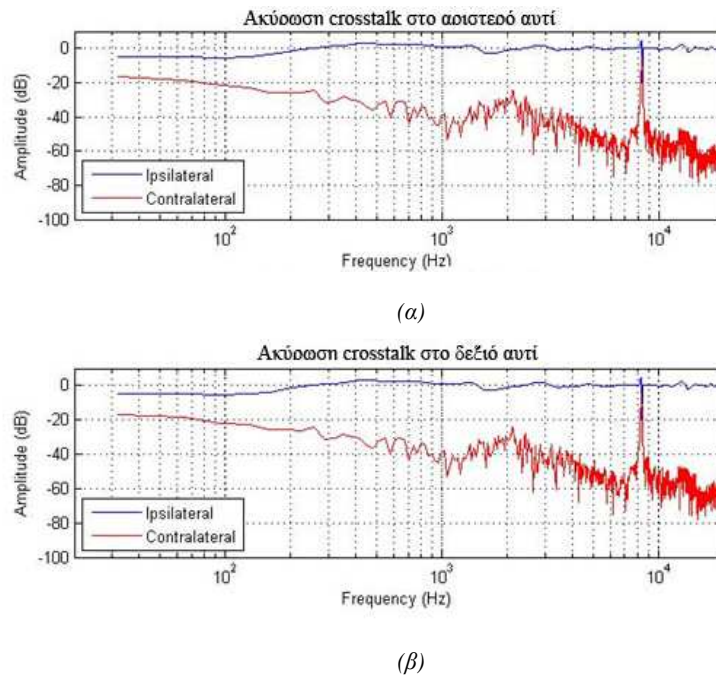
	fcia	fcib	fii
Taps	600	650	1249
Delay	280	280	560
Ser (dB)	44	30	30

Πίνακας 9: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την όγδοη δοκιμή.

Με τις συγκεκριμένες τιμές παρατηρούμε ότι είμαστε σε θέση να επιτύχουμε SER 30dB για όλα τα φίλτρα. Η επίδραση του φίλτρου ξεκινά λίγο πιο μετά από τα 30HZ όπως μπορούμε να παρατηρήσουμε και στην εικόνα 47.



Εικόνα 46: Οι όροι diagonal και off-diagonal για το αριστερό (α, β) και το δεξί κανάλι (γ, δ) – όγδοη δοκιμή.

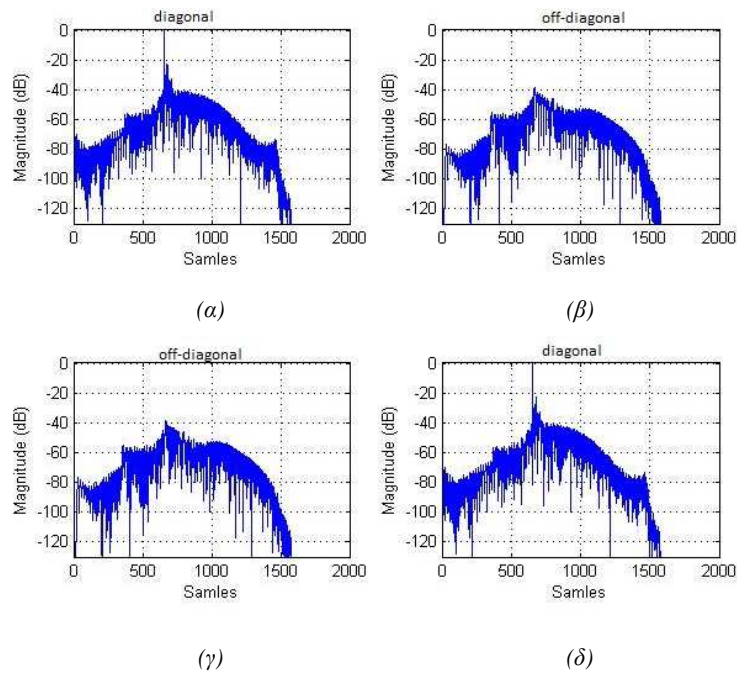


Εικόνα 47: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της όγδοης δοκιμής.

Στην ένατη δοκιμή οι τιμές που χρησιμοποιήσαμε είναι οι παρακάτω

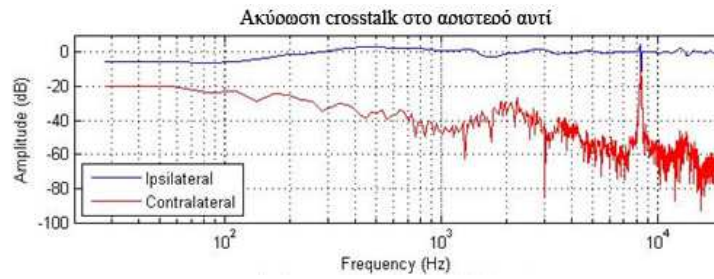
	fcia	fcib	fii
Taps	700	750	1449
Delay	327	327	654
Ser (dB)	44	30	30

Πίνακας 10: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του αναστροφου φίλτρου για την ένατη δοκιμή.

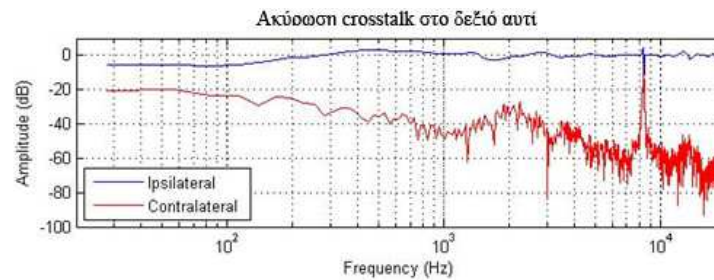


Εικόνα 48: Οι όροι diagonal και off-diagonal για το αριστερό (α, β) και το δεξιό κανάλι (γ, δ) – ένατη δοκιμή.

Σε αυτή την δοκιμή παρατηρούμε στην εικόνα 49 βλέπουμε ότι παρόλο που οι τιμές του SER που επιλέξαμε είναι ίδιες με την προηγούμενη δοκιμή, η μείωση του contralateral όρου στο συχνοτικό φάσμα φαίνεται να είναι ίδια με εξαίρεση την συχνότητα των 3KHz όπου φαίνεται να εξασθενεί περισσότερο. Τέλος η επίδραση του φίλτρου φαίνεται ότι ξεκινάει από ακόμα χαμηλότερα επίπεδα.



(α)



(β)

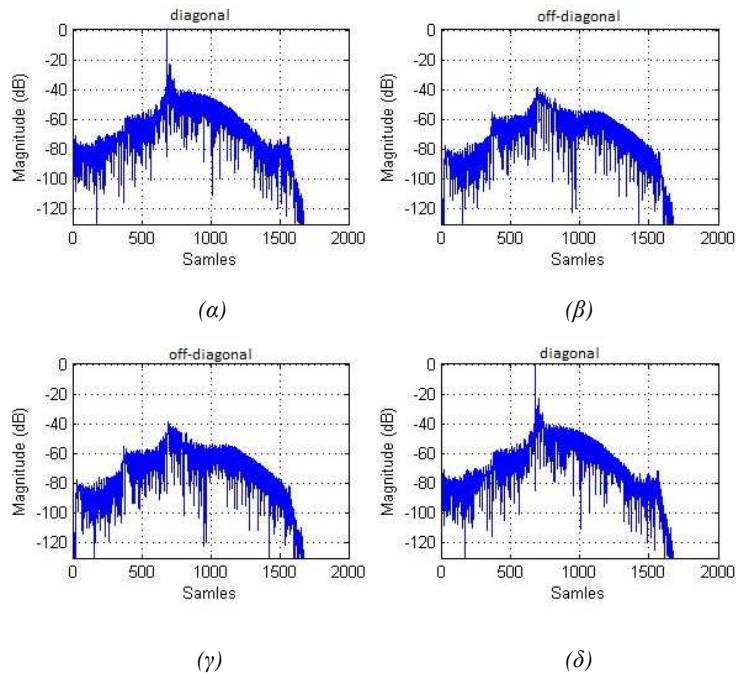
Εικόνα 49: Ακύρωση του crosstalk για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της ένατης δοκιμής.

Στην δέκατη δοκιμή οι τιμές που χρησιμοποιήσαμε είναι οι παρακάτω

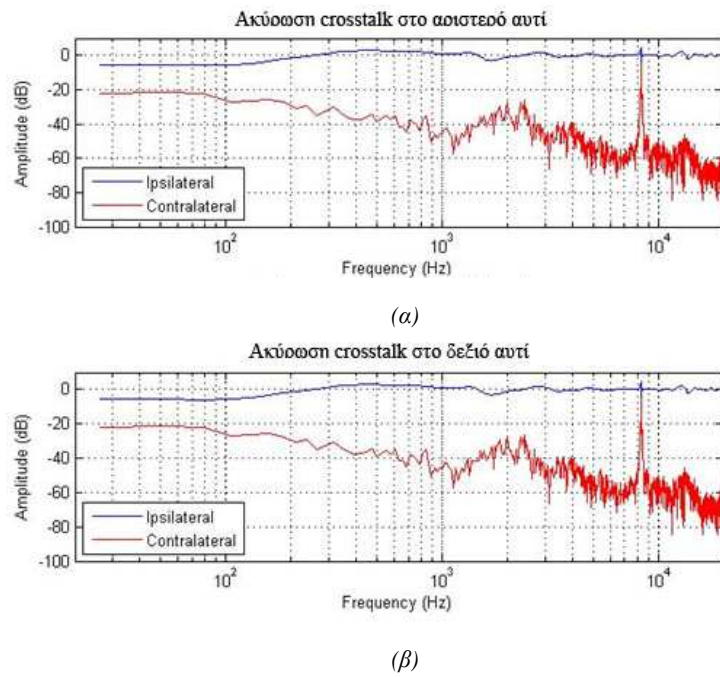
	fcia	fcib	fii
Taps	750	800	1549
Delay	340	340	680
Ser (dB)	44	30	30

Πίνακας 11: Πίνακας παρουσίασης πειραματικών τιμών των παραμέτρων του ανάστροφου φίλτρου για την δέκατη δοκιμή.

Στην εικόνα 51 παρατηρούμε ότι η επίδραση του φίλτρου ξεκινάει ακόμα πιο χαμηλά από ότι πριν γεγονός που δεν έχει μεγάλη σημασία για τόσο χαμηλές συχνότητες. Αυτό που έχει ενδιαφέρον είναι ότι η εξασθένιση που είχαμε στις προηγούμενες τιμές για την συχνότητα των 3KHz στην προηγούμενη δοκιμή φαίνεται να έχει μειωθεί ελάχιστα.



