

Ανάπτυξη ηχητικού συστήματος ακυρωτικής παρεμβολής μεταξύ καναλιών σε ανακλαστικό χώρο

Αλμπέρτο Ερσπάμερ^{1,*}, Χρήστος Γιακόπουλος, Ιωάννης Αντωνιάδης

¹Πολυτεχνειούπολη Ζωγράφου

Ηρώων Πολυτεχνείου 9

15772, Ζωγράφου, Αθήνα

*albertoerspamer@mail.ntua.gr

ΠΕΡΙΛΗΨΗ

Τα συστήματα ακυρωτικής παρεμβολής μεταξύ καναλιών χρησιμοποιούνται για τον αποτελεσματικό έλεγχο των αναπαραγόμενων σημάτων σε σημεία ενδιαφέροντος. Συχνά χρησιμοποιούνται για την δυναμική χωρική αναπροσαρμογή των εκλαμβανόμενων ήχων όταν η χρήση ακουστικών από τους χρήστες δεν είναι εφικτή ή επιθυμητή. Στην παρούσα εργασία μελετάται, προσομοιώνεται και αναπτύσσεται πειραματικά ένα στερεοφωνικό σύστημα ακυρωτικής παρεμβολής σε πραγματικές ανακλαστικές συνθήκες. Δίνεται έμφαση στην απόδοση του συστήματος ως προς την συσχέτιση μεταξύ των επιθυμητών και των αναπαραγόμενων σημάτων καθώς και στην διαφορά επιπέδου έντασης μεταξύ των σημείων αναπαραγωγής. Διερευνάται η απόδοση του συστήματος και σε θέσεις αναπαραγωγής γειτνιάζουσες με την αρχική. Επίσης διερευνάται η επίδραση των ανακλάσεων, του κεφαλιού και της απόστασης των ηχείων από τα σημεία αναπαραγωγής.

Development of a crosstalk cancellation system in a reflective space

ABSTRACT

Crosstalk cancellation systems is a widely used technique for various purposes including precise control of audio signals in the points of interest. These systems are often used for binaural reproduction enabling effective control of the perceived location of sounds with speakers instead of wearable headphones. In this study a crosstalk cancelling system is designed, simulated and tested in a normal sound environment with reflections. The performance of the system is determined considering the correlation between the desired and reproduced signals in the positions of interest for various displacements of the reproduction points, as well as the level difference between the reproduction points. The influence of the head, the level of reflections, and the distance of the reproduction speakers from the reproduction points are also considered.

Εισαγωγή

Η τεχνική της ακυρωτικής παρεμβολής (crosstalk cancellation) είναι ιδανική για εφαρμογές όπου απαιτείται ακριβής έλεγχος του ήχου είτε σε συγκεκριμένα σημεία είτε γενικά στον χώρο. Ανάλογα με την εφαρμογή, τα επιθυμητά σημεία αναπαραγωγής μπορούν να είναι δύο (διφωνική αναπαραγωγή) είτε περισσότερα. Η μέθοδος χρησιμοποιείται συχνά και για χωρική αναπαραγωγή πεδίου (Pressure matching). Στην παρούσα εργασία ασχολούμαστε με την περίπτωση αναπαραγωγής σε δύο σημεία ενδιαφέροντος.

Η ακριβής αναπαραγωγή σημάτων σε δύο σημεία ενδιαφέροντος παρουσιάζει ενδιαφέρον γιατί έτσι μπορούν να ελεγχθούν αποτελεσματικά τα ηχητικά χαρακτηριστικά του προσλαμβανόμενου ήχου στα δύο αυτιά. Με αυτό τον τρόπο είναι δυνατή η δημιουργία ηχητικών παραισθήσεων κάνοντας χρήση των πορισμάτων της ψυχοακουστικής [1]. Καθώς η εκλαμβανόμενη θέση ενός ήχου στον χώρο εξαρτάται από διάφορα χαρακτηριστικά του όπως η διαφορά έντασης, η διαφορά φάσης των σημάτων στα δύο αυτιά, ο λόγος ανακλώμενου προς προσπίπτοντος πεδίου κ.α., η μέθοδος προτιμάται όταν σκοπός είναι ο έλεγχος της προσλαμβανουσας διεύθυνσης και απόστασης των διαφόρων «ηχητικών αντικειμένων» στον χώρο.

