



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Μάθημα: Ψηφιακή Επεξεργασία Σημάτων

3^η Εργαστηριακή Άσκηση

**Θέμα: Συστοιχίες Μικροφώνων (Microphone Arrays) και
Πολυκαναλική Επεξεργασία Σημάτων
(Multichannel Signal Processing)**



Ονοματεπώνυμο: **Αθανασίου Νικόλαος** **ΑΜ:03112074**

Σταυρακάκης Δημήτριος **ΑΜ:03112017**

Ημερομηνία Παράδοσης: 19/6/2015

ΜΕΡΟΣ 1^ο: ΣΥΣΤΟΙΧΙΕΣ ΜΙΚΡΟΦΩΝΩΝ ΚΑΙ ΧΩΡΙΚΟ ΦΙΛΤΡΑΡΙΣΜΑ (SPATIAL FILTERING)

Παράθεση Θεωρίας:

Στο συγκεκριμένο μέρος της άσκησης μελετάται η επίδραση διαφόρων παραγόντων στην διαδικασία φιλτραρίσματος και συνδυασμού σημάτων μέσω συστοιχίας μικροφώνων από πολλαπλά κανάλια για την παραγωγή μιας εξόδου (beamforming). Ειδικότερα θα μελετηθούν οι εξής παράγοντες:

- το πλήθος των μικροφώνων που διαθέτει η συστοιχία
- η απόσταση των μικροφώνων
- η γωνία με την οποία φτάνει το επιθυμητό σήμα στη συστοιχία

Σε ένα σύστημα beamforming πρέπει να χρησιμοποιηθεί το χωρικό ανάλογο της απόκρισης συχνότητας ενός χρονικού φίλτρου, για να μπορεί να υπολογιστεί η έξοδος του για οποιαδήποτε είσοδο. Το χωρικό ανάλογο του είναι το beam pattern, που περιγράφεται με τη σχέση:

$$B(\omega, \vec{\alpha}) = \vec{H}^T(\omega) \cdot \vec{d}(\vec{k}) \Big|_{k=\frac{\omega \vec{\alpha}}{c}}$$

Στην παραπάνω σχέση ο H^T είναι ο ανάστροφος πίνακας που περιέχει την απόκριση συχνότητας κάθε φίλτρου σε κάθε γραμμή της συστοιχίας, ενώ στο διάνυσμα d περιέχονται όλα τα απαραίτητα στοιχεία για τα χωρικά χαρακτηριστικά της συστοιχίας. Καλούμαστε να μελετήσουμε τον delay-and-sum beamformer, στον οποίο κάθε σήμα εισόδου στα μικρόφωνα της συστοιχίας υφίσταται κάποια χρονική καθυστέρηση για τον απαραίτητο χρονικό ευθυγραμμισμό. Ακολούθως όλα τα χρονικά μετατοπισμένα σήματα αθροίζονται και διαιρούνται με το πλήθος των μικροφώνων για να παραχθεί η τελική έξοδος του beamformer. Εφόσον τα μικρόφωνα είναι ισαπέχοντα η διάταξη απλουστεύεται και λέμε τότε ότι πρόκειται για μία ομοιομορφική γραμμική συστοιχία (ΟΓΣ) της οποίας το beam pattern είναι το ακόλουθο:

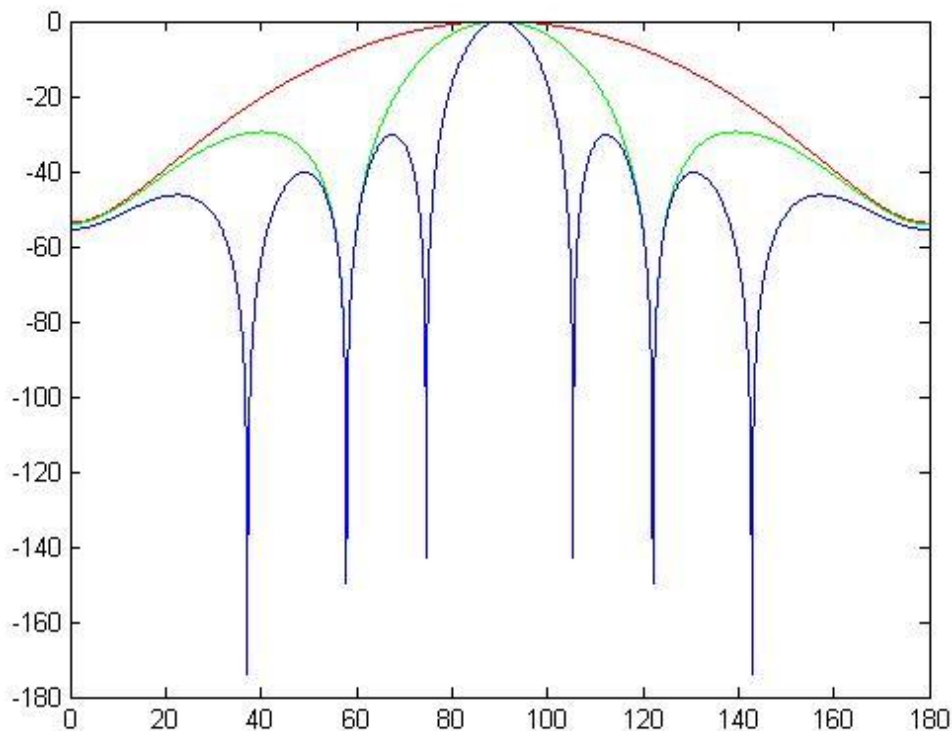
$$B(\omega, \theta) = \frac{1}{N} \cdot \frac{\sin \left[\frac{N}{2} \cdot \frac{\omega}{c} \cdot d (\cos \theta - \cos \theta_s) \right]}{\sin \left[\frac{1}{2} \cdot \frac{\omega}{c} \cdot d (\cos \theta - \cos \theta_s) \right]}$$

Όπου :

- N : πλήθος των μικροφώνων της συστοιχίας
- ω : γωνιακή συχνότητα του σήματος
- θ : πολική γωνία
- θ_s : γωνία με την οποία φτάνει το επιθυμητό σήμα στην συστοιχία.

Μελέτη των χαρακτηριστικών του delay-and-sum beam pattern για ΟΓΣ

Για συχνότητα σήματος $f = 2\text{kHz}$ και $\theta_s = 90^\circ$ σχεδιάζεται το μέτρο του delay-and-sum beam pattern ως συνάρτηση της γωνίας θ . Αρχικά για σταθερή απόσταση μικροφώνων $d = 4\text{cm}$ σχεδιάζουμε στο ίδιο διάγραμμα το μέτρο που αναφέρεται παραπάνω για πλήθος μικροφώνων 4 , 8 και 16. Παρακάτω παρατίθενται τα προαναφερθέντα γραφήματα:

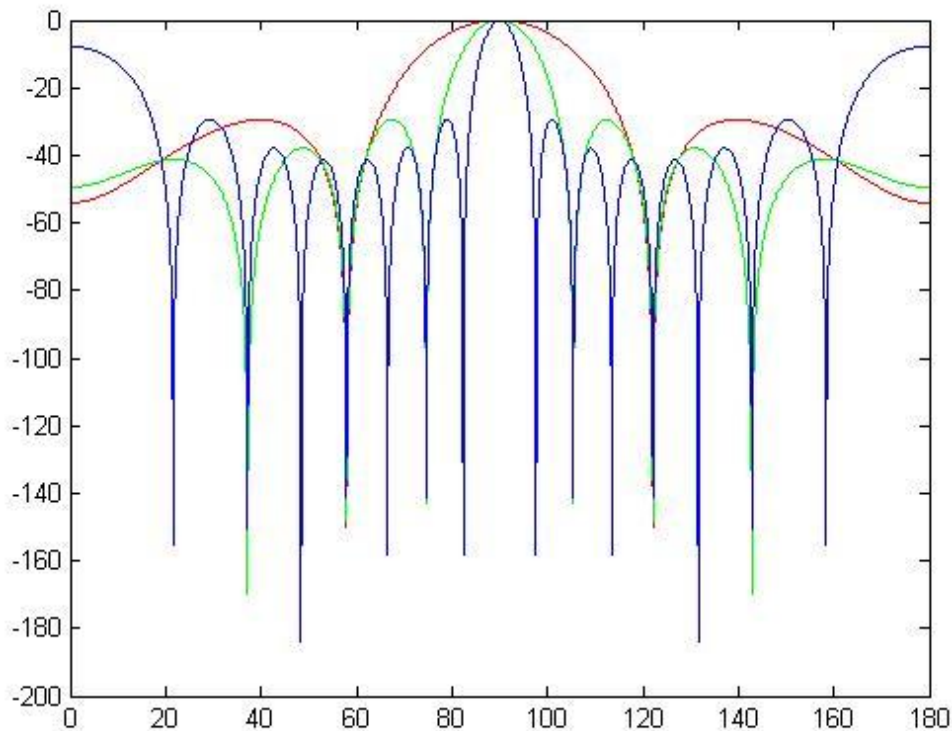


Γραφική παράσταση του μέτρου του delay-and-sum beam pattern ως συνάρτηση της πολικής γωνίας θ για **πλήθος μικροφώνων 4 (κόκκινη γραμμή)** , **8 (πράσινη γραμμή)** και **16 (μπλε γραμμή)**