Εικόνα 50: Οι όροι *diagonal* και *off-diagonal* για το αριστερό (A-B) και το δεξί κανάλι (Γ-Δ) – δέκατη δοκιμή.



Εικόνα 51: Ακύρωση του *crosstalk* για το αριστερό αυτί (α) και το δεξιό αυτί (β) με το ανάστροφο φίλτρο της δέκατης δοκιμής.

4.3.4 Επίλογος

Σε αυτό το κεφάλαιο αναλύσαμε την κατασκευή του αλγορίθμου και παρατηρήσαμε τις τιμές των αποτελεσμάτων για διαφορετικές τιμές δειγμάτων, καθυστέρησης και το δυνατό SER που μπορεί να επιτευχθεί με αυτές. Η ακύρωση crosstalk που επιτεύχθηκε στο συχνοτικό φάσμα κυμαίνεται από -20dB με -60dB ξεκινώντας από της χαμηλές προς της υψηλές περιοχές σε σχέση με τον ipsilateral όρο. Ωστόσο στην συχνότητα των 8KHz παρατηρήθηκε μία κορυφή όπου με την αύξηση των τιμών, πετύχαμε να μειώσουμε το εύρος επίδρασης.

5 Πειραματική διαδικασία.

5.1 Εισαγωγή.

Σε αυτό το κεφάλαιο θα παρουσιαστούν η διάταξη της πειραματικής διαδικασίας, τα χαρακτηριστικά των ακροατών, που πήραν μέρος στο πείραμα, τα HRTFs τα οποία χρησιμοποιήθηκαν για την δημιουργία των ανάστροφων φίλτρων, ο τρόπος δημιουργίας των binaural αρχείων τα οποία χρησιμοποιήθηκαν για τις ακροάσεις και η στατιστική ανάλυση των αποτελεσμάτων του πειράματος.

5.2 Διάταξη πειραματικής διαδικασίας.

Η πειραματική διαδικασία πραγματοποιήθηκε στο εργαστήριο πολυφωνίας και διάδρασης στο Τ.Ε.Ι Κρήτης, παράρτημα Ρεθύμνου, τμήμα μουσικής τεχνολογίας και ακουστικής. Ο εξοπλισμός που χρησιμοποιήθηκε κατά την πειραματική διαδικασία ήταν:

- Υπολογιστής (macbook pro)
- Κάρτα ήχου (motu traveler)
- 3 ηχεία (genelec 8020)

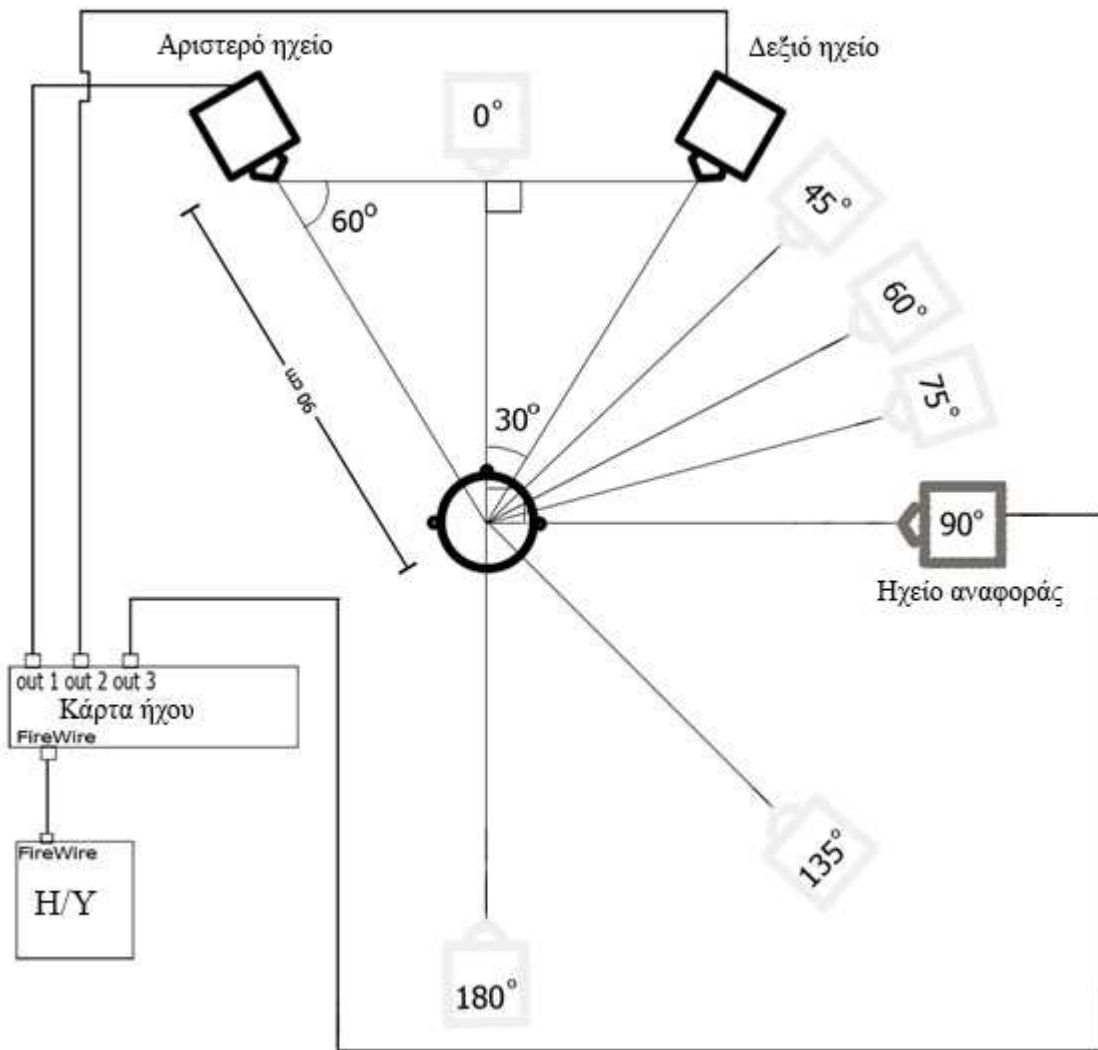
Η διάταξη του ηλεκτροακουστικού συστήματος έχει ως εξής: δύο ηχεία τοποθετημένα δεξιά και αριστερά σε γωνία ± 30 μοιρών σε σχέση με την θέση του ακροατή (διάταξη stereo αναπαραγωγής) σε απόσταση 90cm από τον ακροατή. Επιπροσθέτως χρησιμοποιήθηκε ένα επιπλέον, ίδιου τύπου, ηχείο το οποίο είχε τον ρόλο της πραγματικής ηχητικής πηγής, για σύγκριση από τον ακροατή της θέσης της εκάστοτε εικονικής ηχητικής πηγής με την πραγματική. Το ηχείο αναφοράς τοποθετήθηκε στις γωνίες των 0, 30, 45, 60, 75, 90, 135 και 180 μοιρών. Αντίστοιχα με τις γωνίες των binaural αρχείων ήχου τα οποία αναπαράχθηκαν από τα ηχεία της stereo διάταξης. Η διαδικασία της ακρόασης περιλάμβανε αναπαραγωγή τριών αρχείων ήχου για κάθε γωνία. Το πρώτο αρχείο ήχου ήταν μονοφωνικό και αναπαράχθηκε από το ηχείο αναφοράς για κάθε γωνία που αναφέρθηκε παραπάνω. Το δεύτερο και το τρίτο αρχείο ήχου ήταν binaural αρχεία στα οποία η εικονική ηχητική πηγή ήταν προγραμματισμένη να ηχεί στις γωνίες που αναφέραμε παραπάνω και επιπλέον ήταν φιλτραρισμένα με τα ανάστροφα φίλτρα τα οποία κατασκευάστηκαν με τις τεχνικές της γρήγορης αποσυνέλιξης με κανονικοποίηση

και των ελαχίστων μέσων τετραγώνων αντίστοιχα, για την ακύρωση του crosstalk και της γωνίας της θέσης των ηχείων. Από τους ακροατές ζητήθηκε να συγκρίνουν την θέση των δύο εικονικών ηχητικών πηγών με την ηχητική πηγή αναφοράς και συγκρίνοντας της δύο εικονικές ηχητικές πηγές να ορίσουν ποια από τις δύο είναι πλησιέστερη όσον αφορά την θέση στον χώρο με την πραγματική. Τα αρχεία ήχου που χρησιμοποιήθηκαν κατά την πειραματική διαδικασία ήταν ροζ θόρυβος. Επιλέξαμε ροζ θόρυβο λόγω της ιδιότητας του στην ίση κατανομή ενέργειας ανά οκτάβα. Στην εικόνα 52 παρουσιάζεται η πειραματική διάταξη.

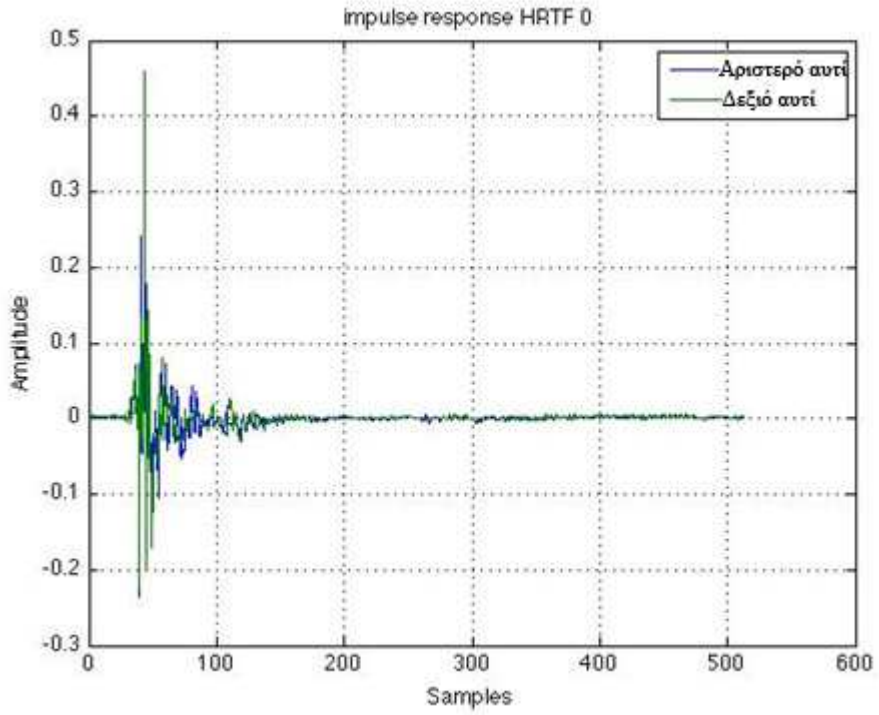
5.3 Δημιουργία binaural αρχείου ήχου.