Η δημιουργία αυτών των ηχητικών ψευδαισθήσεων δίνει την αίσθηση της μετακίνησης του ήχου κάνοντας την ακουστική εμπειρία πιο ρεαλιστική. Συχνά παραδείγματα εφαρμογής των αρχών της ψυχοακουστικής είναι το Dolby Atmos, το amplitude panning [2] κ.α. Τα συστήματα ακυρωτικής παρεμβολής παρουσιάζουν ενδιαφέρον γιατί μπορούν να επιτύχουν πολύ καλύτερα αποτελέσματα όσον αφορά τον ακριβή έλεγχο των ηχητικών ψευδαισθήσεων. Πιθανές πρακτικές εφαρμογές της εν λόγω τεχνολογίας είναι η δημιουργία ενισχυμένων ακουστικών εμπειριών στις τέχνες, η δημιουργία συστημάτων ακουστικής πραγματικότητας, συστήματα αληθοφανούς ακρόασης (περιπτώσεις ακουστικής προσομοίωσης ακριβείας για εκπαίδευση, ιατρικούς σκοπούς κ.α.), ακρόαση δικαναλικής (binaural) μουσικής, έρευνα σχετικά με την ανθρώπινη ηχητική αντιληπτικότητα κ.α.

Ένας βασικός περιορισμός στην δημιουργία ακουστικών ψευδαισθήσεων είναι ο βαθμός ευαισθησίας του ανθρώπινου εγκεφάλου σε μικρές αλλαγές των ηχητικών χαρακτηριστικών των ήχων στα δύο αυτιά. Μεταβολές στη διαφορά της ηχητικής πίεσης στα δύο αυτιά της τάξης μερικών dB, είτε μεταβολές ως προς την χρονική διαφορά των δύο σημάτων της τάξης των 100-1000 us μπορούν να μεταβάλλουν σημαντικά την προσλαμβανόμενα θέση του ήχου. Για τον λόγο αυτό συνήθως προτείνονται δυναμικές μέθοδοι αναπροσαρμογής των εφαρμοζόμενων φίλτρων [3], που ωστόσο εμφανίζουν υπολογιστικούς περιορισμούς ενώ η απόδοση τους επηρεάζεται από τις ανακλάσεις. Συνεπώς η διαπίστωση της απόδοσης ενός τέτοιου συστήματος για μικρές μετακινήσεις από τα σημεία αναπαραγωγής κρίνεται σημαντική.

1. Ανασκόπηση λειτουργίας συστήματος παρεμβολής

Ένα σύστημα ακυρωτικής παρεμβολής περιλαμβάνει δύο βασικά στάδια. Αρχικά πρέπει να γίνει μοντελοποίηση των συναρτήσεων μεταφοράς μεταξύ των πηγών

ήχου (ηχεία) και των σημείων στα οποία ζητούμενο είναι ο έλεγχος του σήματος [4]. Αυτό συνήθως επιτυγχάνεται με τη διαδοχική αναπαραγωγή λευκού θορύβου από τα ηχεία. Κάθε φορά, τα σήματα που καταγράφονται στα μικρόφωνα καταγράφονται. Είτε κατά την διάρκεια των διαταραχών είτε κατόπιν, γίνεται εφαρμογή ενός αλγορίθμου ελαχιστοποίησης σφάλματος για την εύρεση των συναρτήσεων μεταφοράς. Στη συνέχεια, υπολογίζονται τα κατάλληλα φίλτρα που αντισταθμίζουν την επίδραση του φυσικού συστήματος, ώστε να αναπαραχθούν τα επιθυμητά σήματα στα σημεία ενδιαφέροντος.