Παρατηρούμε ότι η αύξηση του αριθμού των μικροφώνων αυξάνει τους λοβούς και πιο συγκεκριμένα οι μεταβολές τους είναι πιο απότομες στο γράφημα του μέτρου του delay and sum beam pattern ως συνάρτηση της γωνίας θ , ενώ στα 0dB βρίσκεται μόνο για $\theta = 90^\circ$, που είναι και το επιθυμητό. Αντίθετα σε μικρότερο πλήθος μικροφώνων, οι λοβοί που σχηματίζονται είναι ομαλότεροι και έτσι παίρνουν τιμές για μεγαλύτερο πλήθος γωνιών.

Στη συνέχεια θα μελετήσουμε την επίδραση του παράγοντα της απόστασης των μικροφώνων μεταβάλλοντας την για τις επιθυμητές τιμές που ζητώντας και σχεδιάζοντας το μέτρο του delay-and-sum beam pattern με διαφορετικά χρώματα όπως και πριν στη μελέτη για το πλήθος των μικροφώνων.

Ακολουθώς παρατίθεται ανάλογη γραφική παράσταση για σταθερό πλήθος μικροφώνων ($N=8$) και μεταβλητές τιμές αποστάσεων μικροφώνων.



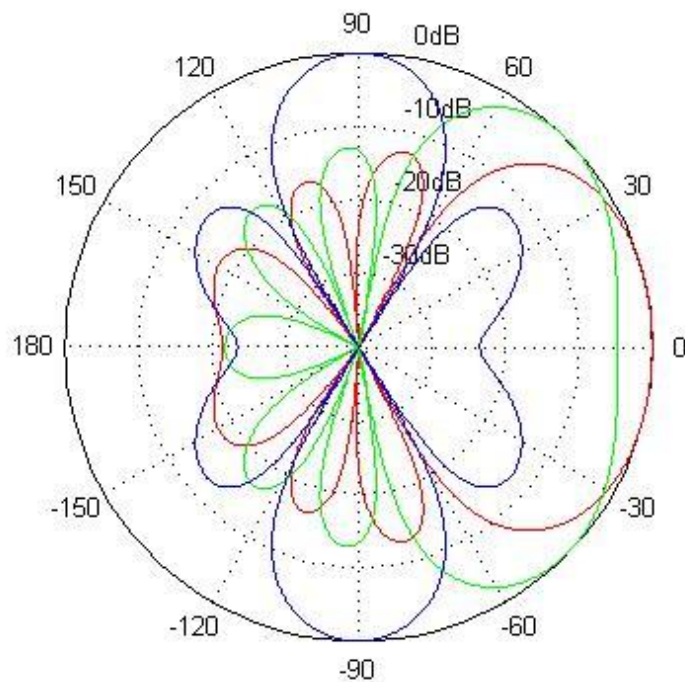
Γραφική παράσταση του μέτρου του *delay-and-sum beam pattern* ως συνάρτηση της πολικής γωνίας θ για απόσταση μικροφώνων **4cm(κόκκινη γραμμή)**, **8cm(πράσινη γραμμή)** και **16cm(μπλε γραμμή)**

Παρατηρούμε ότι όσο μεγαλώνει η απόσταση των μικροφώνων τόσο χειρότερη είναι η απόδοση του beam pattern καθώς εμφανίζει υψηλότερες κορυφές που πλησιάζουν την τιμή 0dB, και όχι μόνο μία για $\theta = 90^\circ$, όπως φαίνεται καθαρά στη μπλε γραμμή.

Αντίθετα, για απόσταση 4 cm που είναι η μικρότερη από αυτές που σχεδιάστηκαν, παρατηρούμε μόνο ο ζητούμενος λοβός να φτάνει στα 0dB, με τους υπόλοιπους, που είναι εμφανώς ομαλότεροι, να έχουν τις κορυφές τους σε αρκετά μικρότερες τιμές.

Στη συνέχεια θεωρούμε μια Ομοιομορφική Γραμμική Συστοιχία όπου τα σήματα έχουν συχνότητα $f = 2 \text{ kHz}$, ενώ η συστοιχία αποτελείται από 8 μικρόφωνα με απόσταση 4 cm.

Για γωνίες άφιξης $\theta_s = 0^\circ, 45^\circ$ και 90° σχεδιάζεται σε κοινό διάγραμμα το μέτρο του *delay-and-sum beam pattern* ως συνάρτηση της γωνίας θ σε πολικό διάγραμμα, όπως φαίνεται στο ακόλουθο σχήμα.



Πολικό διάγραμμα του μέτρου του *delay-and-sum beam pattern* ως συνάρτηση της πολικής γωνίας θ για γωνίες άφιξης 0° (κόκκινη γραμμή), 45° (πράσινη γραμμή) και 90° (μπλε γραμμή)

Παρατήρηση-Συμπέρασμα

- ✚ Ο λοβός με κορυφή στα 0dB υπάρχει και στην αντίθετες συχνότητες από τις προαναφερθείσες, γεγονός που δίνει να καταλάβουμε ότι ο Delay and Sum Beamformer δέχεται σήματα υπό μία συγκεκριμένη γωνία **ανεξαρτήτως θετικής ή αρνητικής διεύθυνσης**.
- ✚ Επίσης παρατηρούμε ότι η πιο απόδοτική περίπτωση του συστήματος είναι **όταν το σήμα φτάνει κάθετα ($\theta=90^\circ$)** στη συστοιχία, αφού είναι ξεκάθαρο ότι στις άλλες περιπτώσεις ο επιθυμητός λοβός έχει μεγαλύτερο εύρος, δηλαδή δεν λειτουργεί μόνο για την επιθυμητή γωνία θ αλλά για περισσότερες, πράγμα που δεν είναι επιθυμητό. Γενικά ο στόχος του Delay and Sum Beamformer είναι να επιτυγχάνει μια πιο “στενή” απόκριση στη διεύθυνση της επιθυμητής γωνίας.

ΜΕΡΟΣ 2^ο: ΕΦΑΡΜΟΓΗ BEAMFORMING ΓΙΑ SPEECH ENHANCEMENT

2.1: Beamforming σε προσομοιωμένα σήματα

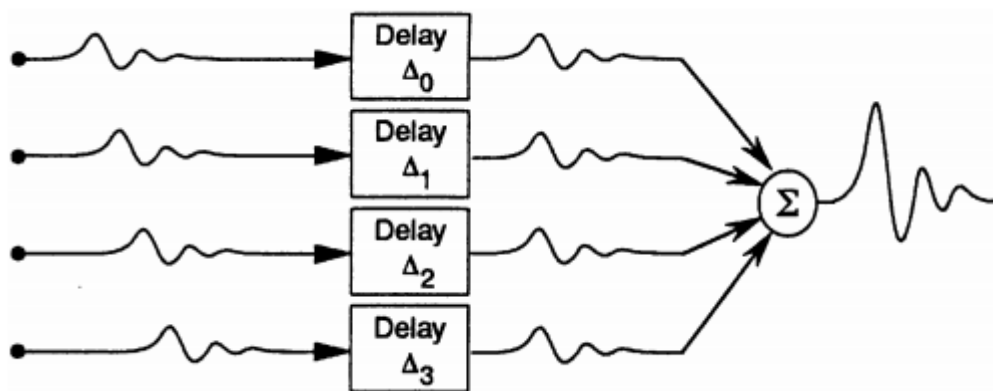
Θεωρούμε ότι έχουμε μια γραμμική συστοιχία **N=7** μικροφώνων με απόσταση $d=4\text{cm}$ τα οποία καταγράφουν σήματα που παράγονται από δύο σημειακές πηγές.