Το binaural σήμα μπορεί να δημιουργηθεί με δύο τρόπους. Ο ένας τρόπος είναι να πραγματοποιηθεί binaural ηχογράφηση με ένα dummy head, το οποίο θα έχει προσαρμοσμένα δύο παντοκατευθυντικά (omni) μικρόφωνα στο ακουστικό κανάλι. Ο άλλος τρόπος είναι με τη χρήση της συνέλιξης (convolution). Παίρνοντας ένα μονοφωνικό σήμα και κάνοντας το συνέλιξη με μία κρουστική απόκριση ενός dummy head δημιουργείται ένα νέο σήμα, το οποίο έχει τα χαρακτηριστικά πλάτους και φάσης, τα οποία θα είχαν εισαχθεί στο σήμα αν αντί για ένα μικρόφωνο, είχε χρησιμοποιηθεί ένα dummy head για την ηχογράφηση του.

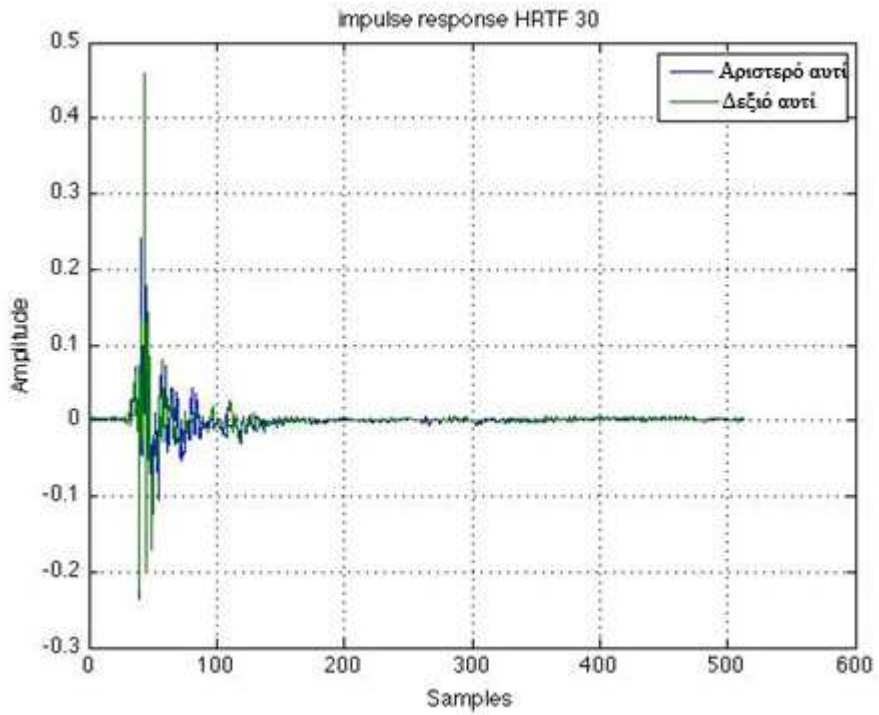
Στη συγκεκριμένη διπλωματική εργασία, για τις ανάγκες του πειράματος χρησιμοποιήθηκε η τεχνική της binaural synthesis, για τη δημιουργία των binaural ηχητικών παραδειγμάτων. Χρησιμοποιήθηκε ένα μονοφωνικό ηχητικό αρχείο ροζ θορύβου, όπως αναφέρθηκε στο κεφάλαιο 4, στο οποίο εφαρμόστηκε συνέλιξη με τα HRTFs των επιθυμητών γωνιών. Για τη διαδικασία της συνέλιξης χρησιμοποιήθηκε η εντολή conv του Matlab. Στις εικόνες 53 ως 60 παρατίθενται τα HRTFs.



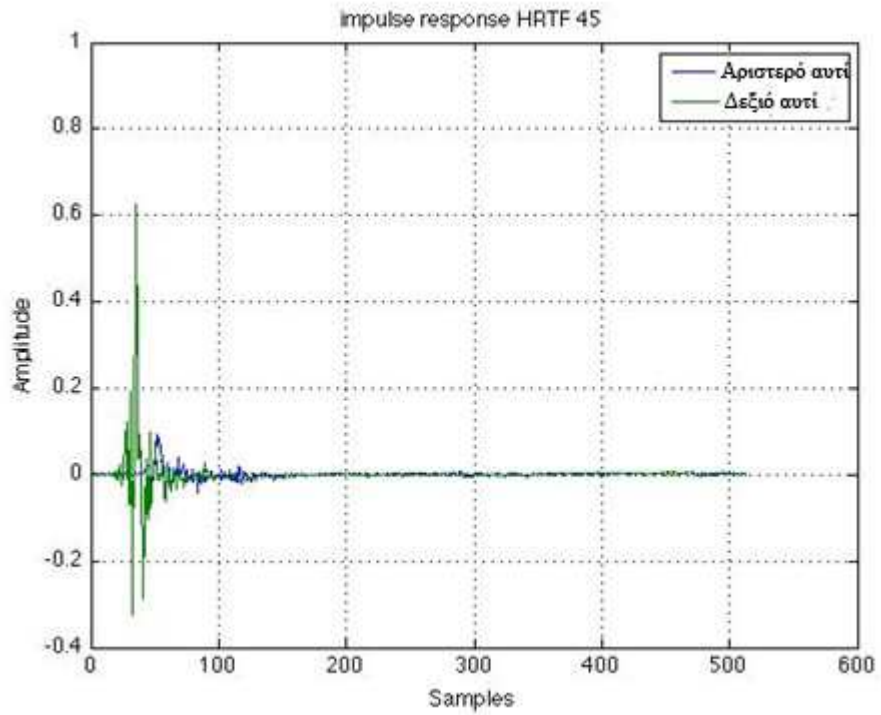
Εικόνα 52: Πειραματική διάταξη. Διακρίνονται τα ηχεία αναπαραγωγής του binaural σήματος σε stereo διάταξη (αριστερό ηχείο, δεξί ηχείο) και το μετακινούμενο ηχείο (ηχείο αναφοράς) στην εκάστοτε γωνία σπού ηχεί η εικονική πηγή.



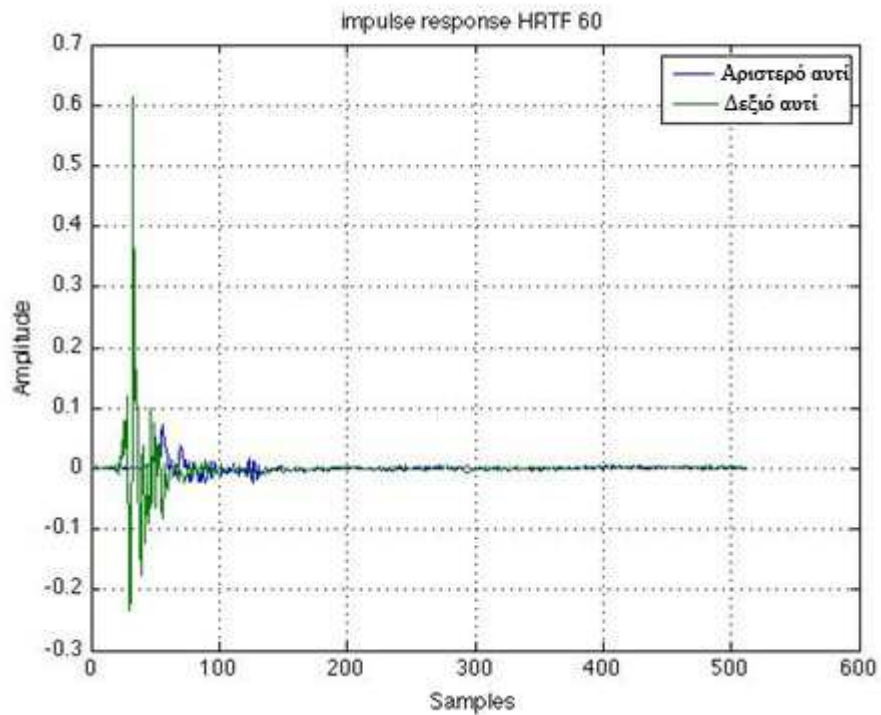
Εικόνα 53: Η κρουστική απόκριση 0° των HRTFs στο πεδίο του χρόνου (IRCAM database).



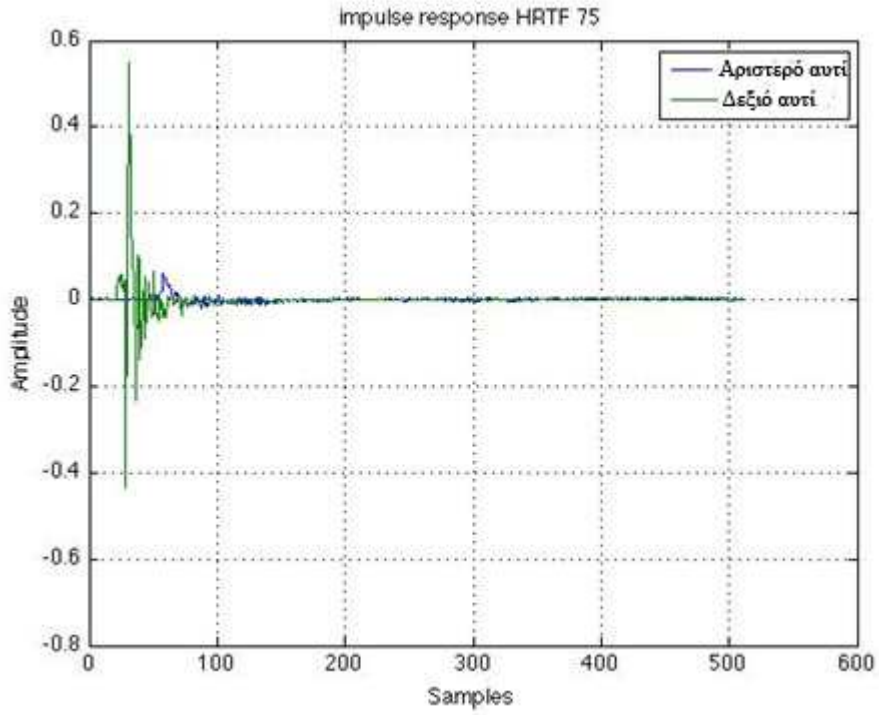
Εικόνα 54: Η κρουστική απόκριση 30° των HRTFs στο πεδίο του χρόνου (IRCAM database).