Στους ανακλαστικούς χώρους, ο υπολογισμός των αντίστροφων συναρτήσεων μεταφοράς προτιμάται να γίνεται στο πεδίο του χρόνου [4], λόγω της παρουσίας στοιχείων ελάχιστης φάσης (minimum phase components) στο φυσικό σύστημα. Τόσο για τις προσομοιώσεις όσο και για το πείραμα που ακολουθούν επιλέχθηκε μέγεθος φίλτρου 5000 taps. Για την μοντελοποίηση των ακουστικών μονοπατιών μεταξύ ηχείων – μικροφώνων επιλέχθηκε step size = 0.00001 ενώ για τον υπολογισμό των φίλτρων αντιστάθμισης step size = 1. Το μεγάλο μέγεθος του φίλτρου ήταν απαραίτητο για την σύγκλιση στο ανακλαστικό περιβάλλον εργαστηρίου ($T60=0.8s$) όπου το σύστημα δοκιμάστηκε.

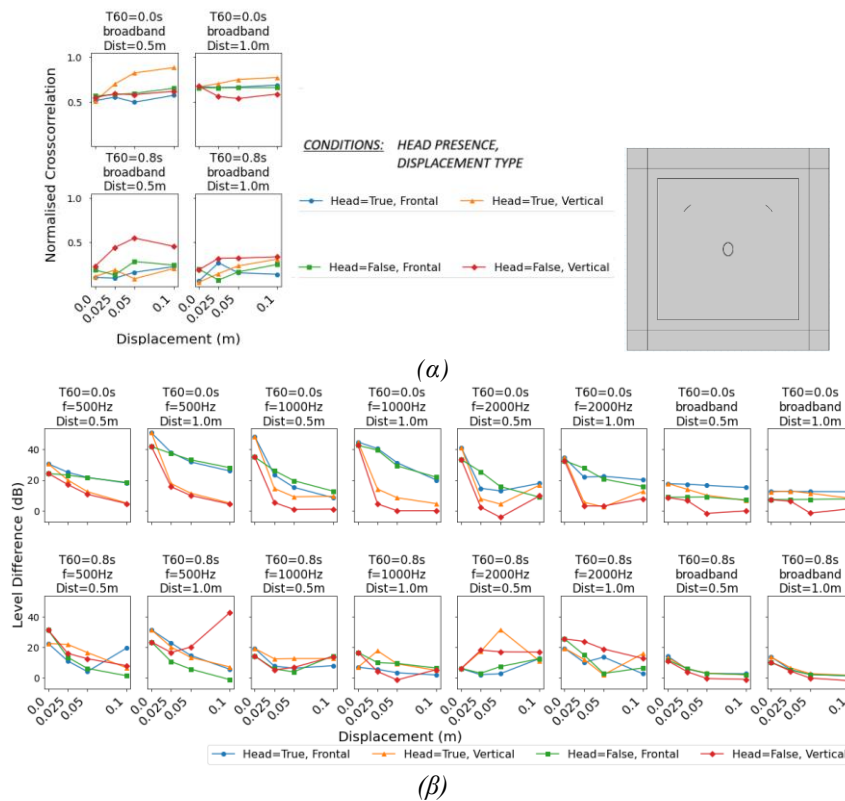
2. Μετρούμενα χαρακτηριστικά (συσχέτιση, διαφορά έντασης)

