

ΕΜΠ

ΣΗΜΜΥ

ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΩΝ

3^η ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ

ΑΚ.ΕΤΟΣ : 2014-2015

ΠΑΡΑΔΟΤΕΑ: 19-6-15

ΟΝΟΜΑ:

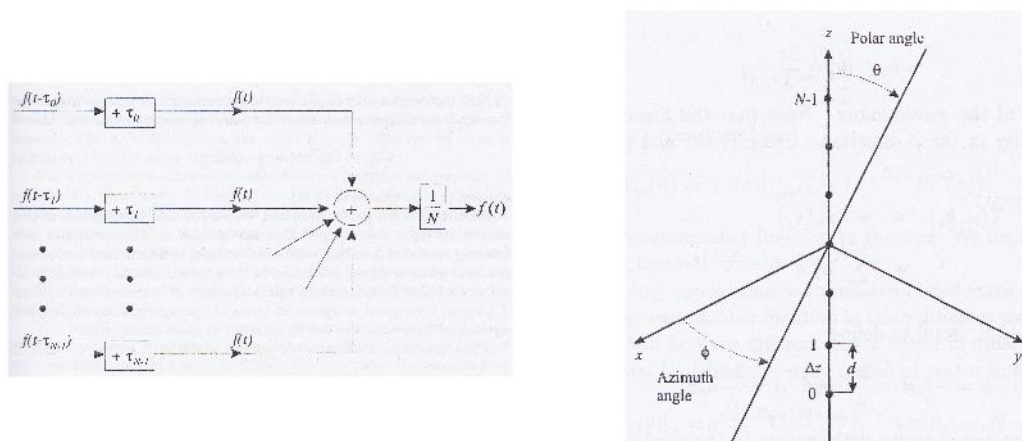
ΚΑΛΑΒΡΥΤΙΝΟΣ ΧΑΡΑΛΑΜΠΟΣ ΑΜ: 3112024

ΘΕΜΑ:

«Συστοιχίες μικροφώνων (Microphone Arrays)
και Πολυκαναλική επεξεργασία
Σημάτων (Multichannel signal Processing).»

ΜΕΡΟΣ 1. Συστοιχίες μικροφώνων και χωρικό φιλτράρισμα (Spatial filtering).

Η χρήση μικροφώνων στο χώρο δίνει τη δυνατότητα καταγραφής και αξιοποίησης τόσο των χρονικών όσο και των χωρικών χαρακτηριστικών των ακουστικών σημάτων. Στο πρώτο μέρος της εργαστηριακής άσκησης είδαμε σε βάθος τον τρόπο φιλτραρίσματος από πολλαπλά κανάλια και την παραγωγή της εξόδου γτ (beamforming). Για την περιγραφή της χωρικής απόκρισης ενός **beamformer** χρησιμοποιείται το beam pattern. Για αυτή την άσκηση θεωρούμε ότι το ηχητικό κύμα είναι επίπεδο, δηλαδή βρισκόμαστε μακριά από την πηγή του σήματος. Ακόμα, τα μικρόφωνα ισαπέχουν και βρίσκονται σε γραμμική διάταξη, οπότε η διάταξη καλείται **ομοιόμορφη γραμμική συστοιχία (ULA)**.



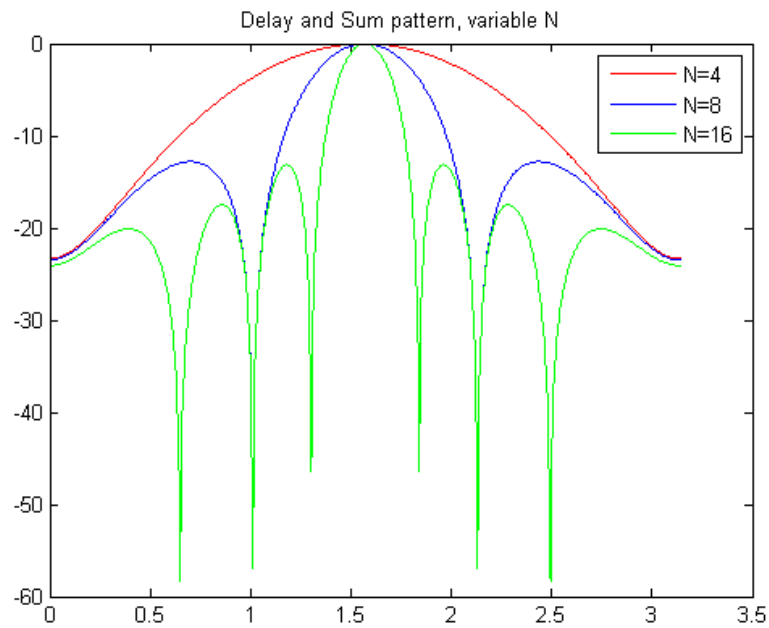
Σχ.1 Delay and sum beamformer (αριστερά) και ομοιόμορφη γραμμική συστοιχία (δεξιά)

Αρχικά θεωρούμε ότι το σήμα φτάνει στη συστοιχία των μικροφώνων με γωνία $\theta_s = 90^\circ$ και ότι η συχνότητα είναι ίση με $f=2\text{kHz}$. Ζητείται το μέτρο του delay-and-sum pattern σε λογαριθμική κλίμακα συναρτήση της γωνίας $\theta=[0, 180^\circ]$ στις ακόλουθες περιπτώσεις:

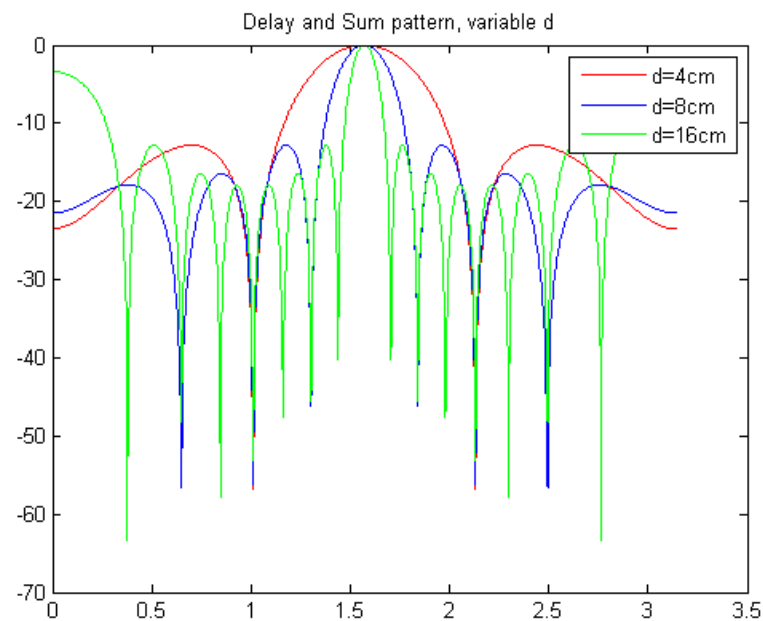
- 1) $d=0.08\text{m}$, $N=4, 8, 16$, όπου d η απόσταση μεταξύ των μικροφώνων και N ο αριθμός των μικροφώνων
- 2) $d=0.04, 0.08, 0.16 \text{ m}$, $N=8$

Στις δύο παραπάνω περιπτώσεις συγκρίνουμε τη μορφή του pattern για διάφορες τιμές στις αποστάσεις μεταξύ των μικροφώνων d και για διαφορετικό αριθμό μικροφώνων N , κρατώντας το ένα από τα δύο σταθερό κάθε φορά.

Τα αποτελέσματα φαίνονται παρακάτω



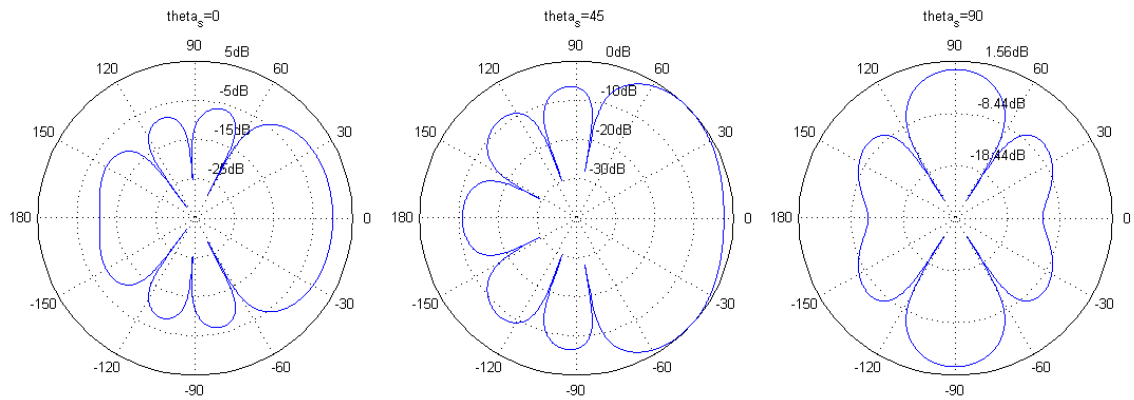
Σχ2.1 Μέτρο του DAS pattern για διαφορετικά N (αριθμός μικροφώνων).



Σχ2.2 Μέτρο του DAS pattern για διαφορετικά d (αποστάσεις ανάμεσα στα μικρόφωνα).

Στα δύο παραπάνω Σχήματα (2.1 και 2.2) κάνουμε τις εξής παρατηρήσεις: Όσο διατηρούμε σταθερή την απόσταση d , με την αύξηση του N αυξάνεται το πλήθος των side lobes, και μειώνεται το πλάτος του κυρίου και των υπολοίπων λοβών. Η αύξηση της απόστασης μεταξύ των μικροφώνων οδηγεί σε στενότερους λοβούς, και περισσότερους grating lobes. Γενικά, προσπαθούμε ο κεντρικός λοβός να έχει μικρό πλάτος, και τα grating lobes να είναι μικρά. Οι πλευρικοί λοβοί δηλώνουν ότι δημιουργείται spatial aliasing, και ο beamformer αδυνατεί να διαχωρίσει το σήμα που προέρχεται από τη steering direction, από τα υπόλοιπα που περνούν από τους άλλους λοβούς.

Στο επόμενο κομμάτι της άσκησης θεωρούμε ότι η συστοιχία αποτελείται από $N=8$ μικρόφωνα, με απόσταση $d=4\text{cm}$. Για την ίδια συχνότητα 2000 κύκλων, σχεδιάζουμε το μέτρο του Delay and sum pattern για $\theta_s = 0^\circ, 45^\circ, 90^\circ$.