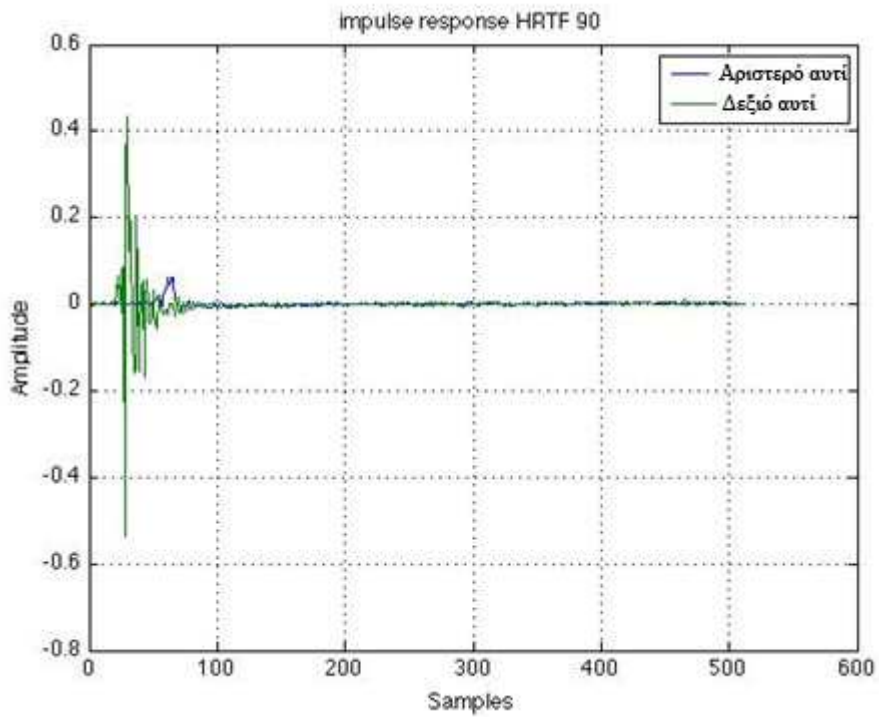
Εικόνα 55: Η κρουστική απόκριση 45° των HRTFs στο πεδίο του χρόνου (IRCAM database).



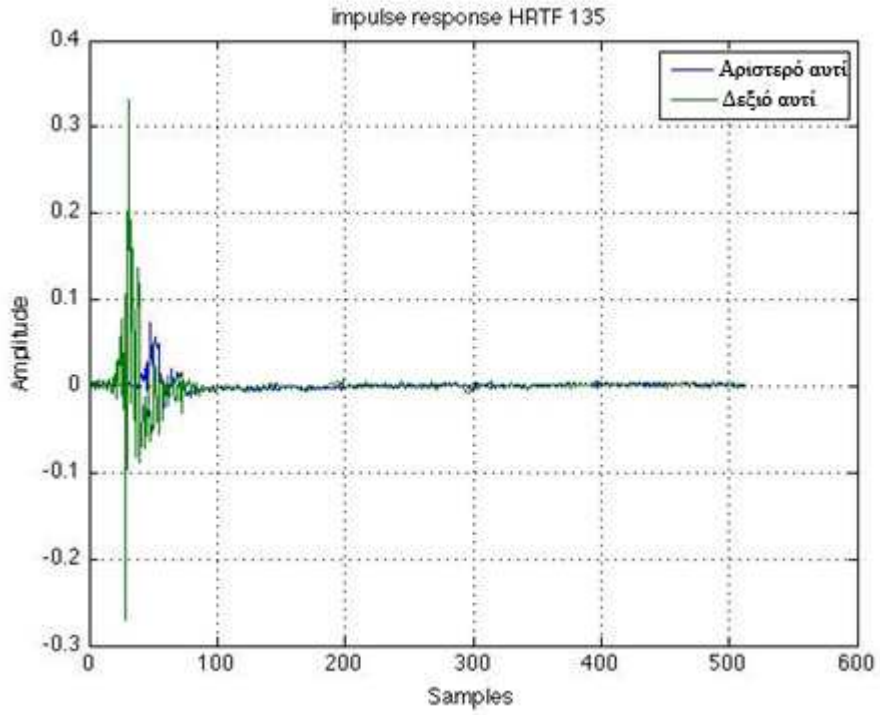
Εικόνα 56: Η κρουστική απόκριση 60° των HRTFs στο πεδίο του χρόνου (IRCAM database).



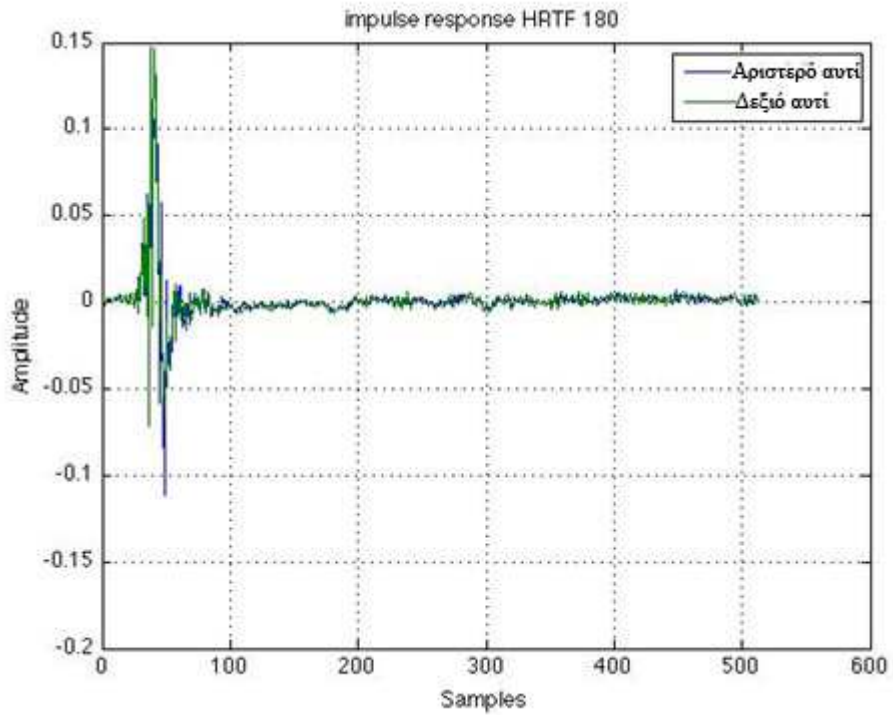
Εικόνα 57: Η κρουστική απόκριση 75° των HRTFs στο πεδίο του χρόνου (IRCAM database).



Εικόνα 58: Η κρουστική απόκριση 90° των HRTFs στο πεδίο του χρόνου (IRCAM database).



Εικόνα 59: Η κρουστική απόκριση 135° των HRTFs στο πεδίο του χρόνου (IRCAM database).



Εικόνα 60: Η κρουστική απόκριση 180° των HRTFs στο πεδίο του χρόνου (IRCAM database).

5.4 Χαρακτηριστικά των ακροατών που έλαβαν μέρος στην πειραματική διαδικασία.

Οι ακροατές που έλαβαν μέρος στην πειραματική διαδικασία επιλέχθηκαν τυχαία. Ο αριθμός των ακροατών ανέρχεται στους 15. Στον πίνακα 12 παρατίθενται οι διαστάσεις της κεφαλής του εκάστοτε ακροατή.

A/A	Πλάτος (cm)	Ύψος (cm)	Βάθος (cm)
1	15.50	21.00	20.50
2	16.50	21.50	18.00
3	15.90	22.00	21.00
4	15.80	20.50	19.00
5	16.00	22.50	22.00
6	14.80	22.00	21.00
7	16.00	22.00	19.00
8	16.70	23.00	20.00
9	15.00	21.00	19.00
10	16.00	22.50	20.00
11	16.70	22.00	21.50
12	15.50	22.50	23.00
13	16.10	22.00	22.50
14	16.50	23.50	20.00
15	17.00	23.00	22.00
M.O	16.00	22.06	20.56

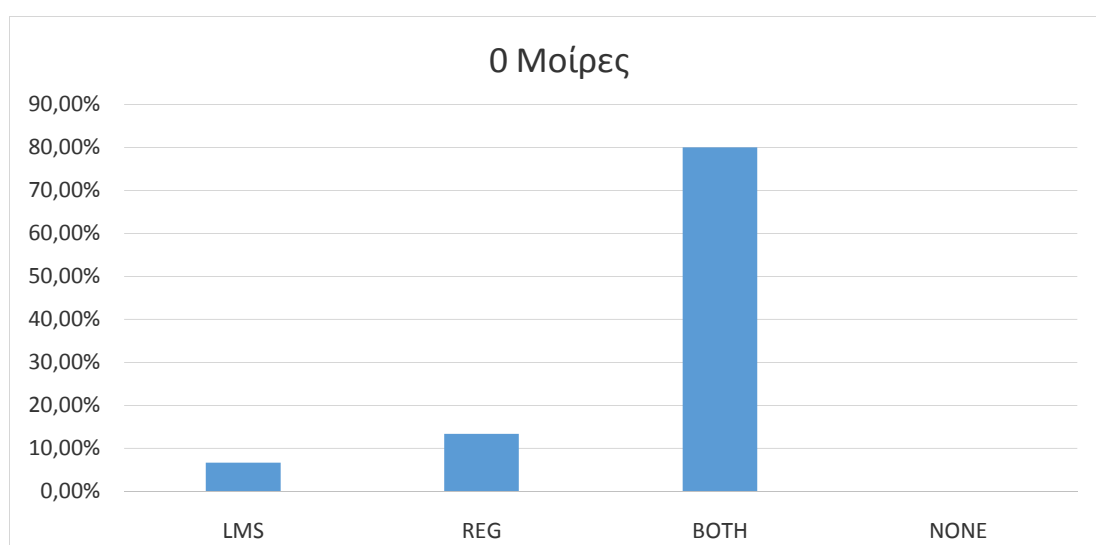
Πίνακας 12: Παρουσίαση των διαστάσεων της κεφαλής του κάθε ακροατή που έλαβε μέρος στην πειραματική διαδικασία και ο μέσος όρος των διαστάσεων του συνόλου των ακροατών.