Η εργασία επικεντρώνεται στην ανάλυση των κανονικοποιημένων συσχετίσεων α) μεταξύ των σημάτων στα δύο σημεία αναπαραγωγής και β) μεταξύ των επιθυμητών και των πραγματικά αναπαραγόμενων σημάτων. Η χρήση κανονικοποιημένων μεγεθών επιτρέπει την ακριβέστερη αξιολόγηση της ποιότητας αναπαραγωγής, μειώνοντας την επίδραση της διαφοράς ενέργειας των σημάτων. Εξετάζεται η απόδοση του συστήματος όσον αφορά και την επιτευχθείσα διαφοροποίηση ως προς το επίπεδο της έντασης στα δύο σημεία ενδιαφέροντος. Τόσο οι συσχετίσεις όσο και οι διαφορές επιπέδου έντασης μετρούνται μέσω προσομοιώσεων και πειραματικών δοκιμών για διάφορες θέσεις μετακίνησης των σημείων αναπαραγωγής. Οι μετακινήσεις είναι είτε πλάγιες είτε εμπρόσθιες και γίνονται ταυτόχρονα και από τα δύο σημεία.

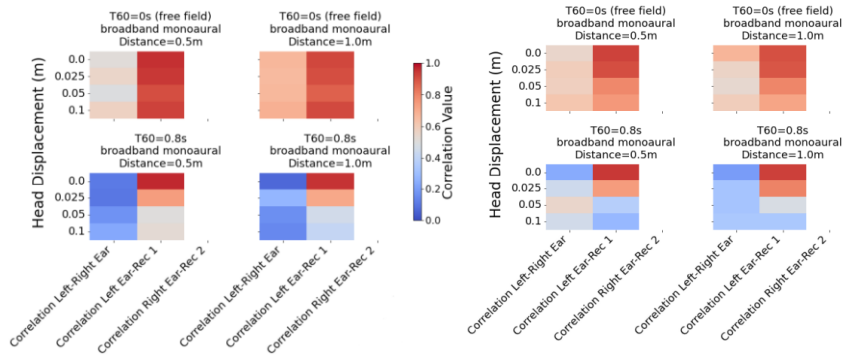
Η απόδοση του συστήματος αναλύεται για τρεις τύπους διαταραχών. Αρχικά, το σύστημα δοκιμάστηκε σε περιπτώσεις όπου η διαταραχή στοχεύει μόνο σε ένα σημείο αναπαραγωγής, ενώ στο άλλο σημείο απαιτείται η ενέργεια να είναι μηδενική. Για αυτήν την περίπτωση, πραγματοποιήθηκαν δοκιμές με τρεις απλές μονοσυχνотικές διαταραχές (500Hz, 1000Hz και 2000Hz) και με μία διαταραχή ευρέως φάσματος (περίπτωση monaural broadband reproduction). Επίσης, εξετάστηκε η πιστότητα αναπαραγωγής δύο στατιστικά ανεξάρτητων σημάτων ίσης ενέργειας (επίσης ευρέως φάσματος) στα δύο σημεία αναπαραγωγής (binaural broadband reproduction). Το εύρος ζώνης των διαταραχών ευρέως φάσματος ορίστηκε στα 200-2500Hz, καθώς ψυχοακουστικά χαρακτηριστικά όπως διαφορά φάσης μεταξύ των προσλαμβανόμενων διαταραχών στα αυτιά είναι σημαντικά σε αυτό το εύρος. Επιπλέον, η επιλογή αυτής της φασματικής περιοχής μειώνει την πολυπλοκότητα των μοντέλων προσομοίωσης, καθιστώντας εφικτή την προσομοίωση χρονικά εξελισσόμενων σημάτων σε εύλογο χρόνο.