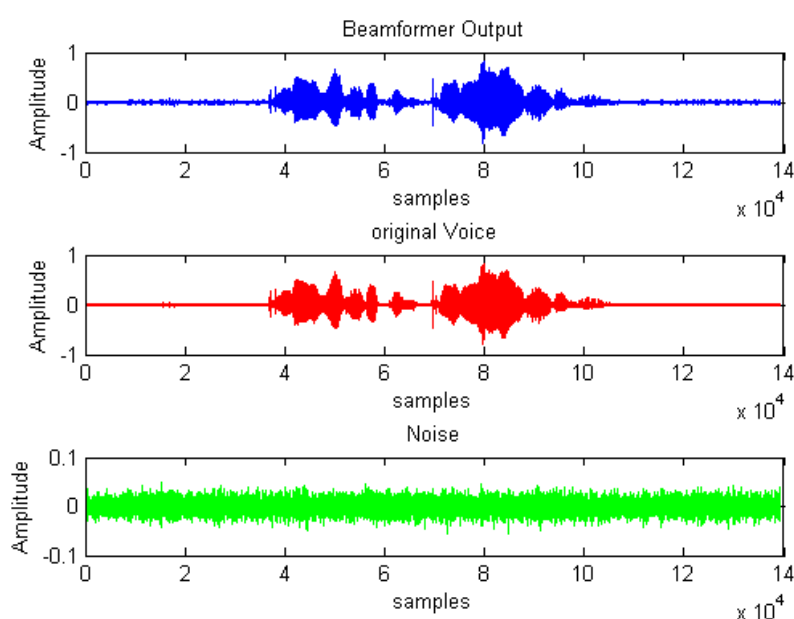


Σχ.3. Μέτρο του DAS beam pattern σε λογαριθμική κλίμακα σε πολικό διάγραμμα, συναρτήσει της γωνίας θ

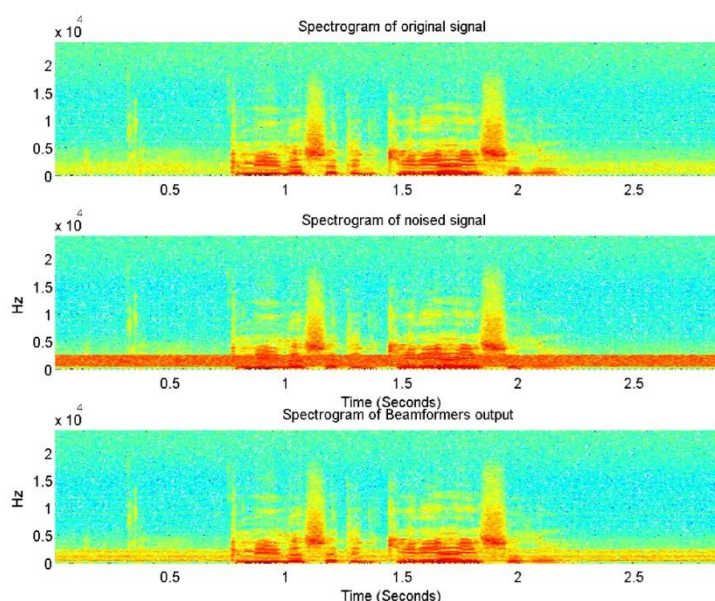
Στα παραπάνω σχήματα παρατηρούμε ότι οι κύριοι λοβοί δείχνουν κάθε φορά στη γωνία θ_s . Στο πρώτο σχήμα παρατηρούμε ότι ο κύριος λοβός έχει κατεύθυνση τις 0 μοίρες, και στο τελευταίο δείχνει τις 90 μοίρες. Στο μεσαίο διάγραμμα η κατεύθυνση του κυρίου λοβού δεν είναι τόσο εμφανής, αλλά αν παρατηρήσουμε το πάνω μισό στο διάγραμμα, φαίνεται ότι ο κύριος λοβός δείχνει τις 45 μοίρες.

ΜΕΡΟΣ 2. Beamforming

Θεωρούμε γραμμική συστοιχία με $N=7$ και $d=4\text{cm}$. Η πηγή παράγει σήμα φωνής με γωνία 45 μοιρών από τη συστοιχία. Ο θόρυβος είναι ζωοπερατός και η ενέργειά τους είναι συγκεντρωμένη στα 50-2500Hz. Παρακάτω φαίνονται οι κυματομορφές εξόδου του beamformer και του καθαρού σήματος φωνής όπως δόθηκε. Στο ίδιο διάγραμμα φαίνεται και η διαφορά τους. Από τη μορφή της διακρίνουμε ότι δεν υπάρχει καθόλου σήμα, οπότε η ευθυγράμμιση των σημάτων έγινε σωστά. Παρατηρούμε ακόμα ότι οι δύο προηγούμενες κυματομορφές βρίσκονται πολύ κοντά μεταξύ τους. Αυτό γίνεται κυρίως επειδή τα σήματα είναι προσομοιωμένα, οπότε η αφαίρεση του θορύβου με αυτή τη μέθοδο, είναι αρκετά αποτελεσματική.



Σχ.4.1 Κυματομορφές εξόδου του beamformer και καθαρό σήμα.



Σχ.4.2 Σπεκτρογραφήματα του καθαρού σήματος, του θορυβώδους σήματος, και της εξόδου του beamformer

Σχετικά με το SNR του θορυβώδους σήματος βρέθηκε να ισούται με $x_{snr} = 4.7313\text{dB}$ ενώ το SNR της εξόδου του beamformer βρέθηκε ίσο με $y_{snr} = 17.6106\text{dB}$.

Παρατηρούμε ότι ορθώς το SNR της εξόδου έχει μεγαλύτερη τιμή, πράγμα που σημαίνει ότι το σήμα έχει λιγότερο θόρυβο. Γενικά, το SNR θέλουμε να είναι όσο το δυνατό πιο μεγάλο γίνεται.