Για τη δημιουργία του ανάστροφου φίλτρου, το οποίο δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, χρησιμοποιήθηκε κρουστική απόκριση HRTF της γωνίας των 30° από τη βάση δεδομένων του IRCAM. Αντίστοιχα, για τη δημιουργία του ανάστροφου φίλτρου με τη μέθοδο των ελαχίστων μέσω των τετραγώνων, χρησιμοποιήθηκε κρουστική απόκριση HRTF της γωνίας των 30°, η οποία δημιουργήθηκε από το τεχνητό κεφάλι KEMAR (Knowles Electronics Manikin for Acoustic Research). Στην περίπτωση του ανάστροφου φίλτρου το οποίο δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσω των τετραγώνων χρησιμοποιήθηκε η συγκεκριμένη κρουστική απόκριση λόγω του μικρού μήκους δειγμάτων. Όπως

αναφέρθηκε στο κεφάλαιο 4.3.2, το μήκος της κρουστικής απόκρισης είναι καθοριστικό όσον αφορά στο χρόνο που απαιτείται για την ολοκλήρωση της επεξεργασίας από τον αλγόριθμο, για τη δημιουργία του ανάστροφου φίλτρου. Στην περίπτωση του ανάστροφου φίλτρου, το οποίο δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, δεν χρησιμοποιήθηκε η κρουστική απόκριση του KEMAR διότι προκαλούσε την έντονη παρουσία συχνότητας κουδουνίσματος κατά την ακρόαση. Χρησιμοποιήθηκε η κρουστική απόκριση από τη βάση δεδομένων του IRCAM διότι το ανάστροφο φίλτρο ήταν εξίσου αποδοτικό όσον αφορά την ακύρωση του crosstalk και της θέσης των ηχείων και επιπροσθέτως παρουσίαζε σημαντική μείωση στη συχνότητα κουδουνίσματος.

5.5 Ανάλυση αποτελεσμάτων.

Στην απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 0°.

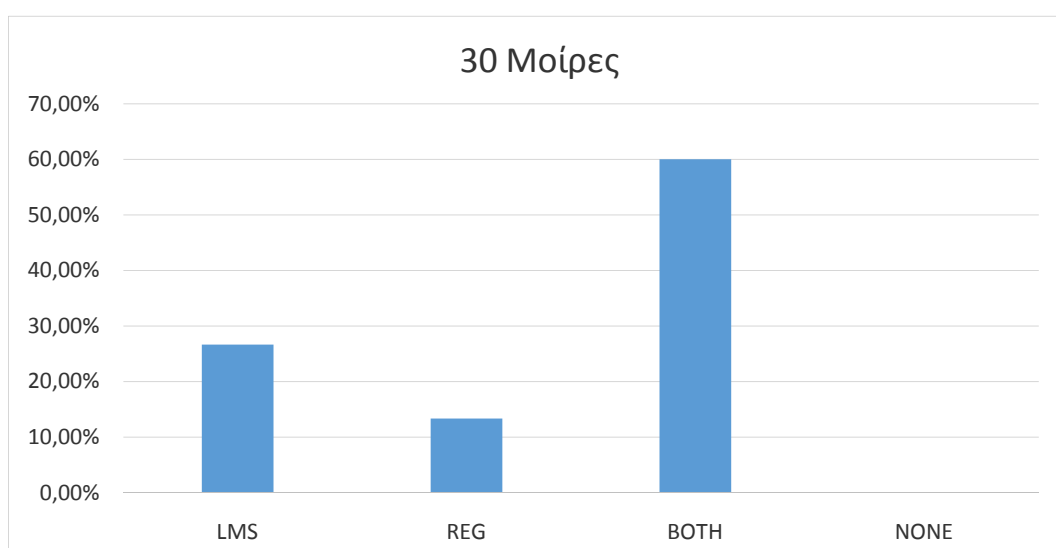


Εικόνα 61: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 0°

Παρατηρώντας την εικόνα 61 συμπεραίνουμε ότι το 80% των ακροατών, οι οποίοι συμμετείχαν στο πείραμα, αντιλήφθηκαν την εικονική ηχητική πηγή στην ίδια θέση σε σχέση με το ηχείο αναφοράς, το οποίο συμβολίζει την πραγματική θέση της ηχητικής πηγής, και για τις δύο μεθόδους δημιουργίας ανάστροφου φίλτρου. Το 13.33% των ακροατών αντιλήφθηκε την εικονική ηχητική η, οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε

με τη μέθοδο των ελαχίστων μέσω τετραγώνων. Το 6.67% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσω τετραγώνων, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Άρα το μεγαλύτερο ποσοστό, με διαφορά, επιβεβαίωσε την αποτελεσματικότητα και των δύο μεθόδων, δημιουργίας ανάστροφου φίλτρου, στη γωνία των 0°.

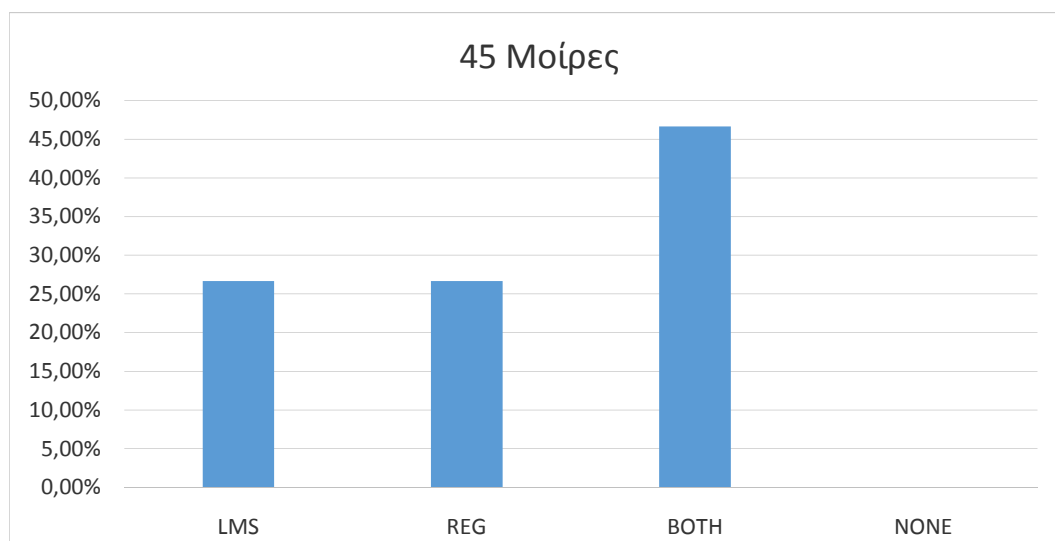
Στην απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 30°.



Εικόνα 62: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 30°

Στην εικόνα 62 βλέπουμε ότι το 60% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή στην ίδια θέση με το ηχείο αναφοράς και για τις δύο μεθόδους δημιουργίας ανάστροφου φίλτρου. Το 26.67% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσω τετραγώνων, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Το 13.33% των ακροατών αντιλήφθηκε την εικονική ηχητική η, οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσω τετραγώνων. Άρα και για τη γωνία των 30°, το μεγαλύτερο ποσοστό των ακροατών επιβεβαίωσε την αποτελεσματική λειτουργία και των δύο μεθόδων.

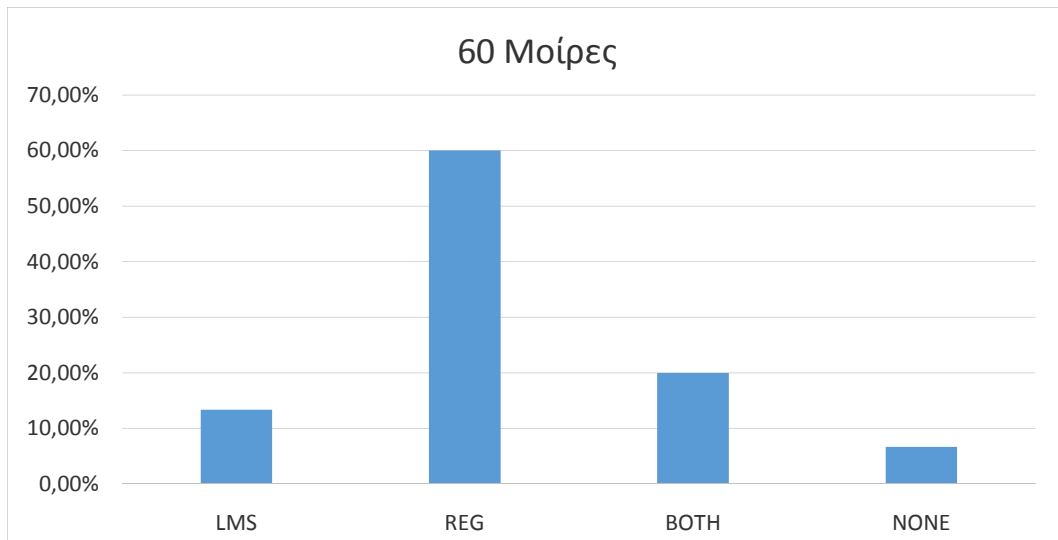
Στην εικόνα 63 απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 45° .



Εικόνα 63: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 45°

Σύμφωνα με το γράφημα της εικόνας 63, το 46.67% των ακροατών αντιλήφθηκαν την εικονική ηχητική πηγή στην ίδια θέση με το ηχείο αναφοράς και για τις δύο μεθόδους δημιουργίας ανάστροφου φίλτρου. Το 26.67% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Το ίδιο ποσοστό των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων. Στην περίπτωση αυτή παρατηρείται ότι, ομοίως με τις 0° και τις 30°, το μεγαλύτερο ποσοστό των ακροατών επιβεβαίωσε την αποτελεσματικότητα και των δύο αλγορίθμων.