3. Προσομοιώσεις

Αναπτύχθηκε προσομοίωση του συστήματος ακυρωτικής παρεμβολής στο περιβάλλον COMSOL. Σχεδιάστηκε ένας δισδιάστατος τετραγωνικό χώρος, όπου μοντελοποιήθηκαν οι τοίχοι, ένα κεφάλι τυπικών διαστάσεων και τα ηχεία ως δύο τόξα διαμέτρου 16 εκατοστών, τοποθετημένα σε απόσταση 0.5 και 1.0 μέτρων από το κέντρο του κεφαλιού. Για τον καθορισμό του χρόνου ανάκλασης T_{60} , παραμετροποιήθηκαν η πυκνότητα και η ταχύτητα διάδοσης του υλικού των τοίχων. Ο σχεδιασμός του κρανίου του κεφαλιού ακολούθησε τις προβλεπόμενες προδιαγραφές σύμφωνα με το [5]. Η μοντελοποίηση των συναρτήσεων μεταφοράς έγινε με διαταραχές χρονικού μήκους 0.9s. Για τον υπολογισμό της μέσης διαφοράς ενέργειας στα σημεία ευρέως φάσματος, υπολογίστηκαν πρώτα τα φάσματα των δύο μικροφώνων και στη συνέχεια η μέση τιμή της διαφοράς τους. Για τα σημεία μιας συχνότητας, η μέση τιμή της διαφοράς υπολογίστηκε σε ένα φασματικό εύρος 10 Hz γύρω από την αντίστοιχη συχνότητα. Τα αποτελέσματα των προσομοιώσεων απεικονίζονται στα σχήματα 3.1-3.4.



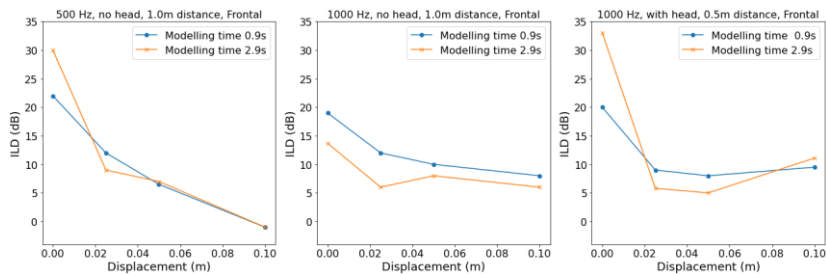
Σχήμα 3.1 Κανονικοποιημένη συσχέτιση (α) και διαφορά επιπέδου έντασης (β) μεταξύ των σημάτων στα δύο σημεία ενδιαφέροντος σε σχέση με την μετατόπιση των σημείων αναπαραγωγής. Η επιθυμητή ενέργεια στο ένα από τα δύο σημεία είναι μηδενική. T_{60} (χρόνος ανάκλασης), και $Dist$ (απόσταση μεταξύ κεφαλιού-ηχείου).



Σχήμα 3.2 Κανονικοποιημένη συσχέτιση μεταξύ των σημάτων (monoaural broadband) στα δύο σημεία ενδιαφέροντος και μεταξύ των επιθυμητών και αναπαραγόμενων σημάτων σε σχέση με τη μετατόπιση των σημείων αναπαραγωγής. Η επιθυμητή ενέργεια στο ένα από τα δύο σημεία είναι μηδενική. Αριστερά: Με κεφάλι - Εμπρός/πίσω μετακινήσεις. Δεξιά: Χωρίς κεφάλι - Πλάγιες μετακινήσεις.