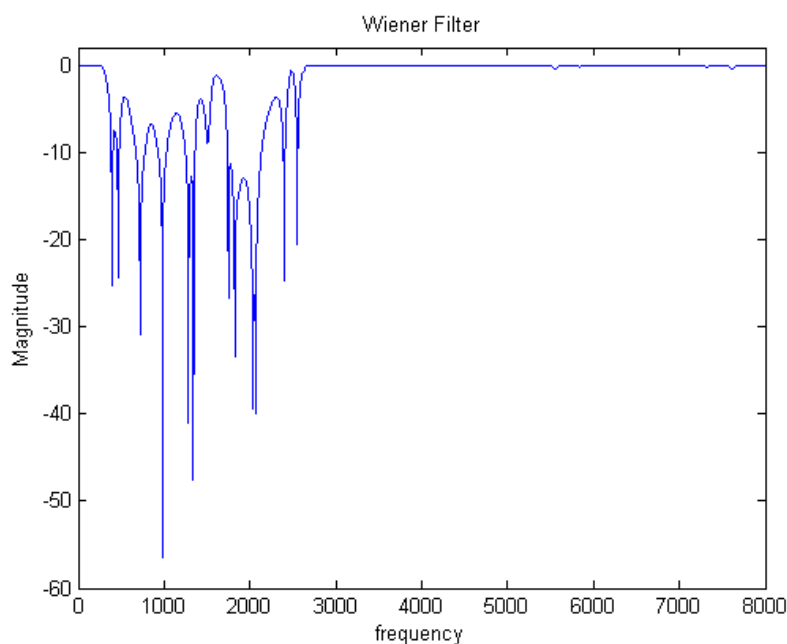
Η έξοδος του beamformer βρίσκεται στο αρχείο με το όνομα `sim_ds.wav`

Β) Στη συνέχεια, συγκρίνουμε τη πολυκαναλική μέθοδο αποθορυβοποίησης που μόλις υλοποιήσαμε, με τη μονοκαναλική μέθοδο με φίλτρο Wiener. Αρχικά, θεωρούμε το πλαίσιο διάρκειας 30ms [0.75,0.78] του σήματος που κατέγραψε το κεντρικό μικρόφωνο. Γνωρίζουμε ότι το πλαίσιο αυτό περιέχει ένα έμφωνο ήχο, και μπορεί να μοντελοποιηθεί ως $x(t)=s(t)+u(t)$ όπου $x(t)$ το σήμα μαζί με το θόρυβο, $s(t)$ το καθαρό σήμα, και $u(t)$ ο θόρυβος. Το καθαρό σήμα μας δίνεται στο φάκελο `material`, (το ίδιο με πριν).

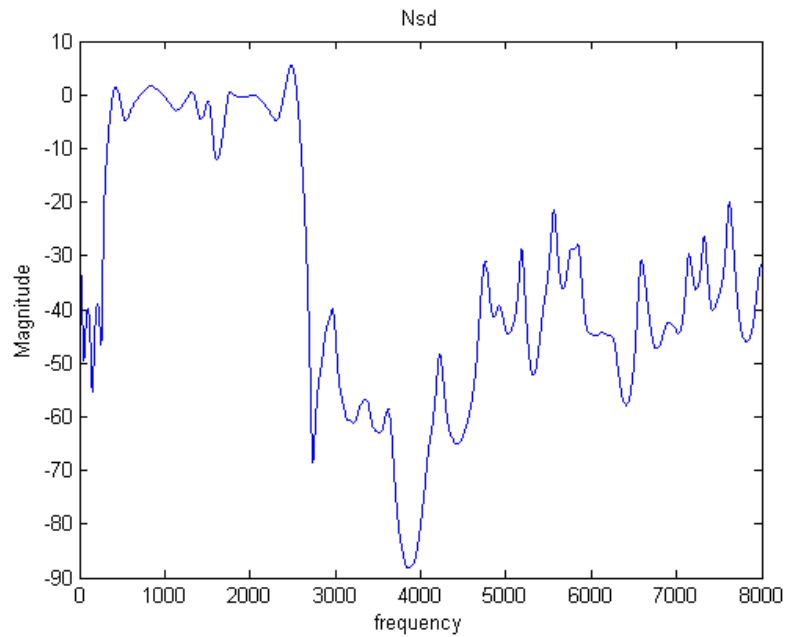
1) Η απόκριση συνόττητα τους φίλτρου Wiener δίνεται ως:

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

Όπου τα φάσματα ισχύος P , υπολογίστηκαν με τη συνάρτηση `pwelch()` ; . Το φάσμα θορύβου P_u , υπολογίστηκε αφού πρώτα αφαιρέσαμε από το θορυβώδες σήμα το καθαρό (`source`), και έτσι προέκυψε ο θόρυβος. Η απόκριση του φίλτρου σε λογαριθμική κλίμακα φαίνεται παρακάτω.



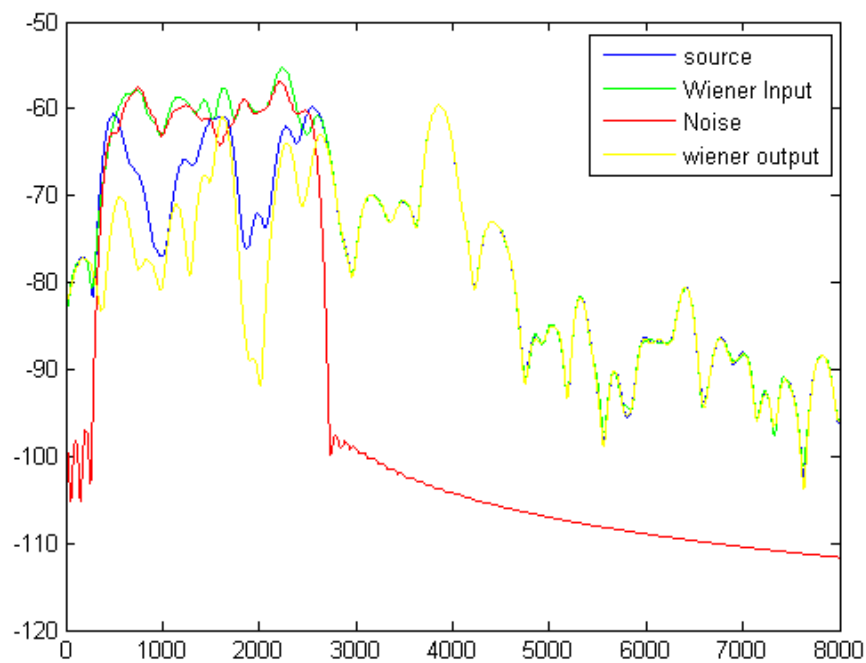
Σχ.5 Απόκριση φίλτρου Wiener για συχνότητα [0,8000]Hz



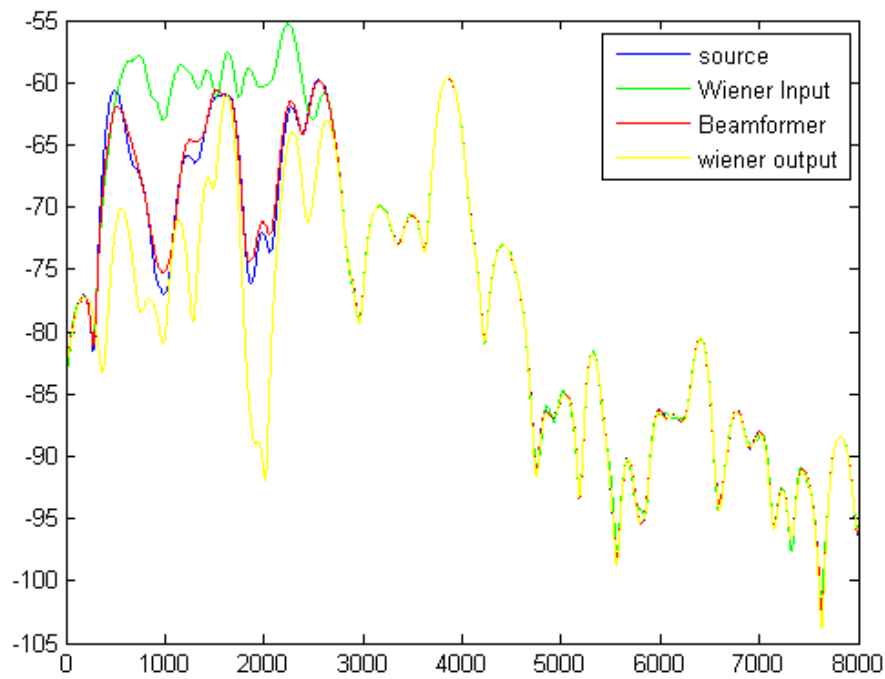
Σχ.6 Nsd σε λογαριθμική κλίμακα

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

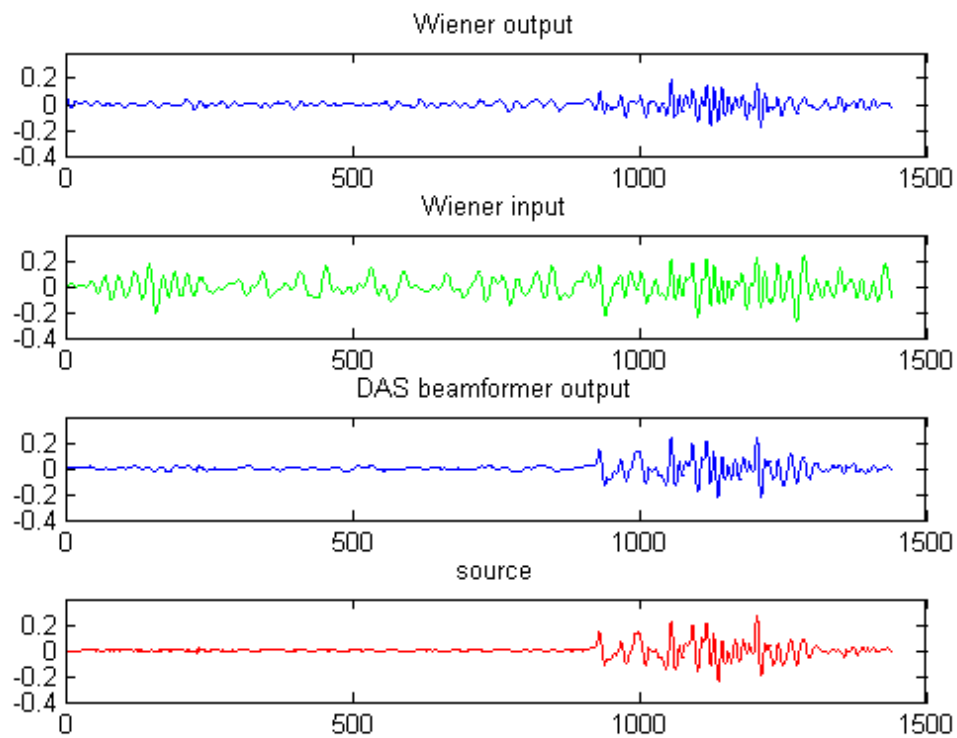
Παρακάτω φαίνονται τα φάσματα ισχύος του καθαρού σήματος, της εισόδου και εξόδου του φίλτρου και του θορύβου. Επιβεβαιώνουμε ότι η ενέργεια του θορύβου βρίσκεται ανάμεσα στα 500-2500Hz.



Σχ.7.1 Φάσματα ισχύος σε λογαριθμική κλίμακα για συχνότητες ως 8000Hz



Σχ.7.2 Φάσματα ισχύος σε λογαριθμική κλίμακα για συχνότητες ως 8000Hz



Σχ.7.3 Κυματομορφές εισόδου και εξόδου του φίλτρου σε σύγκριση με τον beamformer και το καθαρό σήμα

Στα σχήματα 7.1 και 7.2 παρατηρούμε ότι σωστά φαίνεται ότι για συχνότητες μεγαλύτερες του 2500 οι καμπύλες συμπίπτουν, ενώ διαφορά υπάρχει μόνο για f κάτω από 2500 όπου και υπάρχει ο θόρυβος. Το φάσμα του καθαρού σήματος με

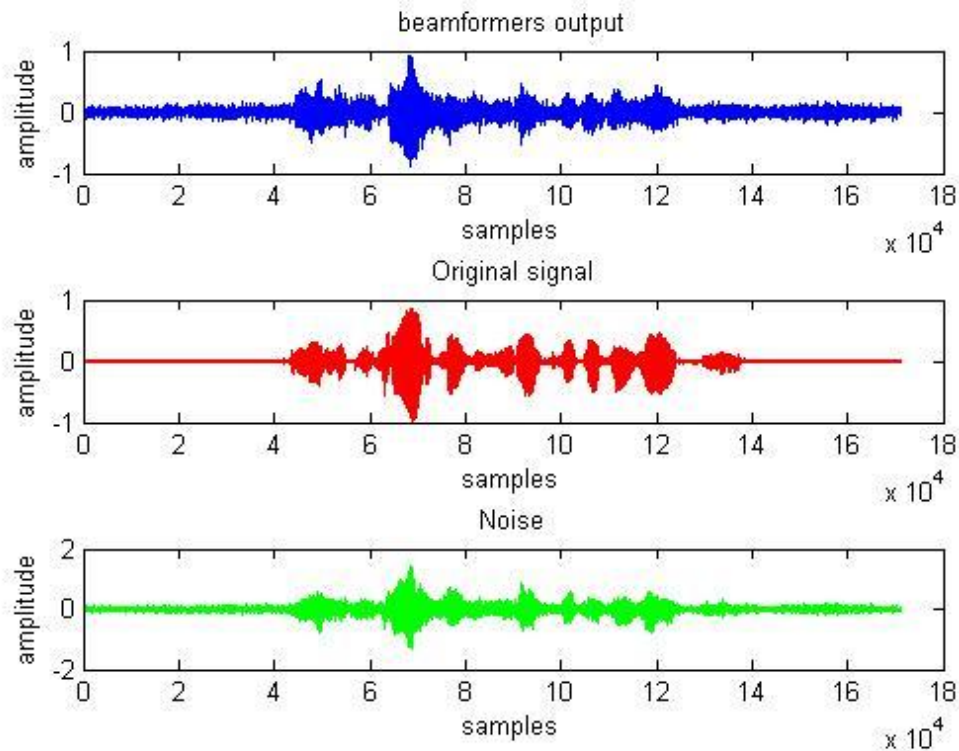
αυτό του σήματος εξόδου του beamformer βρίσκονται πολύ κοντά, εφόσον τα σήματα είναι σχεδόν πανομοιότυπα.

Για τα SNR: Παρατηρούμε ότι οι τιμές στο SNR είναι λογικές, από την άποψη ότι η έξοδος του φίλτρου έχει μεγαλύτερο SNR από την είσοδο. Αυτό σημαίνει ότι ο θόρυβος ελαχιστοποιείται κατά ένα βαθμό. Το SNR της εξόδου είναι κοντά στα 3db και της εισόδου κοντά στο 1.4dB. Το SNR για το συγκεκριμένο πλαίσιο που επιλέχθηκε, με την προηγούμενη μέθοδο είναι στα 11.7 dB.

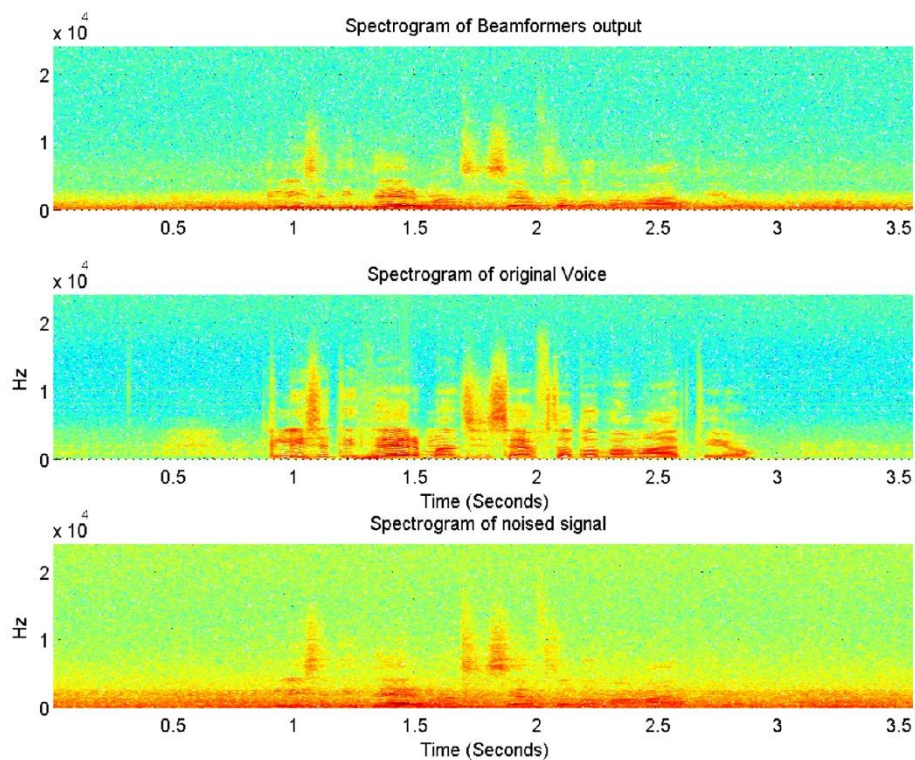
Σημειώνουμε εδώ ότι για το σωστό υπολογισμό των SNRs των κανονικοποιήσαμε πρώτα τα σήματα από ώστε το μέγιστο πλάτος να είναι το 1. Έτσι ο υπολογισμός του λόγου σήματος προς θόρυβο καθίσταται πιο ακριβής. Οι ακριβείς τιμές των SNRs φαίνονται κατά την εκτέλεση του αρχείου.

ΜΕΡΟΣ 2.2. Beamforming σε πραγματικά σήματα

Σε αυτή την περίπτωση τα μικρόφωνα είναι 7, και απέχουν μεταξύ του 8cm. Το σήμα φωνής από την πηγή φτάνει στη συστοιχία σε γωνία 90 μοιρών, οπότε δε θα υπάρχει καθυστέρηση (φτάνει ταυτόχρονα παντού). Χρησιμοποιώντας την ίδια τεχνική με πριν οδηγούμαστε στις παρακάτω γραφικές παραστάσεις. Παρατηρούμε ότι αυτή τη φορά το σήμα εξόδου του **beamformer** περιέχει περισσότερο θόρυβο σε σχέση με πριν, δηλαδή η μέθοδος αυτή δεν είναι το ίδιο αποτελεσματική στην αποθορυβοποίηση πραγματικών σημάτων. Ο θόρυβος σε αυτή την περίπτωση προέρχεται από πολλές διευθύνσεις στο περιβάλλον, και δεν αφαιρέθηκε με την ίδια επιτυχία. Παρ'όλα αυτά, παρατηρείται βελτίωση, η οποία όπως θα δούμε αργότερα μεγιστοποιείται με τη χρήση post filtering.



Σχ.8.1 Κυματομορφές εξόδου του beamformer για πραγματικά σήματα, σε σύγκριση με το καθαρό σήμα και τη διαφορά τους (noise).

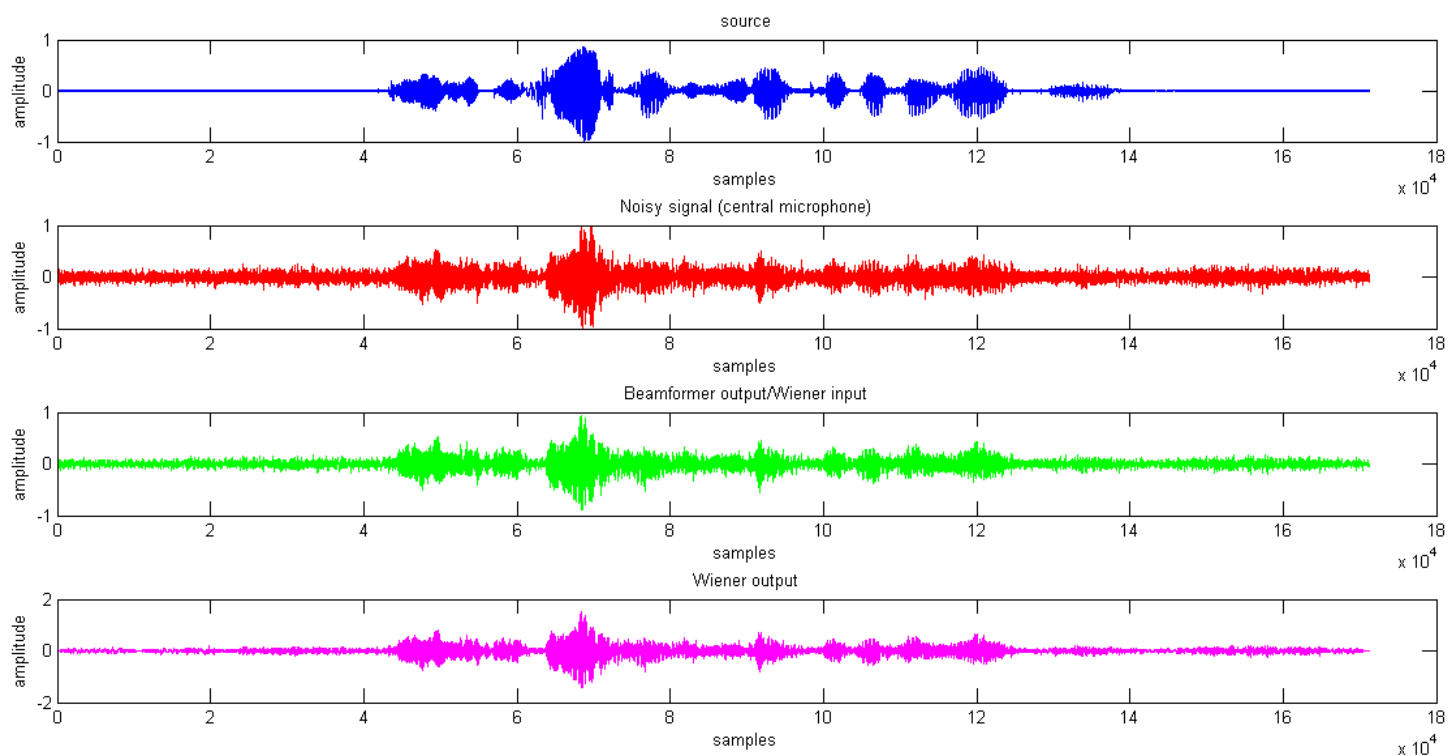


Σχ.8.2 Spectrogram της εξόδου του beamformer, της καθαρής φωνής και του θορυβώδους σήματος (σήμα εισόδου του beamformer)

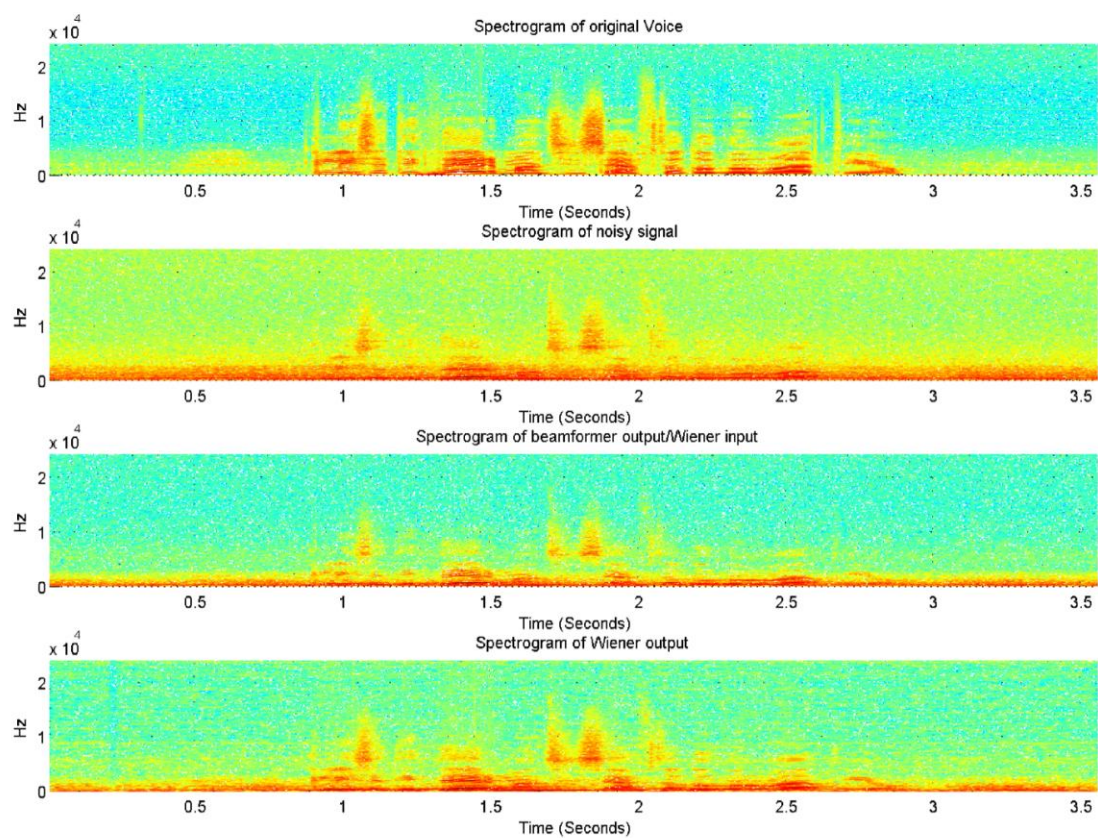
Ζητείται να βρεθεί ο **segmental SNR**, που πρόκειται απλά για ένα μέσο SNR πλαισίων βραχέος χρόνου του σήματος. Για την εύρεσή του χωρίζουμε το σήμα σε frames συγκεκριμένου μήκους, υπολογίζουμε τον SNR και τελικά βρίσκουμε το μέσο όρο. Χρησιμοποιούμε ακόμα κατώφλια των 35db και - 10dB. Χρησιμοποιώντας τη συνάρτηση που υλοποιήσαμε, βλέπουμε ότι υπάρχει μια μικρή βελτίωση στον SNR, η οποία δηλώνει ότι ο θόρυβος στο σήμα δεν αφαιρείται επαρκώς. Σημειώνουμε ότι ο θόρυβος θεωρήθηκε στάσιμος, οπότε υπολογίστηκε μια φορά στην αρχή. Οι τιμές είναι $SSNR_{inp}=4.3131$, $SSNR_{outp}= 6.5022$. Θεωρητικά κάτι τέτοιο φαίνεται σωστό από τη στιγμή που η έξοδος έχει μεγαλύτερη τιμή.

B) Post-filtering Wiener filter

Σε αυτό το κομμάτι εφαρμόζουμε φίλτρο Wiener στην έξοδο που υπολογίσαμε προηγουμένως, με σκοπό την περαιτέρω βελτίωση του σήματος. Χωρίζουμε το σήμα μας σε overlapping frames στα οποία βρίσκουμε κάθε φορά το **φίλτρο Wiener** με τη μέθοδο που χρησιμοποιήθηκε και νωρίτερα. Τέλος, με **overlap add** ανασυνθέτουμε το σήμα που θα υπάρχει στην έξοδο. Σημειώνουμε ότι κατά την ανάλυση χρησιμοποιούμε παράθυρα hamming και ότι ο θόρυβος θεωρήθηκε στάσιμος, οπότε το φάσμα ισχύος του υπολογίστηκε μια φορά στην αρχή και θεωρήθηκε σταθερό. Το αποτέλεσμα φαίνεται παρακάτω όπου παρουσιάζονται τα διαγράμματα κυματομορφών και σπεκτρογραφημάτων των ζητούμενων σημάτων. Τα SSNRs που ζητούνται φαινονται και κατα την εκτέλεση του αρχείου. Το **SSNR** στην είσοδο του φίλτρου ακ ιστην έξοδο υπολογίστηκε και οι τιμές τους είναι 6.502 και 14.9833 αντίστοιχα. Η πρώτη τιμή υπολογίστηκε και στο προηγούμενο ερώτημα.



Σχ.9.1 Κυματομορφές εξόδου του beamformer με post filtering για πραγματικά σήματα, σε σύγκριση με το καθαρό , και αρχικό σήμα



Σχ.9.2 Σπεκτρογραφήματα για τα παραπάνω σήματα. Παρατηρούμε ότι ο θόρυβος που υπάρχει στο αρχικό σήμα έχει κατα ένα μεγάλο μέρος εξαλειφθεί στην τελική έξοδο.

References

- [1] John G. Proakis and Dimitris G. Manolakis. *Digital Signal Processing*. Prentice-Hall, 2010.
- [2] L.R Rabiner and R.W Schafer, *Digital Processing of speech signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978.
- [3] H.L Van Trees, *Optimum Array Processing*, Wiley, 2002
- [4] J. Benesty, J.Chen, Y. Huang, *Microphone Array Signal Processing*. Springer 2008