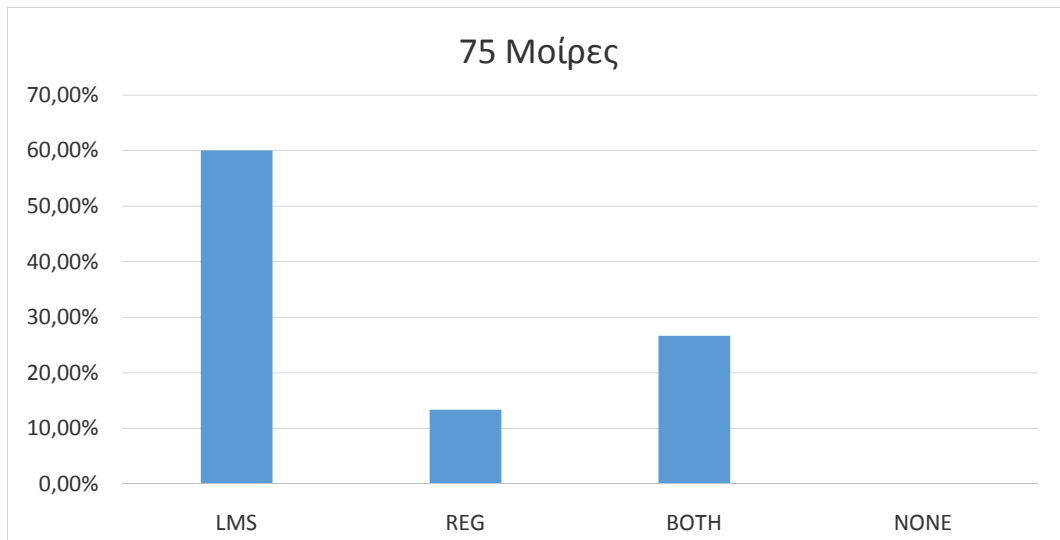
Στην εικόνα 64 απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 60°.



Εικόνα 64: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 60°

Το 60% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων. Ένα ποσοστό ακροατών της τάξης του 20% αντιλήφθηκε την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει εξίσου με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, στο ηχείο αναφοράς. Μικρότερο ποσοστό ακροατών, της τάξης του 13.33%, αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Τέλος το 7% των ακροατών δεν αντιλήφθηκε την εικονική ηχητική πηγή στην επιθυμητή θέση σε σχέση με το ηχείο αναφοράς για καμία από τις δύο μεθόδους. Παρατηρείται ότι στην περίπτωση των 60° το μεγαλύτερο ποσοστό των ακροατών αντιλήφθηκε, κατά την ακρόαση, ότι η μέθοδος για τη δημιουργία ανάστροφου φίλτρου, της γρήγορης αποσυνέλιξης με κανονικοποίηση λειτουργεί αποτελεσματικότερα από τη μέθοδο των ελαχίστων μέσων τετραγώνων.

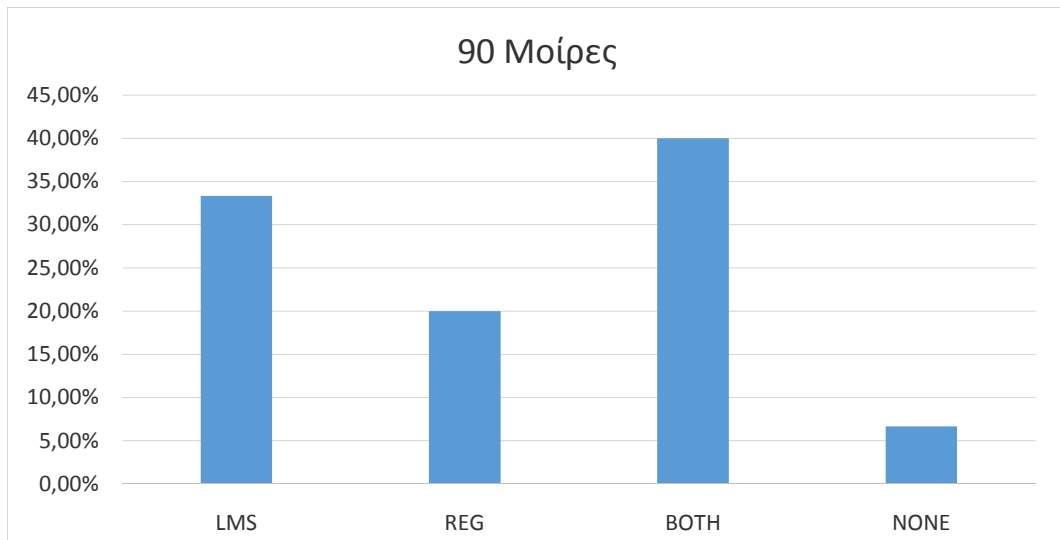
Στην εικόνα 65 απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 75°.



Εικόνα 65: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 75°

Κοιτώντας την εικόνα 65 φαίνεται ότι το 60% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της των ελαχίστων μέσων τετραγώνων, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Το 26.67% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει εξίσου με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, στο ηχείο αναφοράς. Το 13.33% αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων. Στην περίπτωση των 75° το παρατηρείται ότι το μεγαλύτερο ποσοστό των ακροατών αντιλήφθηκε, κατά την ακρόαση, ότι η μέθοδος των ελαχίστων μέσων τετραγώνων για τη δημιουργία ανάστροφου φίλτρου, λειτουργεί αποτελεσματικότερα από τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση.

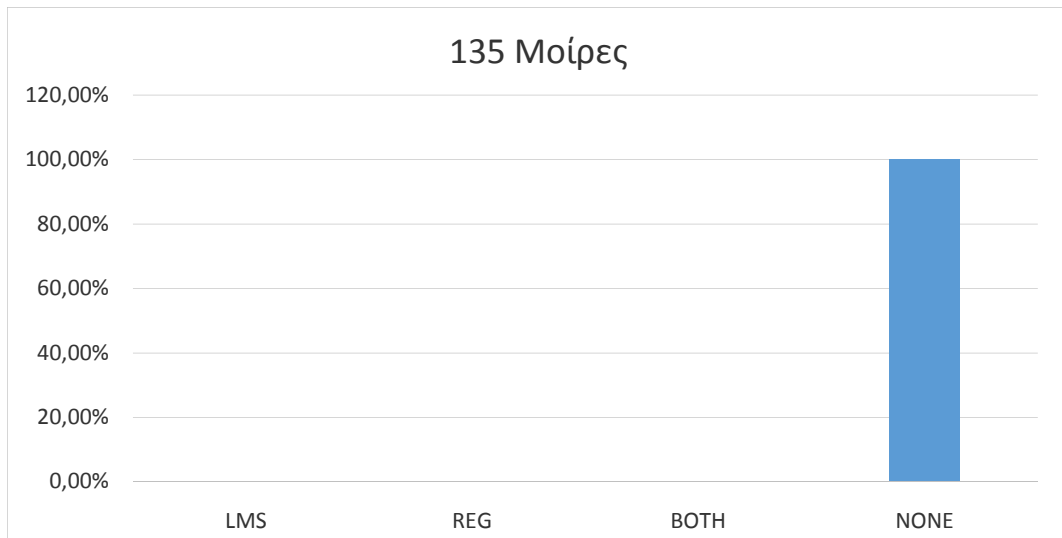
Στην εικόνα 66 απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 90°.



Εικόνα 66: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 90°

Παρατηρώντας την εικόνα 66 συμπεραίνουμε ότι το 40% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή στην ίδια θέση με το ηχείο αναφοράς και για τις δύο μεθόδους δημιουργίας ανάστροφου φίλτρου. Το 33,33%, αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Το 20% των ακροατών αντιλήφθηκε την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει περισσότερο στο ηχείο αναφοράς σε σχέση με την εικονική ηχητική πηγή η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων. Τέλος, το 7% των ακροατών δεν αντιλήφθηκε την εικονική ηχητική πηγή στην επιθυμητή θέση σε σχέση με το ηχείο αναφοράς για καμία από τις δύο μεθόδους. Στην περίπτωση των 90° το μεγαλύτερο ποσοστό των ακροατών επιβεβαίωσε την αποτελεσματική λειτουργία και των δύο μεθόδων.

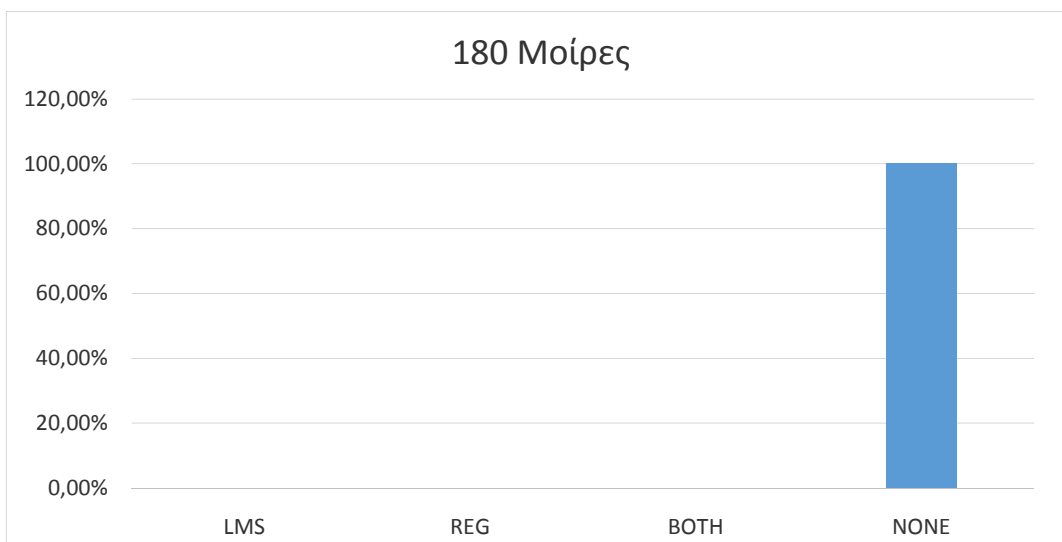
Στην εικόνα 67 απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 135°.



Εικόνα 67: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 135°

Στην περίπτωση των 135° κανείς από τους ακροατές δεν αντιλήφθηκε την εικονική ηχητική πηγή στην επιθυμητή θέση στο χώρο για καμία από της μεθόδους δημιουργίας ανάστροφου φίλτρου.

Στην εικόνα 68 απεικονίζονται τα αποτελέσματα της πειραματικής διαδικασίας για εικονική ηχητική πηγή στις 180°.