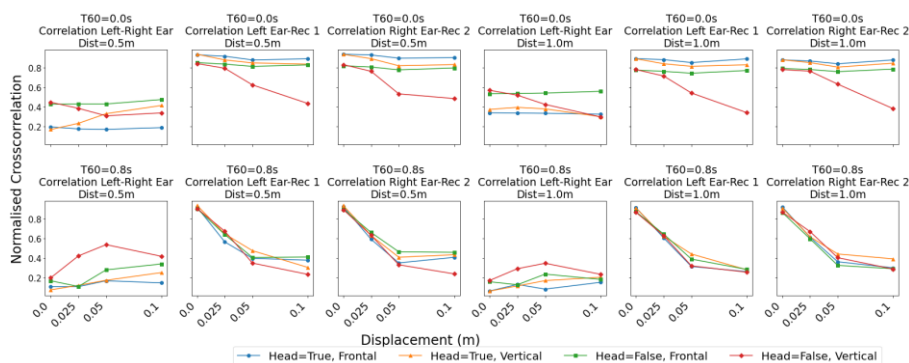
Τα σχήματα 3.1 και 3.2 απεικονίζουν την απόδοση του συστήματος όταν το ένα από τα δύο επιθυμητά σήματα έχει μηδενική ενέργεια. Είναι θεμιτό η κανονικοποιημένη συσχέτιση των σημάτων στα δύο σημεία αναπαραγωγής να είναι ελάχιστη. Για τα σήματα ευρέως φάσματος παρατηρείται ότι σε συνθήκες μηδενικής ανακλαστικότητας, η κανονικοποιημένη συσχέτιση μεταξύ των δύο σημείων είναι μεγαλύτερη συγκριτικά με ανακλαστικά περιβάλλοντα, όπου και εμφανίζεται καλύτερος διαχωρισμός (σχήμα 3.1-α). Ωστόσο, η μεγαλύτερη ανακλαστικότητα οδηγεί σε μειωμένη απόδοση - μείωση στην κανονικοποιημένη συσχέτιση μεταξύ επιθυμητού - αναπαραγόμενου σήματος για μετακινήσεις των σημείων ενδιαφέροντος με ή χωρίς κεφάλι (σχήμα 3.2) και για τους δύο τύπους μετατόπισης. Ταυτόχρονα, παρατηρείται ότι η ύπαρξη κεφαλιού οδηγεί σε αυξημένη στατιστική ανεξαρτησία μεταξύ των αναπαραγόμενων σημάτων.

Οι προσομοιώσεις επίσης δείχνουν την μείωση της απόδοσης ως προς την επιθυμητή διαφορά ενέργειας στα δύο σημεία σε ένα ανακλαστικό περιβάλλον τόσο στα σημεία μοντελοποίησης όσο και στις άλλες θέσεις (σχήμα 3.1-β). Τέλος, σε ανακλαστικά περιβάλλοντα παρατηρήθηκε η εξάρτηση του συστήματος από τον χρόνο μοντελοποίησης των συναρτήσεων μεταφοράς. Για μεγαλύτερο χρόνο μοντελοποίησης (2.9s έναντι 0.9s) η απόδοση στα σημεία μοντελοποίησης είναι μεγαλύτερη ωστόσο σε αυτή την περίπτωση πολύ μικρές μετατοπίσεις οδηγούν σε χειρότερα αποτελέσματα (σχήμα 3.3).



Σχήμα 3.3 Επίδραση του χρόνου μοντελοποίησης των συναρτήσεων μεταφοράς στην απόδοση του συστήματος (για $T60=0.8s$). Η απόδοση βελτιώνεται στα μοντελοποιημένα σημεία αναπαραγωγής αλλά μειώνεται έντονα για μικρές μετακινήσεις.

Τέλος, αξιολογήθηκε η απόδοση του συστήματος για την περίπτωση αναπαραγωγής δύο στατιστικά ανεξάρτητων σημάτων ίσης ενέργειας στα δύο σημεία ενδιαφέροντος. Παρατηρείται ότι, σε μη ανακλαστικό περιβάλλον, η κανονικοποιημένη συσχέτιση μεταξύ των δύο σημείων αναπαραγωγής είναι υψηλότερη, όπως και η συσχέτιση μεταξύ του επιθυμητού και του αναπαραγόμενου σήματος. Αντίθετα, σε ανακλαστικό περιβάλλον, η κανονικοποιημένη συσχέτιση μεταξύ των δύο σημείων αναπαραγωγής είναι μειωμένη, αλλά η συσχέτιση μεταξύ επιθυμητού και αναπαραγόμενου σήματος εξαρτάται πολύ έντονα από τον βαθμό μετακίνησης των σημείων αναπαραγωγής (σχήμα 3.4).