Η μία σημειακή πηγή παράγει σήμα φωνής και βρίσκεται σε γωνία $\theta = 45^\circ$ σε σχέση με τη συστοιχία.

Η δεύτερη παράγει ένα σήμα θορύβου σε γωνία $\theta = 135^\circ$ σε σχέση με τη συστοιχία.

A.Delay-and-sum beamforming

Στο συγκεκριμένο μέρος της άσκησης υλοποιούμε την αποθορυβοποίηση με delay-and-sum beamforming. Η διαδικασία απεικονίζεται παρακάτω:



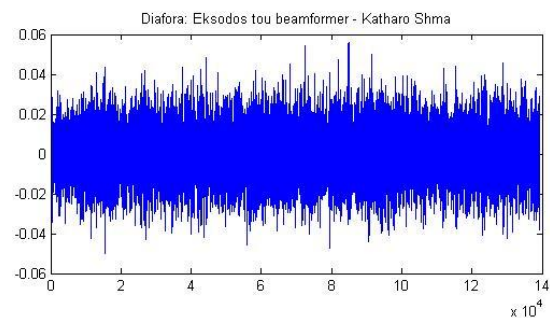
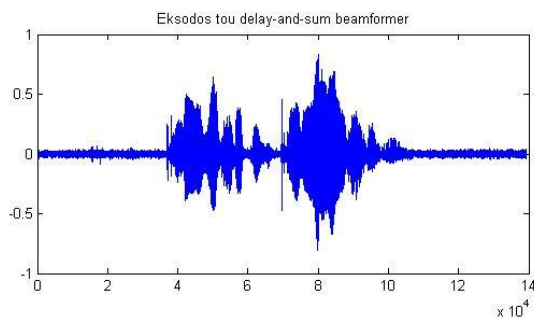
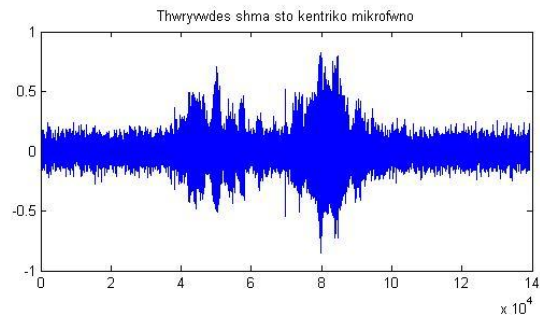
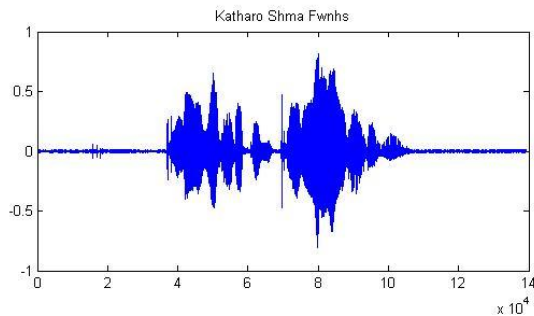
Βηματική Ανάλυση

✚ Αρχικά υλοποιούμε την συνάρτηση *Shift* η οποία μετατοπίζει το διακριτό σήμα που λαμβάνει κατά M (παράμετρος της Shift) δείγματα. Στη συνέχεια καλείται η συνάρτηση αυτή για την χρονική ευθυγράμμιση των 7 σημάτων (που λαμβάνονται από τα μικρόφωνα) στο χρόνο. Στη συνάρτηση δίνονται ως παράμετροι τα 7 αυτά σήματα και οι σταθερές μετατόπισεις υπολογισμένες από τον παρακάτω τύπο:

$$\tau_n = -\frac{(n - \frac{N-1}{2}) * d * \cos \theta}{c} \quad \text{με } n = 0, 1, \dots, 6$$

Αφού αθροίσουμε τα σήματα διαιρούμε με το πλήθος τους δηλαδή $N (=7)$ και έτσι προκύπτει το αποτέλεσμα του beamformer την αποθηκεύουμε ως αρχείο wav με όνομα `sim_ds`.

■ Παρακάτω απεικονίζονται οι κυματομορφές για τα παρακάτω σήματα:

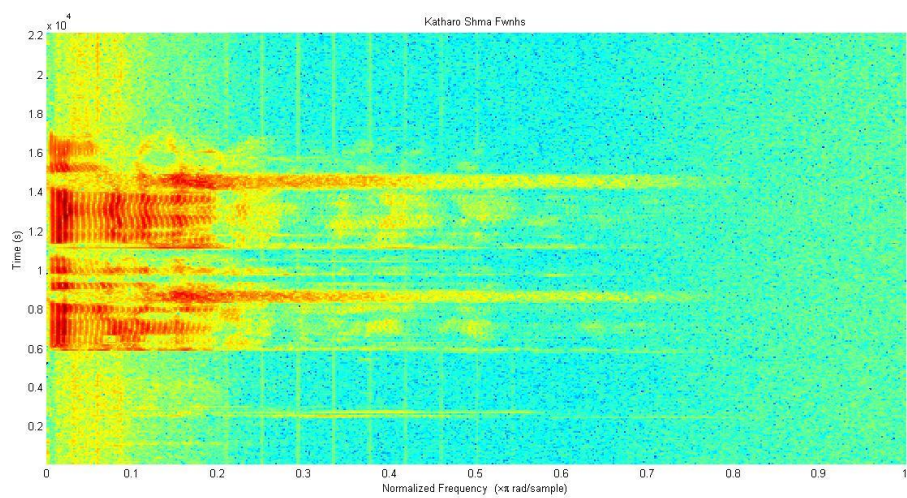


Κυματομορφές για το καθαρό σήμα φωνής, το θορυβώδες σήμα στο κεντρικό μικρόφωνο, της έξοδο του beamformer και το θόρυβο

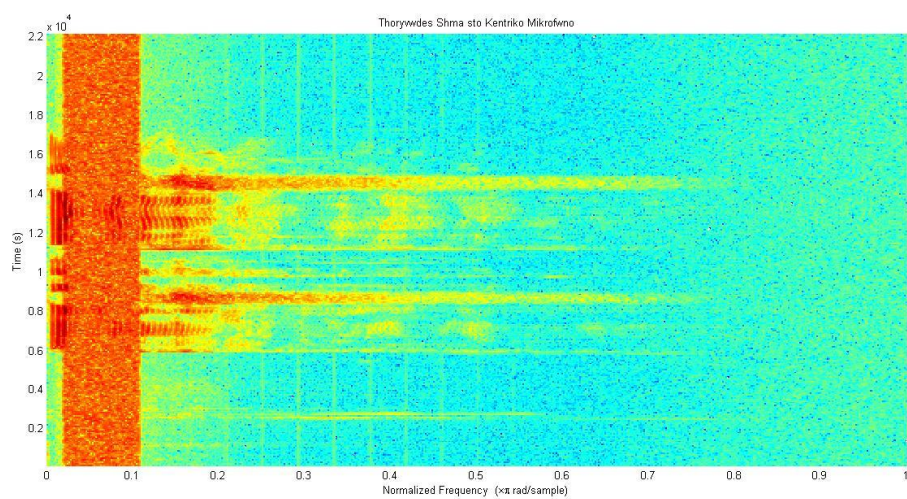
Παρατηρείται ότι η έξοδος του delay and sum beamformer σχεδόν συμπίπτει με το καθαρό σήμα φωνής ενώ στο θορυβώδες σήμα στο κεντρικό μικρόφωνο γίνεται έντονα αισθητή η παρουσία θορύβου γεγονός που το διαφοροποιεί από τα δύο παραπάνω σήματα. Επιπρόσθετα έχει σχεδιαστεί το σήμα της εξόδου του beamformer αν αφαιρεθεί το καθαρό σήμα και δίνεται κατανοητό ότι αυτό αποτελείται μόνο από θόρυβο γεγονός που επιβεβαιώνει τη σωστή λειτουργία του φίλτρου .