Εικόνα 68: Αποτελέσματα πειράματος για την εικονική πηγή στην θέση των 180°

Ομοίως με την περίπτωση των 135°, στην περίπτωση των 180° κανείς από τους ακροατές δεν αντιλήφθηκε την εικονική ηχητική πηγή στην επιθυμητή θέση στο χώρο για καμία από της μεθόδους δημιουργίας ανάστροφου φίλτρου.

5.6 Συμπεράσματα πειραματικής διαδικασίας.

Παρατηρείται στα αποτελέσματα ότι και οι δύο μέθοδοι δημιουργίας ανάστροφου φίλτρου λειτουργούν το ίδιο αποτελεσματικά όσον αφορά την απεικόνιση εικονικών ηχητικών πηγών από τις 0° ως τις 90° . Απαραίτητη προϋπόθεση για τη σωστή ακρόαση, είναι η τήρηση της σωστής θέσης στο λεγόμενο Hot Spot από τον ακροατή. Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα αποτελέσματα των 60° και των 75° .

Στις 60° οι περισσότεροι από τους ακροατές αντιλαμβάνονται την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, να συγκλίνει στην πραγματική ηχητική πηγή περισσότερο από την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων τετραγώνων. Στις 75° συμβαίνει το ακριβώς αντίθετο. Οι περισσότεροι από τους ακροατές, στο ίδιο ποσοστό με την περίπτωση των 60° , αντιλαμβάνονται την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, να συγκλίνει στην πραγματική ηχητική πηγή περισσότερο από την εικονική ηχητική πηγή, η οποία δημιουργήθηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση. Αυτό πιθανότατα ωφελείται στην επίδραση του εκάστοτε φίλτρου στο σήμα, όσον αφορά τη συχνοτική διαμόρφωση. Αυτό σημαίνει ότι συγκεκριμένες συχνοτικές περιοχές, οι οποίες είναι κρίσιμες για την αναγνώριση της θέσης μίας ηχητικής πηγής στο χώρο από τον ακροατή, είναι πιθανόν να αλλοιώθηκαν από την επίδραση των φίλτρων. Άρα το συμπέρασμα που μπορεί να βγει από τα παραπάνω είναι ότι στις 60° το φίλτρο, το οποίο δημιουργήθηκε με τη μέθοδο των ελαχίστων μέσων τετραγώνων, διαμόρφωσε συχνοτικά το σήμα σε κρίσιμη συχνοτική περιοχή για την αναγνώριση της συγκεκριμένης γωνίας από τον ακροατή. Αντίστοιχα ίδιου τύπου επίδραση είναι πιθανόν να επήλθε στο σήμα, στην περίπτωση των 75° , από την επεξεργασία του με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, λαμβάνοντας υπόψιν ότι για τη ρεαλιστική απεικόνιση μίας ηχητικής πηγής στο χώρο απαιτείται ακρίβεια στη συχνοτική απόκριση των HRTFs, της τάξεως του ± 1 dB. Αυτό δικαιολογεί και τα σχόλια των ακροατών στις συγκεκριμένες γωνίες, τα οποία έκαναν λόγο για “θολή” αναπαράσταση της εικονικής ηχητικής πηγής.

Ενδιαφέρον επίσης παρουσιάζουν οι απαντήσεις των ακροατών στις 135° και στις 180°. Ενώ μέχρι τις 90°, τα αποτελέσματα των ακροάσεων είναι ικανοποιητικά, μετά τις 90°, τα αποτελέσματα είναι αρνητικά σε ποσοστό 100% και για τις δύο μεθόδους δημιουργίας ανάστροφου φίλτρου. Στην περίπτωση των 135°, οι ακροατές αντιλήφθηκαν την εικονική ηχητική πηγή στο διάστημα από 80° ως 100°. Στην περίπτωση των 180°, οι ακροατές αντιλήφθηκαν την εικονική ηχητική πηγή στις 0°. Αυτό πιθανότατα οφείλεται στην επίδραση του χώρου κατά την ακρόαση, στην επίδραση του ηλεκτροακουστικού συστήματος και στην χρήση HRTFs μέσου όρου. Οι ιδανικές συνθήκες ακρόασης στην περίπτωση της αναπαραγωγής binaural ηχητικού σήματος από stereo ζεύγος ηχείων είναι:

- Η αναπαραγωγή σε ανηχοϊκό θάλαμο.
- Η αναπαραγωγή σε ηλεκτροακουστικό σύστημα του οποίου η απόκριση έχει ακυρωθεί.
- Η χρήση των HRTFs του κάθε ακροατή ξεχωριστά για τη δημιουργία του binaural αρχείου ήχου.
- Η χρήση των HRTFs του κάθε ακροατή ξεχωριστά για τη δημιουργία του ανάστροφου φίλτρου.

5.7 Επίλογος.

Στο κεφάλαιο 5, παρατέθηκαν οι παράμετροι και οι συνθήκες του πειράματος εντοπισμού εικονικής ηχητικής πηγής, η μέθοδος κατασκευής των binaural αρχείων που χρησιμοποιήθηκαν στο πείραμα και τα χαρακτηριστικά των ακροατών. Επίσης παρουσιάστηκαν και αναλύθηκαν τα αποτελέσματα της πειραματικής διαδικασίας. Τέλος αναφέρθηκαν οι συνθήκες ακρόασης κάτω από τις οποίες ένα σύστημα ηχείων αναπαραγωγής binaural ηχητικού σήματος θα λειτουργούσε ιδανικά.

6. Συμπεράσματα.

Σε αυτό το κεφάλαιο θα παρατεθούν τα συμπεράσματα, τα οποία προέκυψαν από τη σύγκριση μεταξύ των δύο τεχνικών που χρησιμοποιήθηκαν για τη δημιουργία των ανάστροφων φίλτρων. Επιπροσθέτως θα αναφερθούν τα πλεονεκτήματα και τα μειονεκτήματα τους αλλά και τα κοινά χαρακτηριστικά τους.

- Μια σημαντική διαφορά ανάμεσα στις δύο μεθόδους που χρησιμοποιήθηκαν για τη δημιουργία των ανάστροφων φίλτρων, είναι ο χρόνος που χρειάζεται για την ολοκλήρωση του εκάστοτε αλγορίθμου. Η κατασκευή του ανάστροφου φίλτρου με τη μέθοδο των ελαχίστων μέσων τετραγώνων (LMS), πραγματοποιείται στο πεδίο του χρόνου (time domain). Αντιθέτως η κατασκευή του ανάστροφου φίλτρου με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση (fast deconvolution using regularization), πραγματοποιείται στο πεδίο των συχνοτήτων. Όπως έχει αναφερθεί και παραπάνω, ο χρόνος που χρειάζεται για την υλοποίηση ενός φίλτρου στο πεδίο των συχνοτήτων, είναι πολύ μικρότερος από το χρόνο που χρειάζεται για τη δημιουργία ενός φίλτρου στο πεδίο του χρόνου. Πιο συγκεκριμένα, για το φίλτρο των ελαχίστων μέσων τετραγώνων, που χρησιμοποιήθηκε στο πείραμα, ο υπολογιστικός χρόνος που χρειάστηκε για να ολοκληρωθεί ο αλγόριθμος ανέρχεται στα 24 λεπτά. Αντίθετα ο αντίστοιχος χρόνος, που χρειάστηκε για την ολοκλήρωση του αλγορίθμου που κατασκευάστηκε με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, ανέρχεται στα 10 δευτερόλεπτα. Το πλεονέκτημα του LMS είναι ότι δημιουργεί φίλτρα μικρού μήκους τα οποία μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου (real time rendering).
- Ένα σημαντικό πρόβλημα που προκύπτει από το ψηφιακό φιλτράρισμα, είναι η παρουσία συχνότητας κουδουνίσματος (ringing frequency). Παρατηρήθηκε ότι με την κατάλληλη επιλογή κρουστικής απόκρισης HRTF, της γωνίας της θέσης των ηχείων, η συχνότητα κουδουνίσματος ελαττώνεται αισθητά. Αυτός ήταν ένας βασικός περιορισμός στην επιλογή της κρουστικής απόκρισης HRTF, που χρησιμοποιήθηκε για τη δημιουργία ανάστροφου φίλτρου με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση, όπως αναφέρθηκε και στο κεφάλαιο 5.4. Χρησιμοποιήθηκε η κρουστική απόκριση από τη βάση

δεδομένων του IRCAM γιατί το ηχητικό αποτέλεσμα είχε πολύ μικρότερη στάθμη στη συχνότητα κουδουνίσματος σε σχέση με το αντίστοιχο αποτέλεσμα από τη χρησιμοποίηση της κρουστικής απόκρισης HRTF του KEMAR. Για τη δημιουργία ανάστροφου φίλτρου με τη μέθοδο των ελαχίστων μέσων τετραγώνων, χρησιμοποιήθηκε η κρουστική απόκριση του KEMAR λόγω του μικρού της αριθμού δειγμάτων, όπως αναφέρθηκε και στο κεφάλαιο 4.3.2. Στην περίπτωση χρησιμοποίησης της κρουστικής απόκρισης από τη βάση δεδομένων του IRCAM, ο αλγόριθμος χρειάζεται πολύ περισσότερο χρόνο για την ολοκλήρωσή του. Αποτέλεσμα των παραπάνω περιορισμών ήταν η μεγαλύτερη παρουσία της συχνότητας κουδουνίσματος στα αποτελέσματα της μεθόδου των ελαχίστων μέσων τετραγώνων σε σχέση με αυτά της μεθόδου της γρήγορης αποσυνέλιξης με κανονικοποίηση. Παρόλα αυτά το τελικό αποτέλεσμα είναι ικανοποιητικό. Η παρουσία της συχνότητας κουδουνίσματος σε αυτήν την περίπτωση είναι πολύ μικρότερη σε ένταση από την περίπτωση του αποτελέσματος από τη χρήση της κρουστικής απόκρισης του KEMAR με τη μέθοδο της γρήγορης αποσυνέλιξης με κανονικοποίηση.