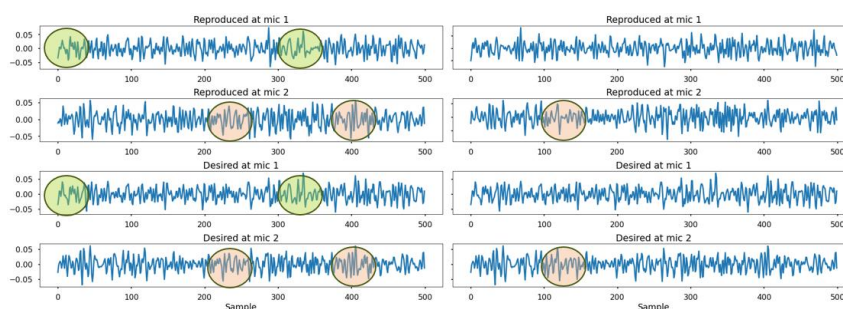
Σχήμα 3.4 Κανονικοποιημένες συσχετίσεις α) μεταξύ των σημάτων στα δύο σημεία ενδιαφέροντος και β) μεταξύ επιθυμητών-αναπαραγόμενων σημάτων σε σχέση με την μετατόπιση των σημείων αναπαραγωγής, για την περίπτωση αναπαραγωγής δύο σημάτων ευρέος φάσματος ίσως ενέργειας.

4. Πειραματική διάταξη

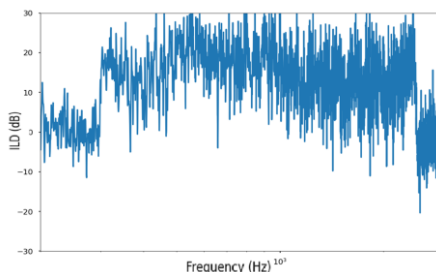
Το σύστημα ακρωτικής παρεμβολής δοκιμάστηκε και πειραματικά σε εργαστηριακές συνθήκες όπου ο μετρούμενος χρόνος ανάκλασης ήταν ίσος με τον αντίστοιχο χρόνο που επιλέχθηκε για τις προσομοιώσεις ($T60=0.8s$). Το σχήμα 4.1 δείχνει την πειραματική διάταξη, ενώ στο σχήμα 4.2 φαίνεται η πολύ καλή απόδοση του συστήματος στα σημεία ενδιαφέροντος για αναπαραγωγή στατιστικά ανεξάρτητων σημάτων ίσης ενέργειας και η μείωση της ομοιότητας μεταξύ επιθυμητών - αναπαραγόμενων σημάτων για πλάγια μετακίνηση 2.5 εκατοστών. Στο σχήμα 4.3 φαίνεται η απόδοση του συστήματος για την περίπτωση αναπαραγωγής σήματος ευρέος φάσματος όπου το ένα εκ των δύο σημείων έχει μηδενική ενέργεια.



Σχήμα 4.1 Πειραματική διάταξη συστήματος ακυρωτικής παρεμβολής και μετρητική μέθοδος για την μέτρηση της απόδοσης για διάφορες μετακινήσεις.



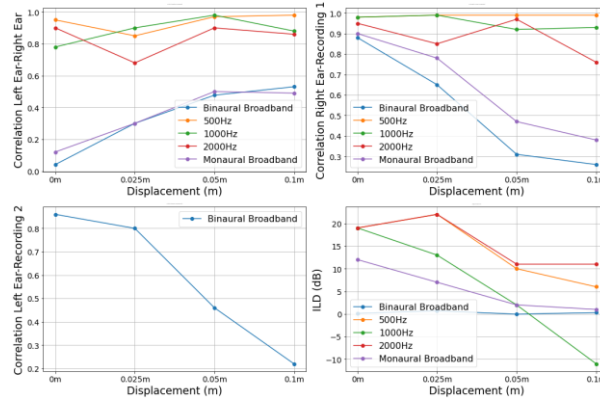
Σχήμα 4.2 Επιθυμητά και πραγματικά σήματα στα δύο σημεία ενδιαφέροντος για μηδενική μετακίνηση (αριστερά) και για μετακίνηση 2.5 cm (δεξιά).



Σχήμα 4.3 Μετρούμενη διαφορά έντασης στα δύο σημεία ενδιαφέροντος για την περίπτωση μονής διαταραχής ευρέος φάσματος με μηδενική μετακίνηση των σημείων ενδιαφέροντος για την πειραματική διάταξη.

Στο σχήμα 4.4 φαίνονται τα αποτελέσματα των μετρήσεων της πειραματικής διάταξης για όλες τις περιπτώσεις αναπαραγωγής. Το πειραματικό σύστημα δοκιμάστηκε για πλάγιες μετατοπίσεις χωρίς κεφάλι.

Σημειώνεται ότι στο πραγματικό πείραμα ο χρόνος των διαταραχών ήταν σημαντικά μεγαλύτερος (10s). Ο θόρυβος προερχόμενος από τα μικρόφωνα και τον ηλεκτρονικό εξοπλισμό οδηγεί στην δυσκολία επίτευξης τέλει σύγκλισης των αλγορίθμων μοντελοποίησης. Αυτό εξηγεί τη μέγιστη διαφοροποίηση των 20dB, σε σύγκριση με τα 30dB που παρατηρήθηκαν στις προσομοιώσεις για τις ίδιες συνθήκες με χρόνο μοντελοποίησης 2.9s (σχήμα 3.3).



Σχήμα 4.4 Μετρούμενα μεγέθη συσχέτισης και διαφοράς έντασης για διάφορες μετατοπίσεις των σημείων ενδιαφέροντος για τις διάφορες περιπτώσεις αναπαραγωγής.

5. Συμπεράσματα

Το προτεινόμενο σύστημα ακρωτικής παρεμβολής κρίνεται ιδιαίτερα αποτελεσματικό στα σημεία μοντελοποίησης. Στο ανακλαστικό περιβάλλον που εξετάστηκε η συσχέτιση μεταξύ των επιθυμητών και των αναπαραγόμενων σημάτων ήταν μικρή (<0.7 και >0.4) για πλάγιες μετακινήσεις μεταξύ 2.5-5 εκατοστών αλλά ικανοποιητική (>0.7) για μετακινήσεις έως 2.5 εκατοστών. Για εμπρόσθιες μετακινήσεις, η συσχέτιση είναι σημαντικά πιο ικανοποιητική. Σε ανηχοικούς χώρους με ύπαρξη κεφαλιού, η συσχέτιση διατηρείται σε υψηλά επίπεδα για όλες τις υπό εξέταση μετακινήσεις.

6. Αναφορές

- [1] P. Zahorik, D. Brungart, D.S. Brungart, and A.W. Bronkhorst, "Auditory distance perception in humans: A summary of past and present research," *Acta Acustica united with Acustica*, 91, pp. 409-420 (2005).
- [2] V. Pulkki, M. Karjalainen, and J. Huopaniemi, "Analyzing Virtual Sound Source Attributes Using a Binaural Auditory Model" *Journal of the Audio Engineering Society*, 47 (4), (1999).
- [3] S. Galvez, M. Menzies, and F. Fazi, "Dynamic audio reproduction with linear loudspeaker arrays," *Journal of the Audio Engineering Society*, 67 (4), pp. 190-200 (2019).
- [4] P.A. Nelson, H. Hamada, and S.J. Elliott, "Adaptive inverse filters for stereophonic sound reproduction," *IEEE Transactions on Signal Processing*, 40, pp. 1621-1632 (1992).
- [5] P.A. Hasgall, F. Di Gennaro, C. Baumgartner, E. Neufeld, B. Lloyd, M.C. Gosselin, *et al.*, "IT'IS Database for thermal and electromagnetic parameters of biological tissues V4.1," IT'IS Foundation (2022).