Παρακάτω παρατίθενται στα σπεκτρογραφήματα των σημάτων όπως δόθηκαν από το Matlab.

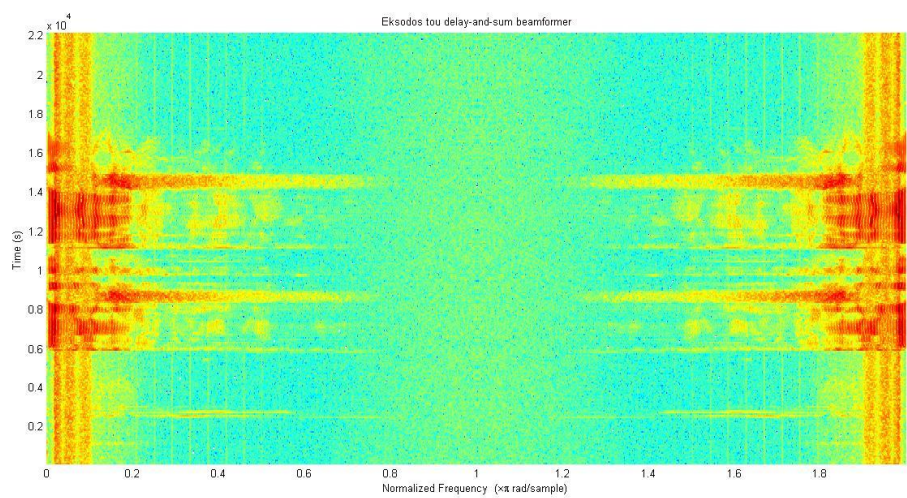
Spectrograms



Σπεκτρογράφημα καθαρού σήματος φωνής



Σπεκτρογράφημα θορυβώδους σήματος στο ΚΕΝΤΡΙΚΟ μικρόφωνο



Σπεκτρογράφημα εξόδου *beamformer*

✚ Αν $s(t) = x(t) + u(t)$, με

- $s(t)$: το θορυβώδες σήμα φωνής
- $x(t)$: το καθαρό σήμα φωνής
- $u(t)$ ο θόρυβος,

SNR του θορυβώδους σήματος στο κεντρικό μικρόφωνο :

$$SNR_{central} = 10 \log_{10} \frac{\sum_{n=1}^{\infty} x^2(t)}{\sum_{n=1}^{\infty} (x(t) - s(t))^2}$$

Αντίστοιχα, το SNR της εξόδου του delay-and-sum beamformer υπολογίζεται από τον τύπο:

$$SNR_{beamformer} = 10 \log_{10} \frac{\sum_{n=1}^{\infty} x^2(t)}{\sum_{n=1}^{\infty} (x(t) - y(t))^2}$$

όπου $y(t)$ η έξοδος του beamformer.

Με τη βοήθεια του Matlab υπολογίζουμε:

- $SNR_{central} = 3$
- $SNR_{beamformer} = 17.5604$

Παρατηρούμε ότι το SNR της εξόδου του beamformer αρκετά μεγαλύτερο από το SNR στο κεντρικό μικρόφωνο, δηλαδή υπάρχει λιγότερος θόρυβος στην έξοδο του φίλτρου, πράγμα που εξακριβώνει την βέλτιστη λειτουργία delay-and-sum beamformer.

B. Μονοκαναλικό Wiener φιλτράρισμα

Για να έχουμε ένα μέτρο σύγκρισης της παραπάνω μεθόδου θα εφαρμοστεί και η διαδικασία αποθρομβοποίησης Wiener filtering. Για την εκτέλεση αυτής της προσομοίωσης λαμβάνουμε ένα τμήμα του σήματος στο κεντρικό μικρόφωνο της συστοιχίας διάρκειας 30ms, από τη στιγμή 0,75sec έως την 0,78sec. Το σήμα αυτό μπορεί να μοντελοποιηθεί ως $x(t) = s(t) + u(t)$ με :

- $s(t)$: καθαρό σήμα φωνής
- $u(t)$: θόρυβος

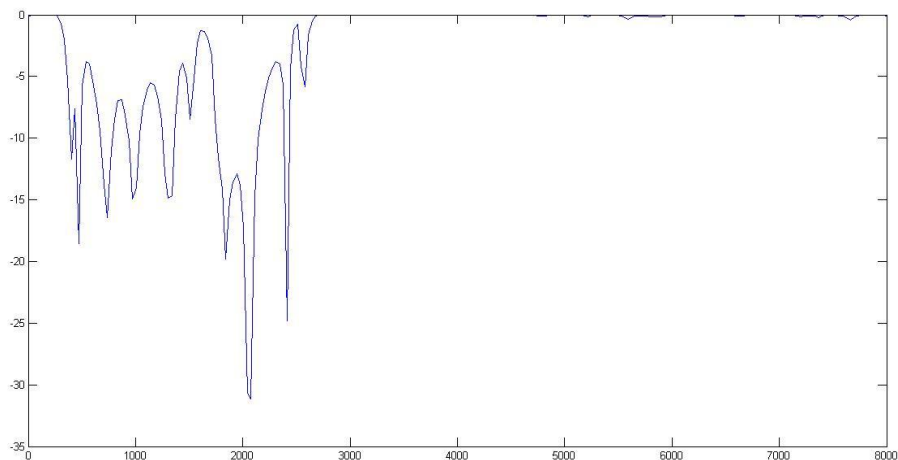
Βηματική Ανάλυση

✚ Αρχικά, πραγματοποιείται κατασκευή της απόκρισης συχνότητας του IIR φίλτρου Wiener με τη βοήθεια της σχέσης:

$$H_w(\omega) = 1 - \frac{P_u(\omega)}{P_x(\omega)} \text{ με:}$$

- $P_u(\omega)$: φάσματα ισχύος του θορύβου
- $P_x(\omega)$: φάσματα ισχύος του σήματος στο κεντρικό μικρόφωνο της συστοιχίας

Σύμφωνα με τα παραπάνω κατασκευάζεται το Wiener φίλτρο και παρακάτω παρατίθεται η γραφική παράσταση της απόκρισης συχνότητας του [0 KHz, 8 KHz].

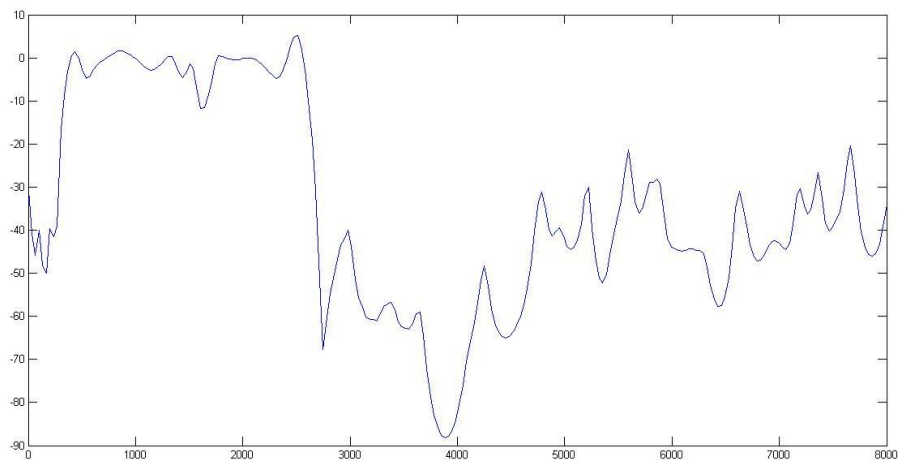


Απόκριση συχνότητας του Wiener φίλτρου σε dB για συχνότητες από 0 έως 8kHz

✚ Το φίλτρο Wiener θα προκαλέσει παραμόρφωση στο σήμα της οποίας η επίδραση θα υπολογιστεί μέσω του speech distortion index το οποίο υπολογίζεται από την παρακάτω σχέση:

$$n_{sd}(\omega) = |1 - H_w(\omega)|^2$$

Το speech distortion index υπολογίστηκε μέσω του Matlab και παρακάτω παρουσιάζεται η γραφική του παράσταση σε dB από 0 έως 8kHz.

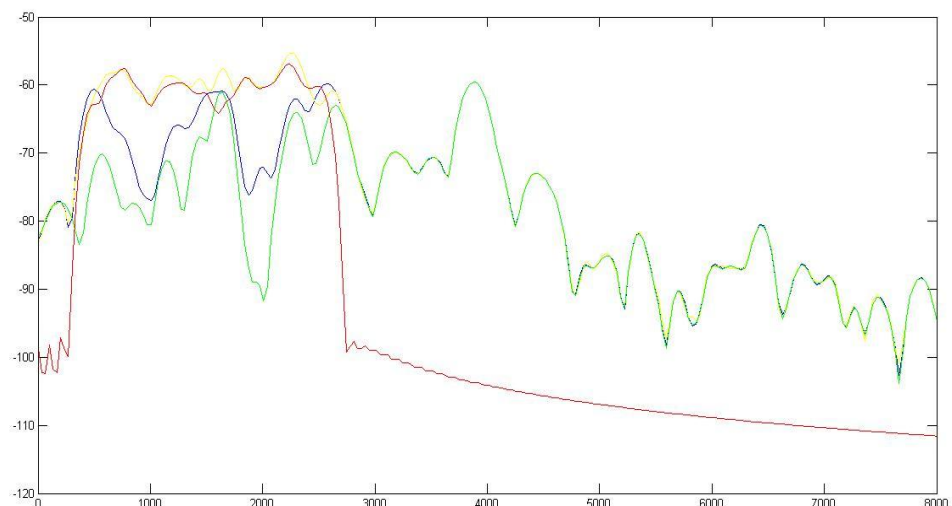


Speech distortion index σε dB για συχνότητες από 0 έως 8kHz

Από το παραπάνω διάγραμμα βλέπουμε ότι ο δείκτης έχει τιμή κοντά στα 0dB από τα 400 μέχρι και τα 2500Hz πράγμα που σημαίνει ότι σε αυτές τις συχνότητες ο δείκτης αυτός είναι κοντά στη μονάδα δηλαδή το φίλτρο θα προκαλέσει εκεί παραμόρφωση στο σήμα εισόδου του. Στις άλλες περιοχές συχνοτήτων παρατηρούνται αρνητικές τιμές (σε dB), δηλαδή έχει τιμή κοντά στο 0 και επομένως δεν θα παραμορφώσει το σήμα. Είναι λοιπόν πολύ πιθανό να έχουμε παραμόρφωση στην προαναφερθείσα ζώνη συχνοτήτων (400Hz-2,5kHz) και ιδιαίτερα κοντά στα 2kHz, όπου ο δείκτης λαμβάνει την υψηλότερη τιμή του.

✚ Στη συνέχεια εφαρμόζεται στο θορυβοποιημένο σήμα του κεντρικού μικροφώνου της συστοιχίας το Wiener φιλτράρισμα. Ακολούθως, παρατίθενται σε κοινό διάγραμμα τα φάσματα ισχύος του

- καθαρού σήματος φωνής $s(t)$
- της θορυβώδους εισόδου $x(t)$
- της εξόδου του φίλτρου Wiener
- του θορύβου.



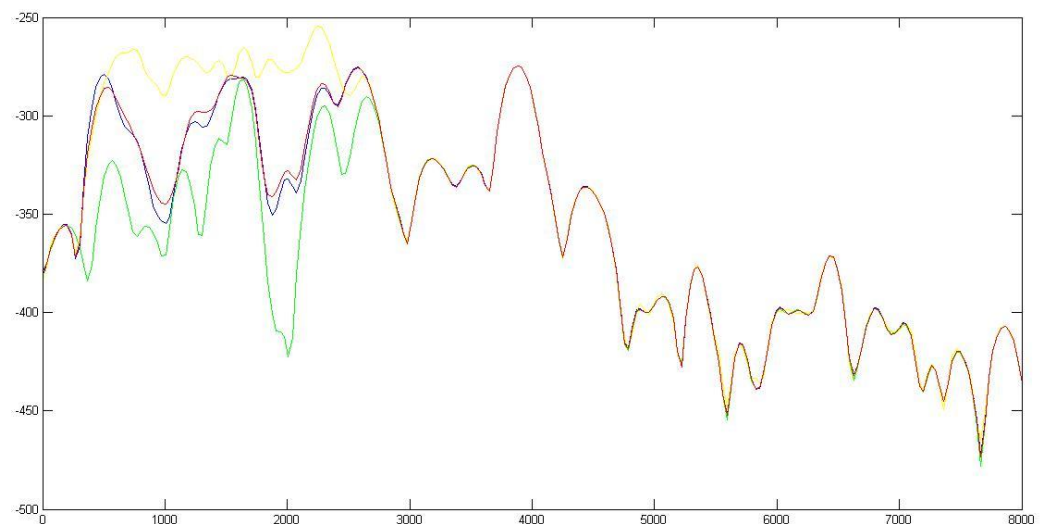
Όπως ήταν αναμενόμενο, το φάσμα ισχύος του σήματος εξόδου παρουσιάζει μία παραμόρφωση από του αρχικού γύρω από τα 2kHz, γεγονός που οφείλεται στη συμπεριφορά του φίλτρου. Επίσης, επειδή ο θόρυβος είναι συγκεντρωμένος στο διάστημα [0.5kHz , 2.5kHz], όπως φαίνεται και στο σχήμα, το φάσμα ισχύος του θορυβώδους σήματος αποκλίνει από το πραγματικό σε αυτήν την περιοχή. Στις περιοχές, όπου δεν έχουμε θόρυβο και το φίλτρο δεν προκαλεί παραμόρφωση, το φάσμα ισχύος και του καθαρού και του θορυβώδους σήματος όπως και της εξόδου σχεδόν ταυτίζονται.

✚ Μέσω του Matlab βρίσκουμε ότι το SNR της εισόδου είναι -3.0666dB, ενώ του σήματος που προκύπτει από την έξοδο του Wiener φίλτρου είναι 5.0403dB. Παρατηρούμε ότι το σήμα στην έξοδο παρουσίασε βελτίωση στο SNR κατά 8.1069dB.

Εφαρμόζουμε το πλαίσιο με το θόρυβο σε σύστημα delay-and-sum beamforming και βρίσκουμε ότι το SNR της εξόδου του είναι 11.8039dB, δηλαδή είχαμε βελτίωση κατά 14.8705dB σε σχέση με το αρχικό σήμα.

Συνεχίζοντας παρατίθενται σε κοινή γραφική παράσταση τα φάσματα ισχύος του

- καθαρού σήματος φωνής $s(t)$
- της θορυβώδους εισόδου $x(t)$
- της εξόδου του φίλτρου Wiener
- της εξόδου της πολυκαναλικής μεθόδου



Παρατηρούμε ότι το φάσμα ισχύος του προκύπτοντος σήματος από την πολυκαναλική μέθοδο ακολουθεί καλύτερα αυτό του καθαρού σήματος φωνής σε σχέση με αυτό της εξόδου της μονοκαναλικής μέθοδου Wiener γεγονός που γίνεται σημαντικά αισθητό γύρω από την συχνότητα των 2KHz στην οποία παρατηρήθηκε και προηγούμενως-στο κομμάτι υπολογισμού του *speech distortion index*- ότι προσθέεται σημαντική παραμόρφωση από το φίλτρο Wiener. Αυτό λοιπόν συμβαίνει επειδή το beamforming δεν προκαλεί παραμορφώσεις όπως το φίλτρο Wiener.

2.2 Beamforming σε πραγματικά σήματα

A) Delay-and-Sum Beamforming

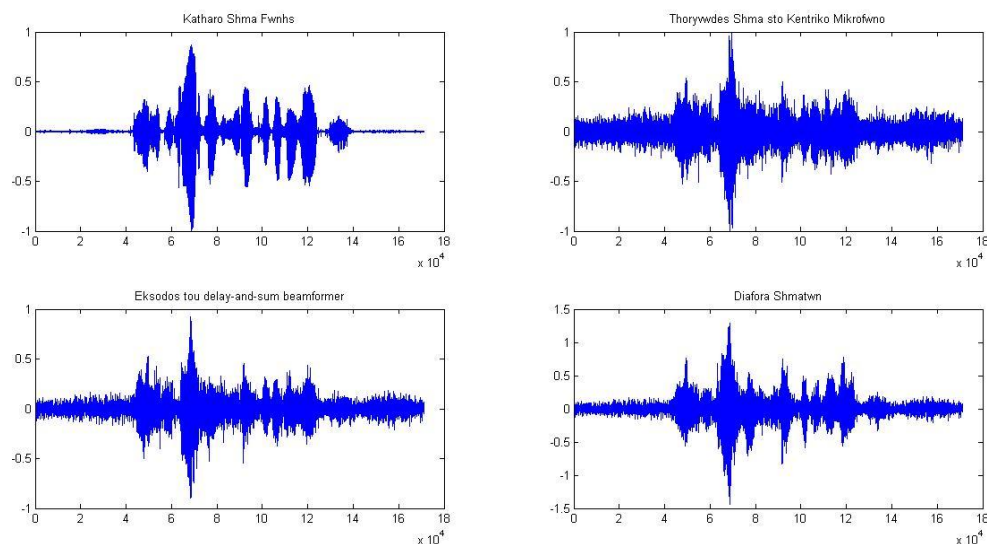
Θεωρούμε πάλι μια γραμμική συστοιχία $N=7$ μικροφώνων με απόσταση $d=8\text{cm}$ τα οποία βρίσκονται σε θορυβώδες δωμάτιο και καταγράφουν το σήμα που εκφωνεί ένας άνθρωπος σε γωνία $\theta = 90^\circ$.

Όπως και πριν έτσι και τώρα αποθορυβοποιείται το σήμα φωνής με τη μέθοδο delay-and-sum beamforming.

Βηματική Ανάλυση

✚ Χρησιμοποιούμε ξανά τη συνάρτηση *Shift* με σκοπό τον υπολογισμό των βαρών και εφαρμόζουμε beamforming απλά για 7 νέα σήματα.

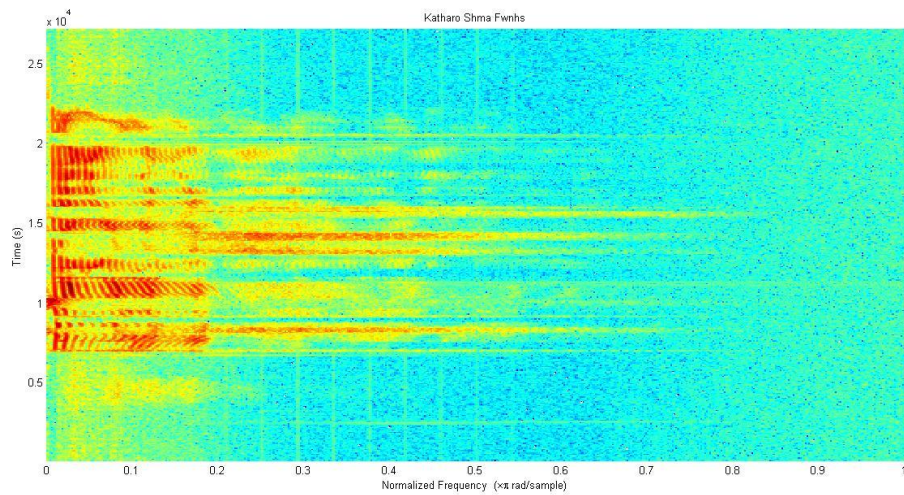
✚ Στη συνέχεια απεικονίζονται οι κυματομορφές για τα εξής σήματα:



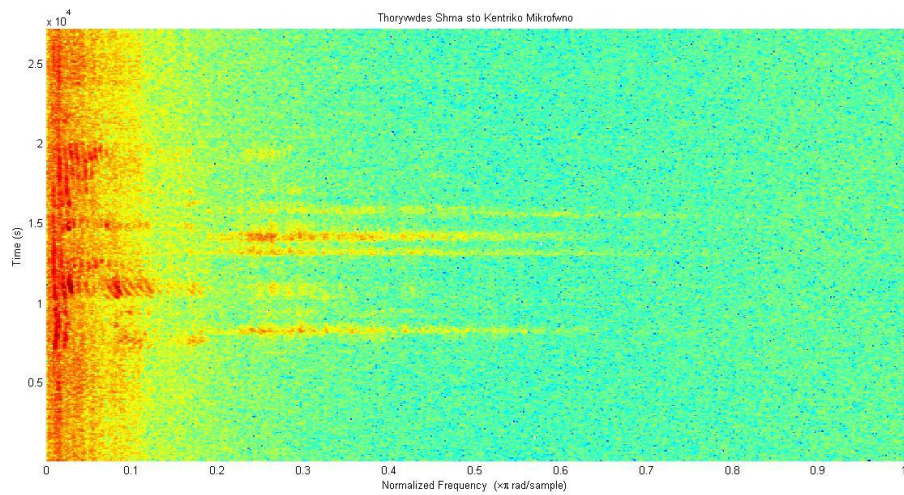
Κυματομορφές για το καθαρό σήμα φωνής, το θορυβώδες σήμα στο κεντρικό μικρόφωνο, της έξοδο του beamformer και το θόρυβο

Παρατηρείται ότι ο delay-and-sum beamformer δεν είναι εξίσου αποδοτικός σε αυτή την περίπτωση, αφού στην περίπτωση του diffuse noise field τα σήματα θορύβου εμφανίζουν μεγάλη συσχέτιση μεταξύ μικροφώνων.

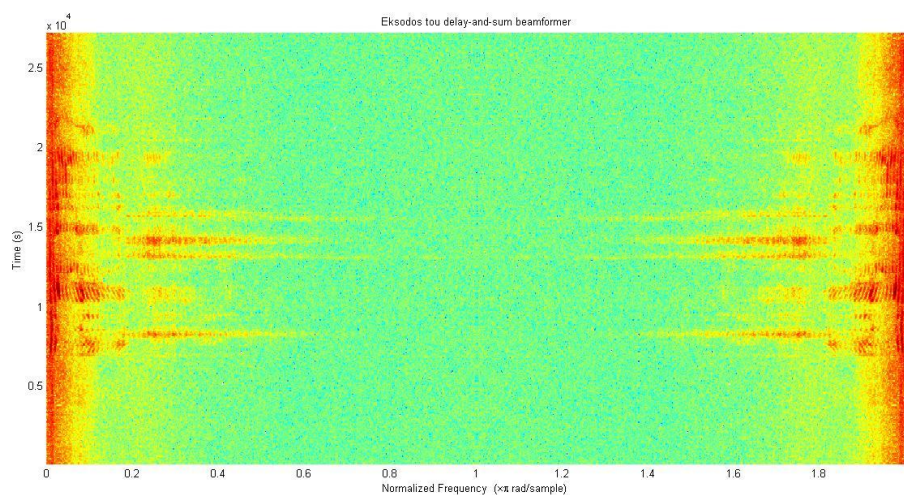
Spectrograms



Σπεκτρογράφημα καθαρού σήματος φωνής



Σπεκτρογράφημα θορυβώδους σήματος στο κεντρικό μικρόφωνο



Σπεκτρογράφημα σήματος από την έξοδο του beamformer

✚ Επειδή το ολικό SNR δεν μπορεί να μας δώσει σαφή αντίληψη για την ακουστική «εντύπωση» ποιότητας που έχει ο άνθρωπος για το σήμα φωνής θα υπολογίσουμε και το segmental SNR (SSNR) το οποίο είναι κάτι σαν ένα μέσο όρο SNR των πλαισίων βραχέως χρόνου του σήματος φωνής:

$$SSNR = \frac{1}{M} \sum_{m=0}^{M-1} \log_{10} \frac{\sum_{n=Lm}^{Lm+L-1} s^2(n)}{\sum_{n=Lm}^{Lm+L-1} v^2(n)} \text{ με:}$$

- L: το μήκος των πλαισίων,
- M: το πλήθος των πλαισίων
- s(n): το σήμα φωνής
- v(n): θόρυβος

Η ισχύς του σήματος θορύβου υπολογίζεται μόνο μέσω των αρχικών δειγμάτων αφού στην αρχή του σήματος είναι μεγάλη η παρουσία θορύβου.

Αν θεωρήσουμε ότι ο θόρυβος είναι στάσιμος τότε η ισχύς του μένει αμετάβλητη χρονικά επομένως η τιμή που υπολογίζεται αρχικά διατηρείται ίδια στους υπολογισμούς.

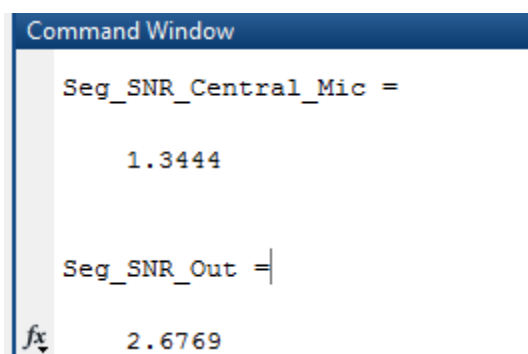
Η συνάρτηση υπολογισμού του SSNR που δημιουργήθηκε, ονομάζεται *SegmentalSNR* στην οποία εισάγεται ως παράμετρος το πλήθος M των πλαισίων, υπολογίζεται το SNR για κάθε πλαίσιο και τελικά λαμβάνεται ο μέσος όρος.

Το SSNR στο κεντρικό μικρόφωνο υπολογίζεται:

$$SSNR_{\text{central}} = 1.3444$$

ενώ το SSNR στην έξοδο του του beamformer είναι:

$$SSNR_{\text{beamformer}} = 2.6769$$



```
Command Window

Seg_SNR_Central_Mic =

    1.3444

Seg_SNR_Out =

fx 2.6769
```

Συμπέρασμα

Παρατηρούμε ότι το SSNR έχει βελτιωθεί αλλά δεν έχει μεγάλη διαφορά από το SSNR του σήματος που περιέχει θόρυβο.

Αυτό διαπιστώνεται και ακουστικά καθώς τα δύο σήματα συμπεραίνουμε ότι είναι σχεδόν ίδια.

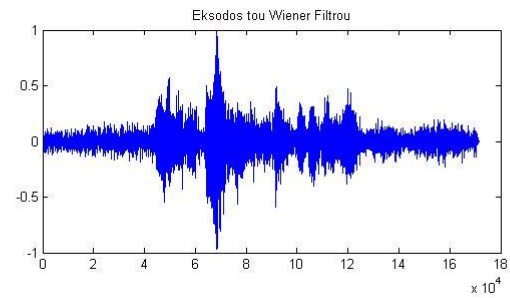
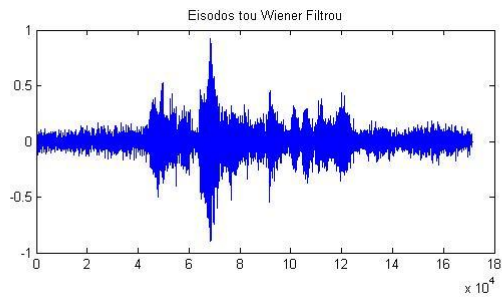
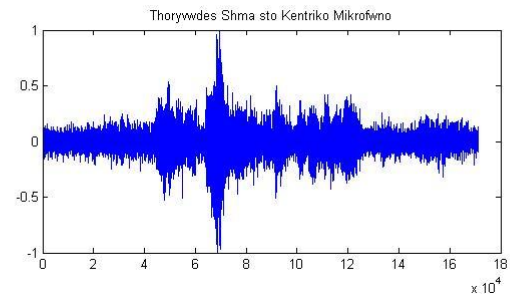
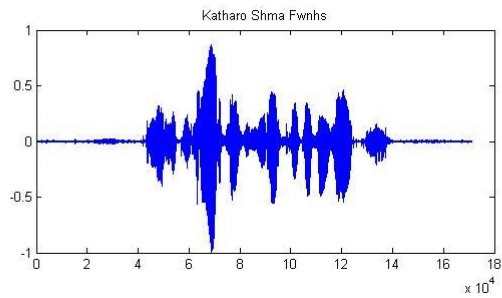
Συνεπώς επιβεβαιώνεται και πάλι ότι ο delay-and-sum beamformer δεν έχει καλή απόδοση στην περίπτωση του diffuse noise.

B)Post-filtering με Wiener φίλτρο

Θεωρούμε κι εδώ μια γραμμική συστοιχία $N=7$ μικροφώνων με απόσταση $d=8\text{cm}$ τα οποία βρίσκονται σε θορυβώδες δωμάτιο και καταγράφουν το σήμα που εκφωνεί ένας άνθρωπος σε γωνία $\theta = 90^\circ$. Θα εφαρμόσουμε εδώ αυτή τη μέθοδο του Post-Filtering με Wiener φίλτρο γιατί, όπως επισημαίνεται και στην εκφώνηση της άσκησης στην περίπτωση του diffuse noise field, ο delay-and-sum beamformer δεν έχει καλή απόδοση. Αυτό το γεγονός οφείλεται στο ότι τα σήματα θορύβου εμφανίζουν μεγάλη συσχέτιση μεταξύ μικροφώνων, ειδικά στις χαμηλές συχνότητες.

Βηματική Ανάλυση

- ✚ Χρησιμοποιούμε εδώ τη συνάρτηση *Shift* με σκοπό τον υπολογισμό των βαρών και εφαρμόζουμε beamforming απλά για τα 7 σήματα που είχαμε και στο προηγούμενο ερώτημα. Έτσι έχουμε ένα νέο σήμα που παίρνουμε από την έξοδο του beamformer.
 - ✚ Επίσης, από το αρχικό μας σήμα, βρίσκουμε με τη συνάρτηση *pwelch* του matlab το φάσμα ισχύος του θορύβου, από το πρώτο παράθυρο του σήματος όπου υπάρχει μόνο θόρυβος. Αυτό το κάνουμε καθώς θεωρούμε πως ο θόρυβος είναι στάσιμος, και εξ αυτού το φάσμα ισχύος του δε μεταβάλλεται με το χρόνο.
 - ✚ Στη συνέχεια “παραθυρώνουμε” το νέο μας σήμα με παράθυρα *hamming* μήκους 30 ms και *overlap* 50%. Για κάθε ένα από αυτά τα παράθυρα, υπολογίζουμε το φάσμα ισχύος του και βρίσκουμε τη συνάρτηση μεταφοράς του κατάλληλου φίλτρου Wiener, όπως πράξαμε και σε προηγούμενο ερώτημα της παρούσας εργαστηριακής άσκησης. Φιλτράρουμε, λοιπόν το κάθε παράθυρο με το αντίστοιχο Wiener φίλτρο και με *overlap-add* βρίσκουμε το νέο μας σήμα που είναι στην έξοδο της συστοιχίας των φίλτρων.
- Στις επόμενες σελίδες παρατίθενται τα αποτελέσματα που ζητούνται:

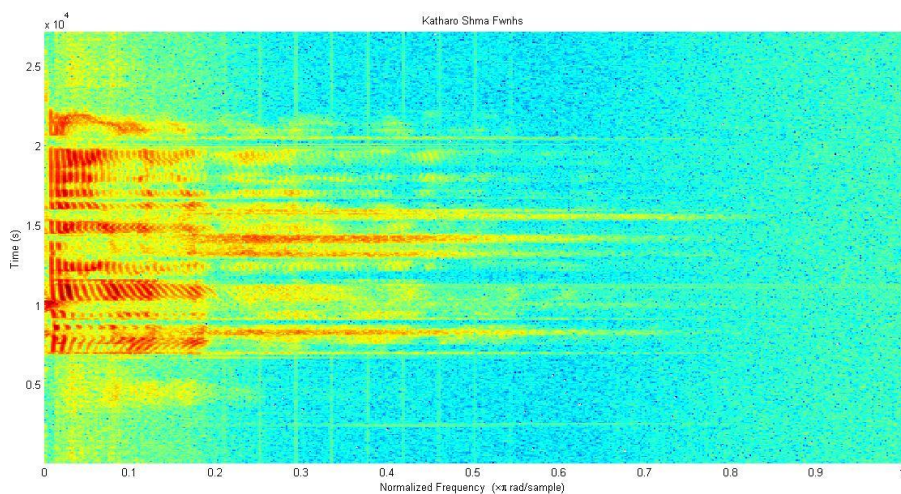


Κυματομορφές για το καθαρό σήμα φωνής, το θορυβώδες σήμα στο κεντρικό μικρόφωνο, της είσοδο του Wiener φίλτρου και την έξοδο του Wiener φίλτρου.

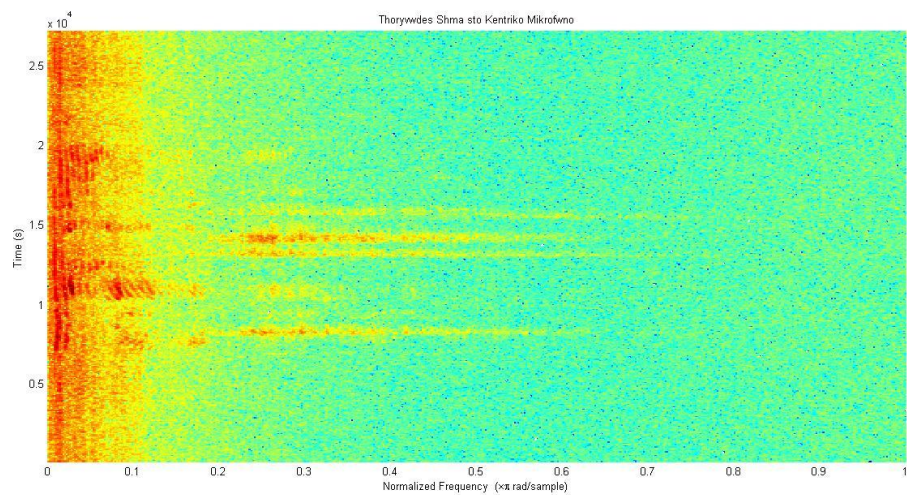
Παρατηρούμε πως με το Wiener φιλτράρισμα έχουμε μια σχετική βελτίωση του τελικού μας σήματος σε σχέση με την απλή εφαρμογή της μεθόδου του Delay-and-Sum beamformer.

Παρακάτω παρατίθενται και στα σπεκτρογραφήματα των παραπάνω σημάτων:

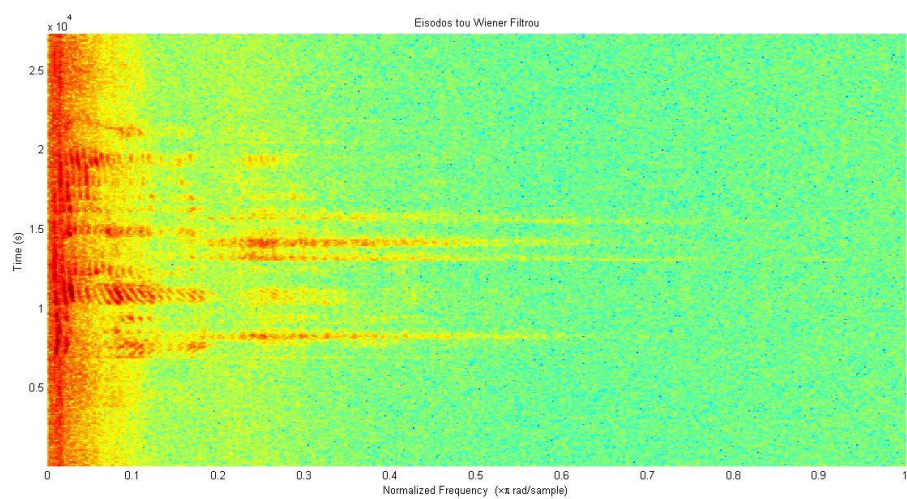
Spectrograms



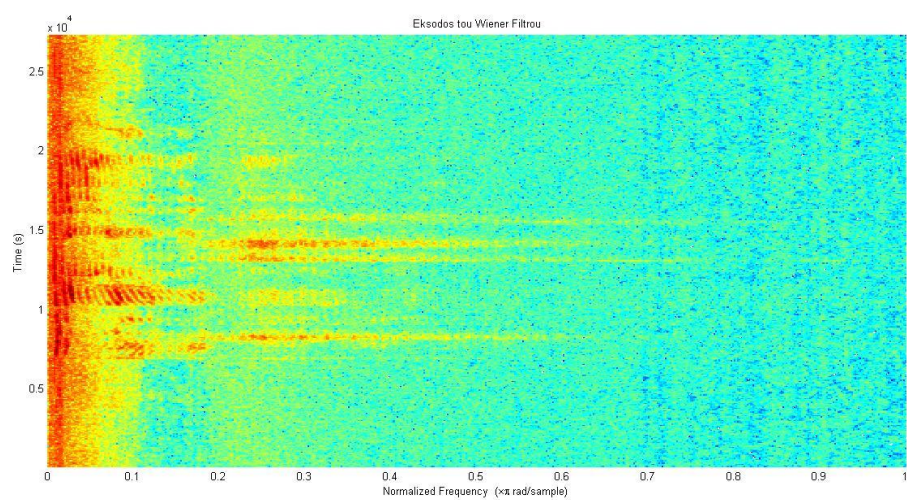
Σπεκτρογράφημα καθαρού σήματος φωνής



Σπεκτρογράφημα θορυβώδους σήματος στο κεντρικό μικρόφωνο



Σπεκτρογράφημα εξόδου beamformer – εισόδου Wiener φίλτρου



Σπεκτρογράφημα εξόδου Wiener φίλτρου

Παρατηρούμε ότι μεταξύ του 3^{ου} και του 4^{ου} σπεκτρογραφήματος υπάρχει ελάττωση του πλάτους στις υψηλές συχνότητες (αρχικά ήταν πράσινες-κίτρινες ενώ μετά έγιναν γαλάζιες) από το οποίο μπορούμε να συμπεράνουμε την ελάττωση του θορύβου στο σήμα μας μετά το wiener φιλτράρισμα.

Στη συνέχεια υπολογίσαμε τις τιμές των SSNRs που ζητούνται, πάλι χρησιμοποιώντας την ίδια συνάρτηση με προηγουμένως. (Segmental_SNR)

Αυτή μας έδωσε τα εξής αποτελέσματα:

- $SSNR_{\text{εισόδου στο φίλτρο}} = 3,3018$
- $SSNR_{\text{εξόδου από το φίλτρο}} = 4,1589$

Παρατηρούμε ότι υπάρχει μια σχετική βελτίωση μεταξύ των 2 SSNRs. Ωστόσο η διαφορά τους δεν είναι και πολύ μεγάλη. Παρόλα αυτά υπάρχει βελτίωση με το Wiener φιλτράρισμα σε σχέση με τη μέθοδο delay and sum Beamforming.

Στη συνέχεια υπολογίσαμε και το μέσο όρο των SSNRs των 7 σημάτων από τα μικρόφωνα και του $SSNR_{\text{εισόδου στο φίλτρο}}$ και βρήκαμε ότι ο ζητούμενος μέσος όρος έχει την τιμή 1,4887.

Η συνολική διαφορά από το μέσο όρο αυτών των SSNR και του τελικού μας σήματος είναι:

$$4,1589 - 1,4687 = 2,6902$$

Τα παραπάνω αποτελέσματα προήλθαν από το matlab:

```
Command Window

SSNR_before_Wiener_Filter =

    3.3018

|
SSNR_after_Wiener_Filter =

fx 4.1589

meanSSNR =

    1.4687

fx >> |
```