- Ένα χαρακτηριστικό που παρατηρήθηκε στα ηχητικά αποτελέσματα και των δύο μεθόδων είναι ο «χρωματισμός» του σήματος. Αυτό μπορεί κανείς να το παρατηρήσει ακούγοντας το binaural αρχείο ήχου πριν φιλτραριστεί από τα ανάστροφα φίλτρα και μετά το φιλτράρισμα. Παρατηρώντας τη χροιά του ήχου, δίνεται η αίσθηση, σαν να έχει γίνει κάποιου είδους ισοστάθμιση (equalization) στο σήμα χωρίς να έχουν όμως αλλοιωθεί τα βασικά χαρακτηριστικά της χροιάς.
- Στην παρούσα διπλωματική εργασία δεν υπήρχε η δυνατότητα να πραγματοποιηθεί η πειραματική διαδικασία σε ιδανικές συνθήκες ακρόασης. Η έλλειψη ανηχοϊκού θαλάμου περιορίσε το εύρος, όσον αφορά τη γωνία, λειτουργίας των εικονικών ηχητικών πηγών στις 90°. Από τις 90° μέχρι τις 180° η εικονική ηχητική πηγή ακούγεται σε λάθος θέση. Η έλλειψη μικροφώνων in-ear, τα οποία χρησιμεύουν στην λήψη κρουστικής απόκρισης HRTF για κάθε ακροατή ξεχωριστά, επηρέασε τα αποτελέσματα του πειράματος. Οι μικροδιαφορές στις διαστάσεις των κεφαλιών των ακροατών σε σχέση με τις διαστάσεις των κεφαλιών, που χρησιμοποιήθηκαν για τη

λήψη των HRTFs, επηρέασαν τα αποτελέσματα. Αυτές οι μικροδιαφορές επηρεάζουν, ως προς τη θέση της πηγής, το ηχητικό αποτέλεσμα που ακούει ο κάθε ακροατής.

- Οι ακροατές, που συμμετείχαν στην πειραματική διαδικασία, ανέφεραν ότι αντιλήφθηκαν τη θέση της εικονικής ηχητικής πηγής να ποικίλει ανάλογα με τη συχνотική περιοχή που έστρεφαν την προσοχή τους. Όταν δηλαδή εστίαζαν στην περιοχή από τα περίπου 500 Hz μέχρι τα 8000 Hz εντόπιζαν την εικονική ηχητική πηγή στη σωστή θέση. Στην χαμηλή περιοχή, από τα περίπου 500 Hz και κάτω, ανέφεραν μία «θολή» εικόνα, όσον αφορά τη θέση της πηγής. Από τα περίπου 8000 Hz και πάνω ανέφεραν ότι εντόπιζαν την εικονική ηχητική πηγή στη γωνία περίπου των 30° . Αυτό οφείλεται στη φύση της λειτουργίας του ψυχοακουστικού συστήματος, να μην μπορεί να προσδιορίσει στο χώρο τις χαμηλές συχνότητες, και να μην μπορεί να ξεγελαστεί στις υψηλές συχνότητες λόγω της μεγάλης τους κατευθυντικότητας.
- Μία εξαιρετικά σημαντική παράμετρος στην ακρόαση binaural ηχητικού σήματος από stereo ζεύγος ηχείων τοποθετημένο στις 30° , είναι η τήρηση της θέσης στο hot spot από τον ακροατή. Η απόκλιση της κεφαλής από το hot spot της τάξης των 75-100mm είναι ικανή να καταστρέψει την αίσθηση της παρουσίας της πηγής στην επιθυμητή θέση. Παρατηρήθηκε ότι οι εικονικές ηχητικές πηγές με περιορισμένο συχνотικό φάσμα το οποίο κυμαίνεται στα περίπου 800Hz με 4 KHz δεν απαιτούν την απόλυτη τήρηση του hot spot από τον ακροατή. Παρατηρήθηκε ότι σε αυτή την συχνотική περιοχή το hot spot είναι διευρυμένο
- Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, χρησιμοποιήθηκαν, για τη δημιουργία των φίλτρων και των binaural αρχείων ήχου, κρουστικές αποκρίσεις HRTFs κοινές για όλους τους ακροατές και όχι για τον κάθε ένα ξεχωριστά βάσει των διαστάσεων της κεφαλής και της γεωμετρίας που τον χαρακτηρίζει. Παρ' όλα αυτά, όπως έδειξαν και τα αποτελέσματα της πειραματικής διαδικασίας τα φίλτρα λειτούργησαν αποτελεσματικά στην πλειοψηφία των ακροατών.

Λαμβάνοντας υπόψη όλα τα παραπάνω συμπεράσματα και με βάση τα αποτελέσματα της πειραματικής διαδικασίας μπορεί κανείς να αντιληφθεί ότι η

αναπαραγωγή binaural αρχείων ήχου από stereo ζεύγος ηχείων τοποθετημένα στις 30° μπορεί να επιτευχθεί με επιτυχία σε όλες τις ενδιάμεσες θέσεις από τις 0 έως και τις 90 μοίρες παρ' όλους τους περιορισμούς που την διέπουν. Αυτό επιβεβαιώνει ότι binaural αρχεία ήχου στα οποία έχει εφαρμοστεί ανάστροφο φίλτρο για την ακύρωση του crosstalk και της θέσης των ηχείων, μπορούν εύκολα να χρησιμοποιηθούν σε μουσικές παραγωγές, ηχητικές επενδύσεις οι οποίες θα αναπαραχθούν από συμβατικά ηχεία σε stereo διάταξη σε συμβατικούς χώρους ακρόασης.

Βιβλιογραφία

Ξενόγλωσση

Atal, B. and Schroeder, M. R. (1962). Apparent sound source translator. US patent, 3 236 949.

Bouer, B. (1961). Stereophonic earphones and binaural loudspeakers. J Audio Eng. Soc., 9:148-151.

Christensen F, Jensen CB, Møller H (2000). The design of VALDEMAR – an artificial head for binaural recording purposes. Proc. 109th AES Convention, Los Angeles, Preprint 4404.

Cooper, D. and Bauck, J. (1989). Prospects of transaural recording. J. Audio Eng. Soc., 37(2).

Damaske, P. (1971). Head related stereophony with loudspeaker reproduction. J. Acous- tic. Soc. Am., 50(4).

Kim, S. M. and Wang, S. (2003). A wiener filter approach to the binaural reproduction of a stereo sound. J. Acoustic. Soc. Am., 114(1).

Kirkeby, O., Nelson, P., Hamada, H., and Orduna-Bustamante, F. (1998a). Fast decon- volution of multichannel systems using regularisation. IEEE trans. of speech and audio processing, 6(2).

Kirkeby, O. and Nelson, P. A. (1998). The “stereo dipole”, a virtual source imaging system using two closely related loudspeakers. J. Audio Eng. Soc., 46(5).

Kirkeby, O. and Nelson, P. (1999). Digital filter design for inversion problems in sound reproduction. J. Audio Eng. Soc., 47(7/8).

Koring, J. and Wightman, F. L. (1992). A model of head related transfer function based on principal component analysis and minimum phase reconstruction. J. Acoustic. Soc. Am., 91(4).

Manias, S. (2008). Subjective evaluation of inverse filtering in Virtual Acoustic Imaging Systems – Inversion efficacy under optimal conditions, University of Southampton.

Miyoshi, M. and Kaneda, Y. (1988). Head related stereophony with loudspeaker reproduction. *IEEE trans. on acoust. in speech and signals processing*, 36:145–152.

Mouchtaris, A., Reveliotis, P., and Kyriakakis, C. (2000). Inverse filter design for immersive audio rendering over loudspeakers. *IEEE trans. on multimedia*, 2(2).

Nelson, P. A., Orduna-Bustamante, F., and Hamada, H. (1995). Inverse filter design and equalisation zones in multichannel sound reproduction. *IEEE trans. on speech and audio processing*, 3(3).

Nelson, P. A., Hamada, H., and Elliot, S. (1996). Adaptive inverse filtering for stereophonic sound reproduction. *IEEE trans. of signal processing*, 100(7).

Nelson, P. A., Orduna-Bustamante, F., and Hamada, H. (1995). Inverse filter design and equalisation zones in multichannel sound reproduction. *IEEE trans. on speech and audio processing*, 3(3).

Papadopoulos, T. (2006). Inverse filtering for virtual acoustic imaging systems. PhD thesis, University of Southampton.

Sakamoto, N., Gotoh, T., Kogure, T., and Shimbo, M. (1981). Controlling sound localisation in stereophonic reproduction. *J. Audio Eng. Soc.*, 29(22).

Vorlander M. (2008). *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. Springer. RWTH edition.

Ελληνική

Ξενικάκης Δ. (2006). Σημειώσεις Ηχοληψία Ι. ΤΕΙ Κρήτης, Σχολή Μουσικής Τεχνολογίας και Ακουστικής.

Internet

<http://recherche.ircam.fr/equipes/salles/listen/download.html>

http://en.wikipedia.org/wiki/Least_mean_squares_filter

Πρόσθετη βιβλιογραφία

GRASS (1994). Kemar product data and specifications. GRASS sound and vibration.

Lopez, J. and Gonzalez, A. (2001). Experimental evaluation of cross-talk cancellation regarding loudspeaker's angle. *IEEE trans. of signals processing*, 8(1).

Moore, B. (1989). *An introduction to the psychology of hearing*. Academic Press Limited, third edition.

Ovcharenko, A., Cho, A. J., and Chong, U. (2007). Front-back confusion resolution in three dimensional sound localisation using databases with a dummy head. *J. Acoust. Soc. Am.*, 117(1).

Wenzel, E .M., Arruda, M. A., Kistler, D. J., and Wightman, F. L. (1993). Localisation using non-individualised head related transfer functions. *J. Acoustic. Soc. Am.*, 94:858–867.

Widrow B. and Stearns S. D, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

Yang, J. and Gan, W. S. (2000). Method of improving performance of 3d audio systems. *IEEE proceedings of ICSP*.

Παράρτημα Α

Α.1 Ερωτηματολόγιο πειραματικής διαδικασίας

Πείραμα υποκειμενικού ηχητικού εντοπισμού.

Στην πειραματική διαδικασία που θα ακολουθήσει πρόκειται να πραγματοποιήσετε ακρόαση εικονικών ηχητικών πηγών. Θα σας ζητηθεί να συγκρίνετε τη θέση 2 εικονικών ηχητικών πηγών με μία πραγματική ηχητική πηγή και να ορίσετε ποια από τις δύο εικονικές ηχητικές πηγές είναι πλησιέστερη, όσον αφορά τη θέση στο χώρο, με την πραγματική.

Συμπληρώστε τον παρακάτω πίνακα.

Ερώτηση: Ποια από τις **Εικονικές πηγές** ακούτε πιο κοντά στην **Πραγματική πηγή**;
Συμπληρώστε με **X** το αντίστοιχο πεδίο.

Μοίρες	Εικονική Α (LMS)	Εικονική Β (Reg)	Και τις δύο	Καμία
0				
30				
45				
60				
75				
90				
135				
180				

Σχόλια: