

Όνομα: Άγγελος

Επώνυμο: Ευστρατίου

ΑΜ: 03121113

Ψηφιακή Επεξεργασία Σημάτων

1η Εργαστηριακή Άσκηση

Μέρος 1^ο – Σύστημα Εντοπισμού Τηλεφωνικών Τόνων

Περιγραφή

Σκοπός της άσκησης είναι να δημιουργήσουμε σήματα που είναι το άθροισμα 2 ημιτόνων συγκεκριμένης συχνότητας και στη συνέχεια από τα σήματα να ανακατασκευάσουμε τους αριθμούς που αντιπροσωπεύουν. Κάθε δυάδα συχνοτήτων αντιστοιχεί σε έναν και μοναδικό αριθμό του δεκαδικού συστήματος σύμφωνα με τον παρακάτω πίνακα.

	Ω_column		
Ω_row	0.9273	1.0247	1.1328
0.5346	1	2	3
0.5906	4	5	6
0.6535	7	8	9
0.7217		0	

Πίνακας 1.1 - Διακριτές Συχνότητες για Τηλεφωνικούς Τόνους για Συχνότητα Δειγματοληψίας 8192 Hz.

1.1. Δημιουργία 10 διαφορετικά σήματα με 1000 δείγματα

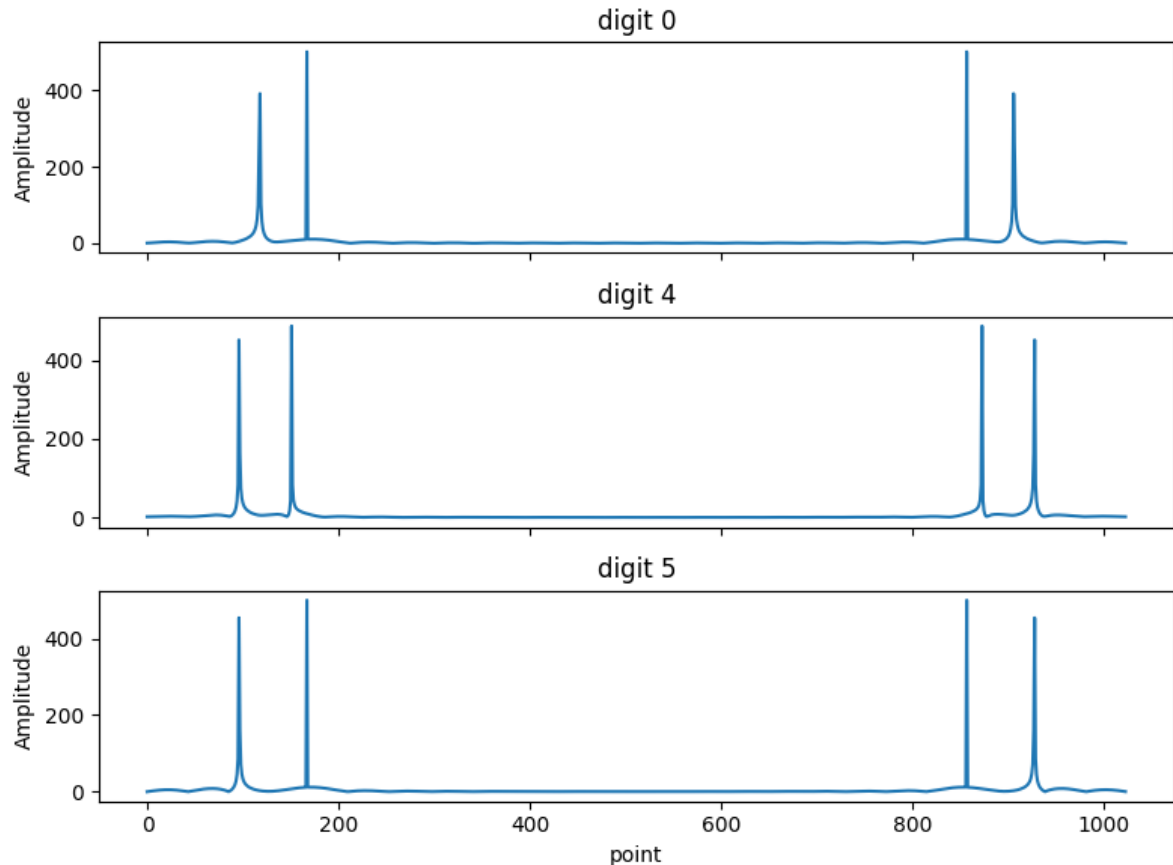
Για κάθε αριθμό στο δεκαδικό σύστημα μπορούμε να φτιάξουμε το αντίστοιχο σήμα που προκύπτει σύμφωνα με τον παραπάνω πίνακα. Έπειτα, μπορούμε να αποθηκεύσουμε όλα αυτά τα σήματα σε έναν 10x1000 πίνακα που θα τον ονομάσουμε `tone`, όπου το στοιχείο `tone[i][j]` είναι το j+1-οστό δείγμα του σήματος που κωδικοποιεί τον αριθμό i.

1.2. Υπολογισμός DFT των σημάτων

Χρησιμοποιώντας την συνάρτηση `fft` της βιβλιοθήκης `numpy` στα σήματα `tone[4]`, `tone[5]`, `tone[0]` υπολογίζουμε τον διακριτό μετασχηματισμό `fourier` των σημάτων. Δηλαδή ένα διάνυσμα με τους συντελεστές του μετασχηματισμένου σήματος. Επειδή τα σήματα έχουν 1000 δείγματα και ζητείται ο DFT 1024 σημείων, στα σήματα προσθέτουμε 24 μηδενικά δείγματα. Στη συνέχεια, υπολογίζουμε το μέτρο για κάθε ένα από τα 1024 σημεία κάθε διανύσματος, οπότε βρίσκουμε τα $|D_0[k]|$, $|D_4[k]|$ και $|D_5[k]|$. Απεικονίζουμε τα αποτελέσματα στο διάγραμμα 1.1. Παρατηρούμε πως υπάρχουν δύο μεγάλες κορυφές σε κάθε εικόνα, που αντιστοιχούν στις συχνότητες των σημάτων. Πιο συγκεκριμένα, η κορυφή που αντικατοπτρίζει τη συχνότητα ενός ημιτινοειδούς σήματος $\sin(\Omega \cdot n)$, βρίσκεται περίπου στο $(1024 / 2\pi) \cdot \Omega$ σημείο.

1.3. Δημιουργία αρχείου `tone_sequence.wav`

Για τον σκοπό του ερωτήματος, φτιάχνουμε ένα διάνυσμα με αριθμούς τα ψηφία του αριθμού μητρώου και το ονομάζουμε `identification_number`. Έπειτα αντιστοιχούμε το i – οστό δείγμα του διανύσματος, `identification_number[i]`, με το σήμα `tone[identification_number]`. Προσθέτουμε όλα αυτά τα σήματα σε ένα μεγάλο σήμα, το ένα μετά το άλλο, και τα διαχωρίζουμε με ένα διάνυσμα



Διάγραμμα 1.1 – Το μέτρο του fft των σημάτων tone[0], tone[4], tone[5]

που αποτελείται από 100 μηδενικά. Τέλος, αποθηκεύουμε το προκύπτον σήμα σε ένα .wav αρχείο με όνομα tone_sequence.wav.

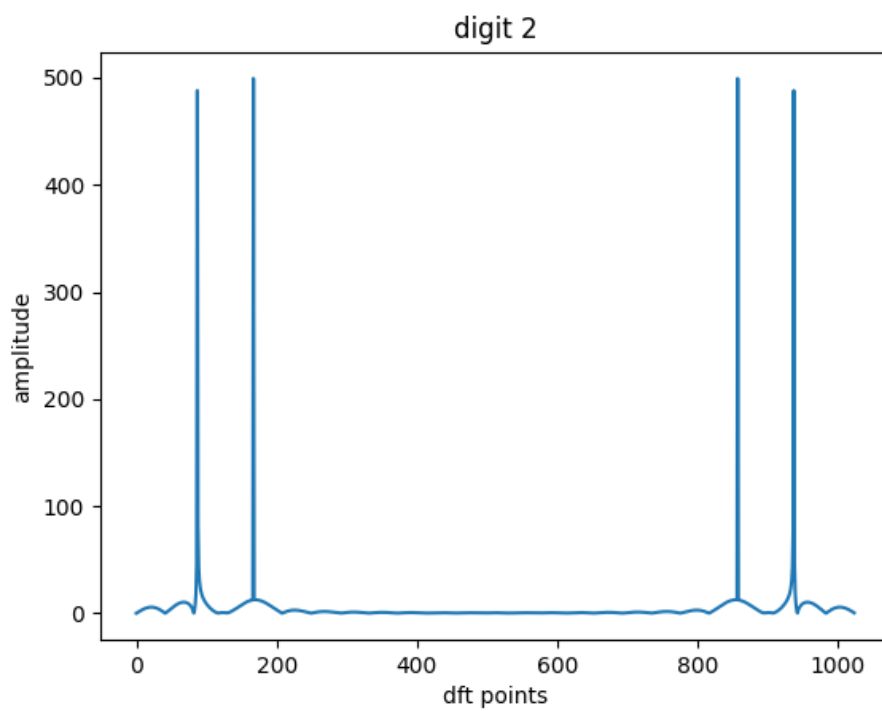
1.4. Υπολογισμός fft παραθυροποιημένων σημάτων

Θέλουμε να ανακατασκευάσουμε τα ψηφία του σήματος που δημιουργήσαμε στο προηγούμενο ερώτημα. Για να το καταφέρουμε αυτό πρέπει αρχικά να αποκόψουμε τα σήματα που αντιστοιχούν σε κάθε ψηφίο, δηλαδή τα `tone[identification_number]`. Αυτό θα το καταφέρουμε με τη χρήση παραθύρων. Πολλαπλασιάζουμε το παράθυρο με το μέρος του σήματος που θέλουμε να αποκόψουμε και έτσι διαχωρίζουμε τα σήματα. Παίρνουμε πάλι τον fft του αποκομένου σήματος και στη συνέχεια παίρνουμε το μέτρο των δειγμάτων του σήματος τα οποία μπορούμε να απεικονίσουμε. Στα διαγράμματα 1.2 και 1.3 φαίνονται τα αποτελέσματα για το 4^ο ψηφίο του αριθμού μητρώου (03121113) δηλαδή το 2. Την πρώτη φορά χρησιμοποιούμε ορθογώνια παράθυρα και τη δεύτερη παράθυρα hamming. Παρατηρούμε ότι τα αποτελέσματα είναι αρκετά καλύτερα με την χρήση hamming window επειδή οι κορυφές φαίνονται εντονότερα ενώ το πλάτος γύρω από τις κορυφές έχει μηδενιστεί.

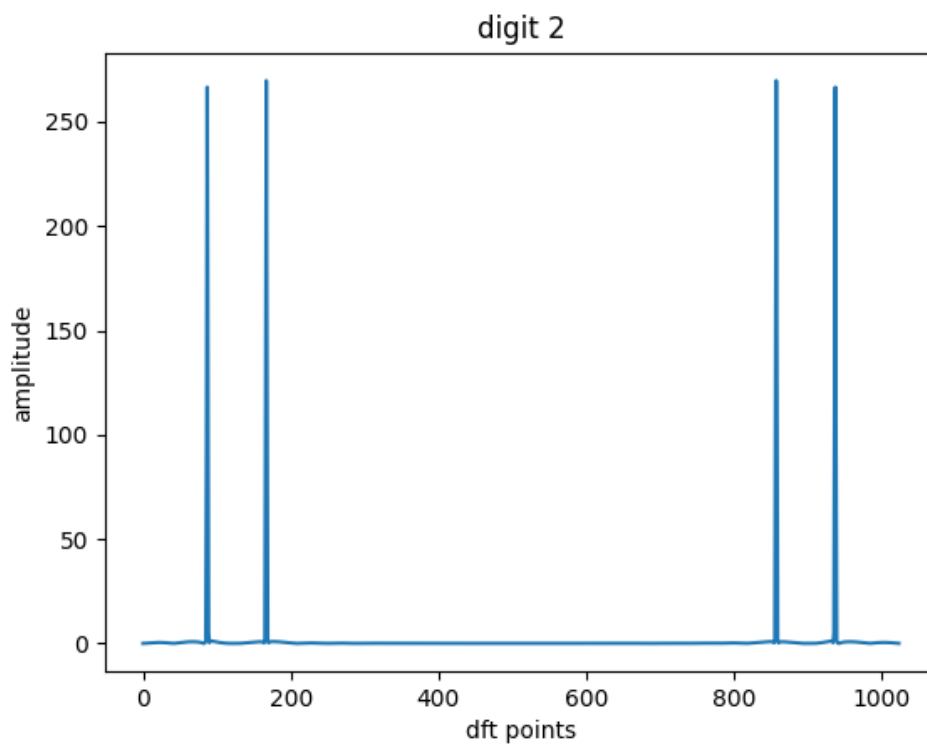
1.5.

1.6. Υπολογισμός ακολουθίας ψηφίων από το κωδικοποιημένο σήμα

Για να φτιάξουμε μια συνάρτηση που αποκοδικοποιεί ένα σήμα που δημιουργείται σύμφωνα με το ερώτημα 1.3. και επιστρέφει τα ψηφία σε ένα διάνυσμα, ακολουθούμε την εξής διαδικασία.



Διάγραμμα 1.2 – Μέτρο του fft του σήματος που προκύπτει από παραθυροποίηση το σήματος tone_sequence.wav με χρήση rectangular window



Διάγραμμα 1.3 – Μέτρο του fft του σήματος που προκύπτει από παραθυροποίηση το σήματος tone_sequence.wav με χρήση hamming window

Αρχικά, όπως κάναμε στο ερώτημα 1.4. απομονώνουμε τα σήματα που κωδικοποιούν το κάθε ψηφίο με τη χρήση παραθύρων. Έπειτα, υπολογίζουμε τον `fft` για κάθε παραθυροποιημένο σήμα και στη συνέχεια υπολογίζουμε το μέτρο του κάθε σημείου του μετασχηματισμένου σήματος. Όπως είδαμε στο προηγούμενο ερώτημα το αποτέλεσμα είναι ένα διάγραμμα με κορυφές σε σημεία που αντιστοιχούν στις συχνότητες των ημιτόνων του αρχικού σήματος. Χρησιμοποιώντας την συνάρτηση `find_peaks` της `scipy` βρίσκουμε τα σημεία που αντιστοιχούν σε αυτές τις κορυφές. Από το ερώτημα 1.2. είδαμε πως πρέπει να κανονικοποιήσουμε τα σημεία αυτά με την τιμή $(1024 / 2\pi)$ για να βρούμε τις πραγματικές συχνότητες των αρχικών σημάτων. Τώρα μένει μονάχα να αντιστοιχήσουμε τις συχνότητες αυτές με το σωστό ψηφίο σύμφωνα με τον πίνακα 1.1. Για να το κάνουμε αυτό, δημιουργούμε μια συνάρτηση, την `frequencies_to_digit`. Αυτή η συνάρτηση επιστρέφει το ψηφίο που η δυάδα συχνότητων του ταιριάζει καλύτερα με τις συχνότητες που υπολογίσαμε. Εδώ με τη λέξη ταιριάζει, εννοούμε πως ελαχιστοποιεί την ευκλείδεια απόσταση των δύο ζευγαριών συχνότητων. Κάνοντας αυτήν τη διαδικασία για κάθε παραθυροποιημένο σήμα, βρίσκουμε την ζητούμενη ακολουθία ψηφίων, την οποία και επιστρέφουμε. Δοκιμάζουμε αυτή τη συνάρτηση (`ttdecode`) με όρισμα εισόδου το σήμα `tone_sequence` που φτιάξαμε στο ερώτημα 1.3 και χαιρόμαστε όταν βλέπουμε πως πράγματι επιστρέφει τον αριθμό μητρώου 0 3 1 2 1 1 3.

1.7. Αποκωδικοποίηση σημάτων αποθηκευμένων σε αρχεία με χρήση της `ttdecode`.

Φορτώνουμε τα σήματα `easy_sig`, `medium_sig` και `hard_sig` και τα αποκωδικοποιούμε με τη συνάρτηση που φτιάξαμε στο προηγούμενο ερώτημα. Τελικά βρίσκουμε τα ακόλουθα αποτελέσματα:

`easy_sig`: 9 3 9 2 6 8 4 8

`medium_sig`: 2 7 3 3 2 5 4 2 5 4

`hard_sig`: 6 9 2 6 3 1 9 9 1 0

Μέρος 2^ο - Χαρακτηριστικά Βραχέος Χρόνου Σημάτων Φωνής και Μουσικής

Περιγραφή

Σε αυτό το μέρος θέλουμε να πάρουμε πληροφορίες για το σήμα καθώς περνάει ο χρόνος. Αυτό το καταφέρνουμε, χρησιμοποιώντας ένα παράθυρο που απομονώνει χρονικά τμήματα του σήματος. Στη συνέχεια, μπορούμε να επεξεργαστούμε το κάθε τμήμα και να εξάγουμε χρήσιμα χαρακτηριστικά του σήματος.

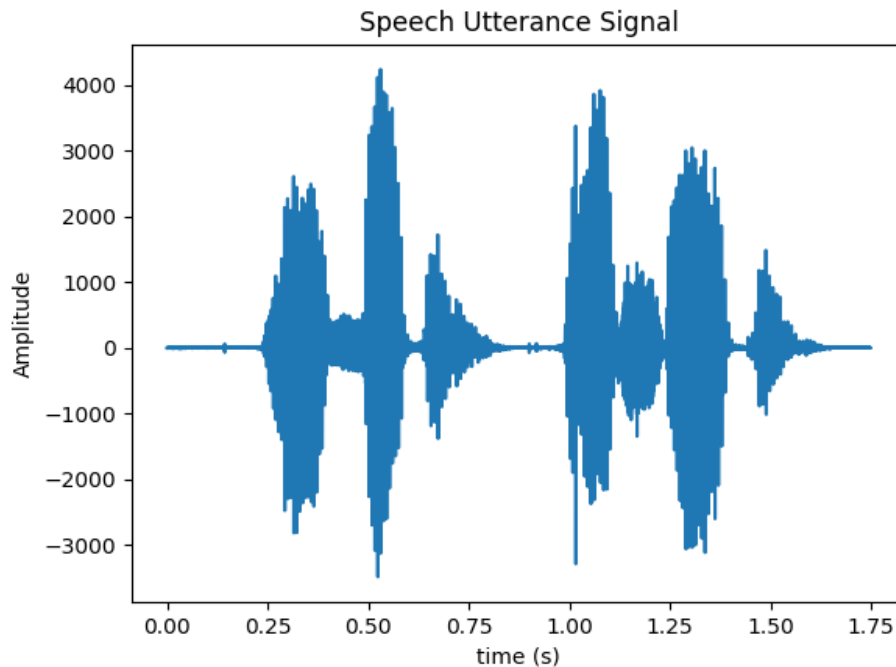
2.1. Φόρτωση αρχείου

Φορτώνουμε το αρχείο `"speech_utterance.wav"` σε μια `numpy array` και τη συχνότητα δειγματοληψίας του σήματος. Το σήμα είναι ένα σήμα ήχου όπου ένας άνθρωπος λέει `"Oh mother, I saw them"`. Το σήμα διαρκεί περίπου 1.7 s και δειγματοληπτείται στα 16 kHz. Η απεικόνισή του στο πεδίο του χρόνου φαίνεται στο διάγραμμα 2.1.

2.2. Υπολογίζουμε την ενέργεια βραχέος χρόνου του σήματος καθώς και τον ρυθμό εναλλαγής προσήμου

Για τον υπολογισμό της ενέργειας βραχέος χρόνου που δίνεται από τον τύπο:

$$E_n = \sum_{m=-\infty}^{\infty} |x[m]|^2 w[n-m]$$



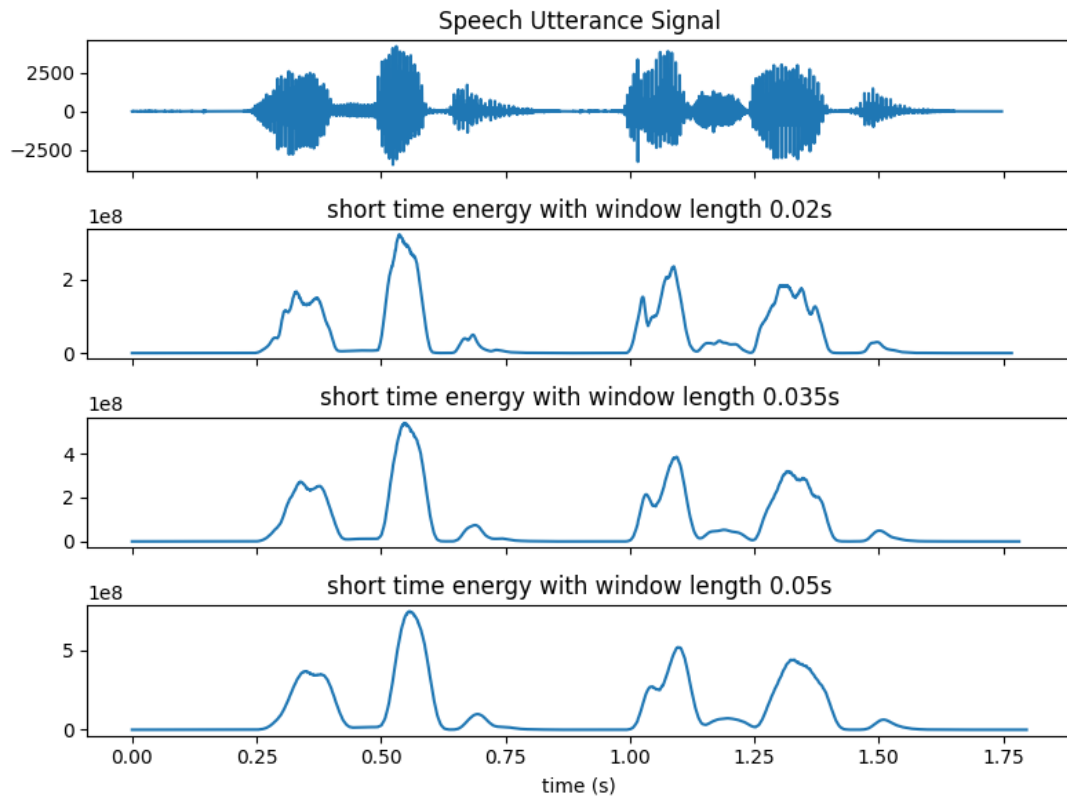
Διάγραμμα 2.1 – Σήμα φωνής από το αρχείο “speech_utterance.wav”

Παρατηρούμε πως συμπίπτει με τη συνέλιξη του σήματος $|x[n]|^2$ με το παράθυρο $w[n]$. Έτσι, με τη χρήση της συνάρτησης `convolve` της `numpy` μπορούμε να μειώσουμε τους υπολογισμούς που χρειάζονται για την εύρεση της ενέργειας βραχέος χρόνου. Η ίδια ιδέα λειτουργεί εξίσου καλά και για τον υπολογισμό του ρυθμού εναλλαγής προσήμου του σήματος. Χρησιμοποιούμε παράθυρα `hamming` μήκους 20-50 ms και απεικονίζουμε τα αποτελέσματα στα διαγράμματα 2.2 και 2.3.

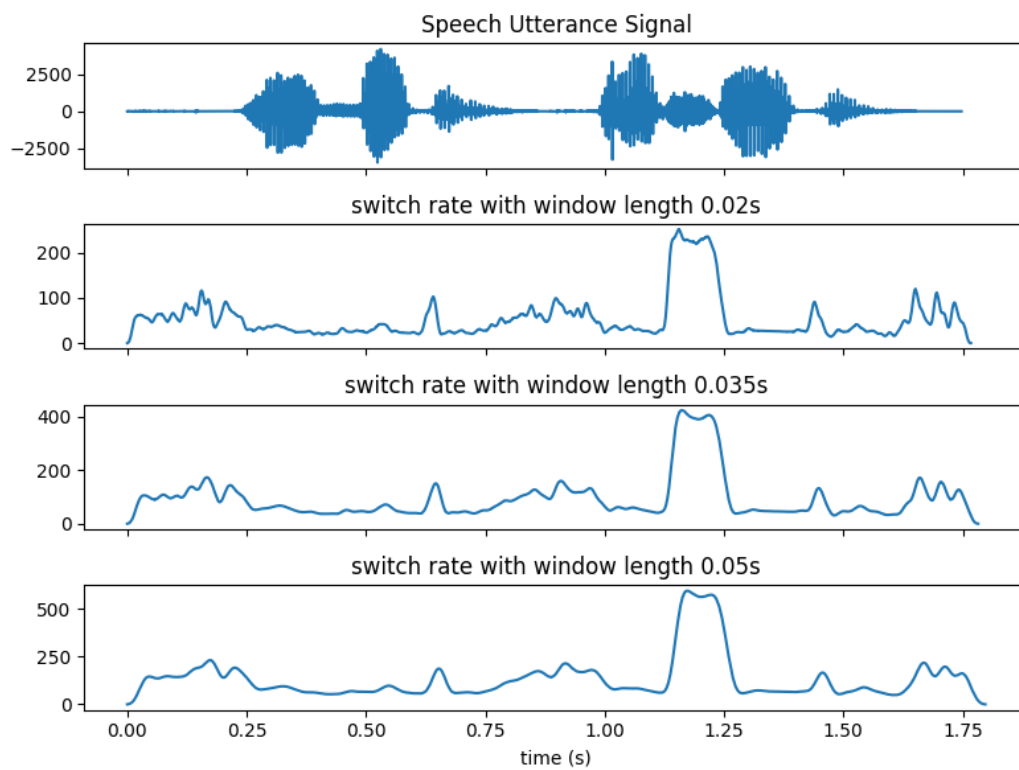
Καθώς μεγαλώνει το μήκος του παραθύρου, βλέπουμε πως εξομαλύνεται το υπολογιζόμενο σήμα και γενικά παίρνει μεγαλύτερες τιμές, επειδή προσθέτουμε περισσότερους όρους. Με τη χρήση του διαγράμματος ενέργειας βραχέος χρόνου μπορούμε να διαχωρίσουμε τη σιωπή από τη φωνή. Όταν ο άνθρωπος μιλάει η τιμή της ενέργειας παίρνει μη μηδενικές τιμές ενώ όταν δε μιλάει η τιμή της ενέργειας είναι περίπου μηδενική. Από το ρυθμό εναλλαγής προσήμου μπορούμε να διαχωρίσουμε σε αρκετές περιπτώσεις τα φωνήεντα από τα σύμφωνα. Συνήθως τα φωνήεντα συγκεντρώνουν την ενέργειά τους σε χαμηλότερες συχνότητες από τα σύμφωνα. Αυτό σημαίνει ότι περιμένουμε να δούμε περισσότερες εναλλαγές προσήμου στα σύμφωνα. Αυτό φαίνεται εύκολα στο διάγραμμα 2.3 όπου οι μεγαλύτερες τιμές του ρυθμού εναλλαγής προσήμου φαίνονται όταν ο άνθρωπος προφέρει το “s” από το “saw” και το “m” από το “them”.

2.3. Υπολογισμός μετασχηματισμού Fourier βραχέος χρόνου

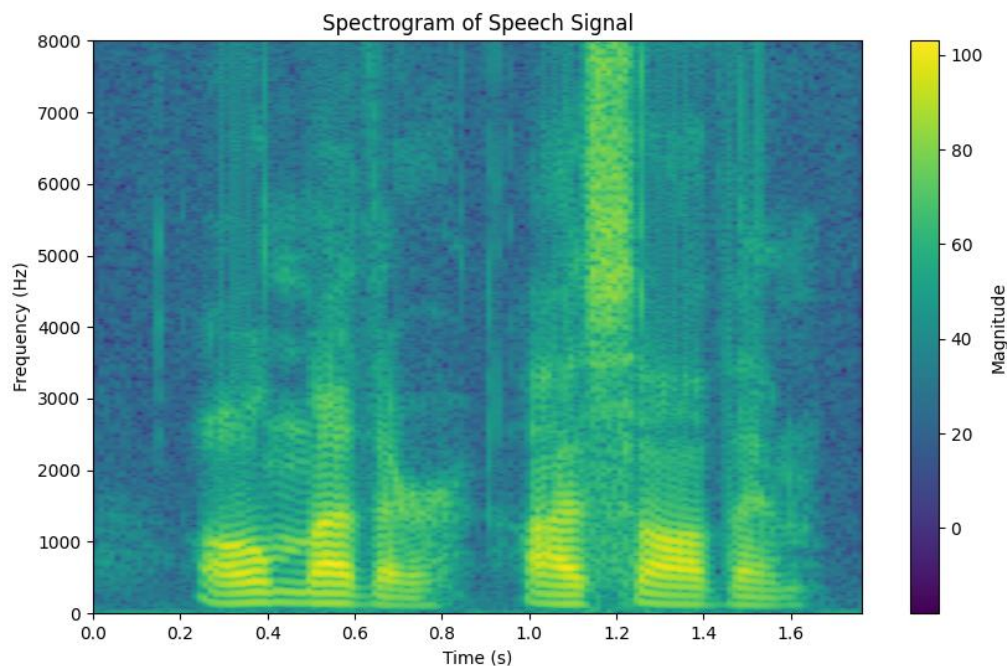
Χρησιμοποιώντας την συνάρτηση `ShortTimeFFT` της `scipy` υπολογίζουμε τον μετασχηματισμό Fourier βραχέος χρόνου του σήματος φωνής και έπειτα χρησιμοποιούμε το αποτέλεσμα για να βρούμε το φασματικό κέντρο και τη φασματική ροή. Παίρνουμε μήκος παραθύρου 25ms και βάζουμε τα παράθυρα να έχουν επικάλυψη 50%. Τα αποτελέσματα φαίνονται στα διαγράμματα 2.4 – 2.6.



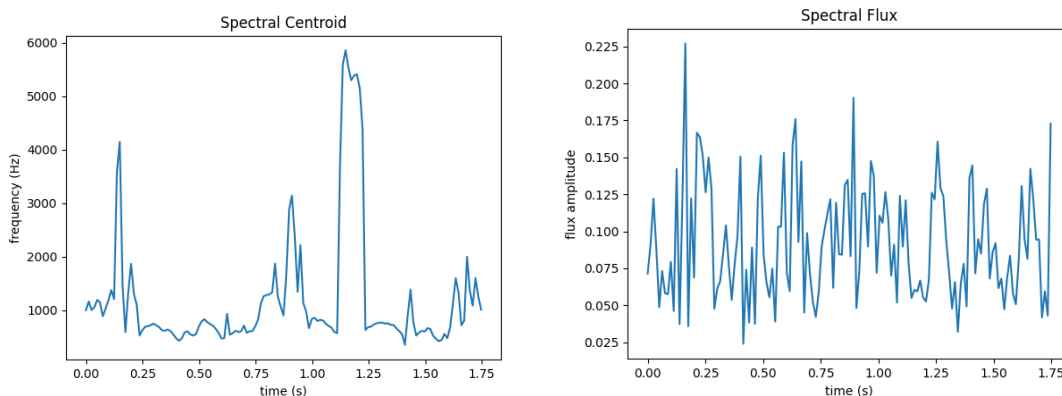
Διάγραμμα 2.2 – Ενέργεια βραχέος χρόνου σήματος



Διάγραμμα 2.3 – Ρυθμός εναλλαγής προσήμου σήματος



Διάγραμμα 2.4 – Φασματογράφημα σήματος φωνής (πλάτος σε dB)

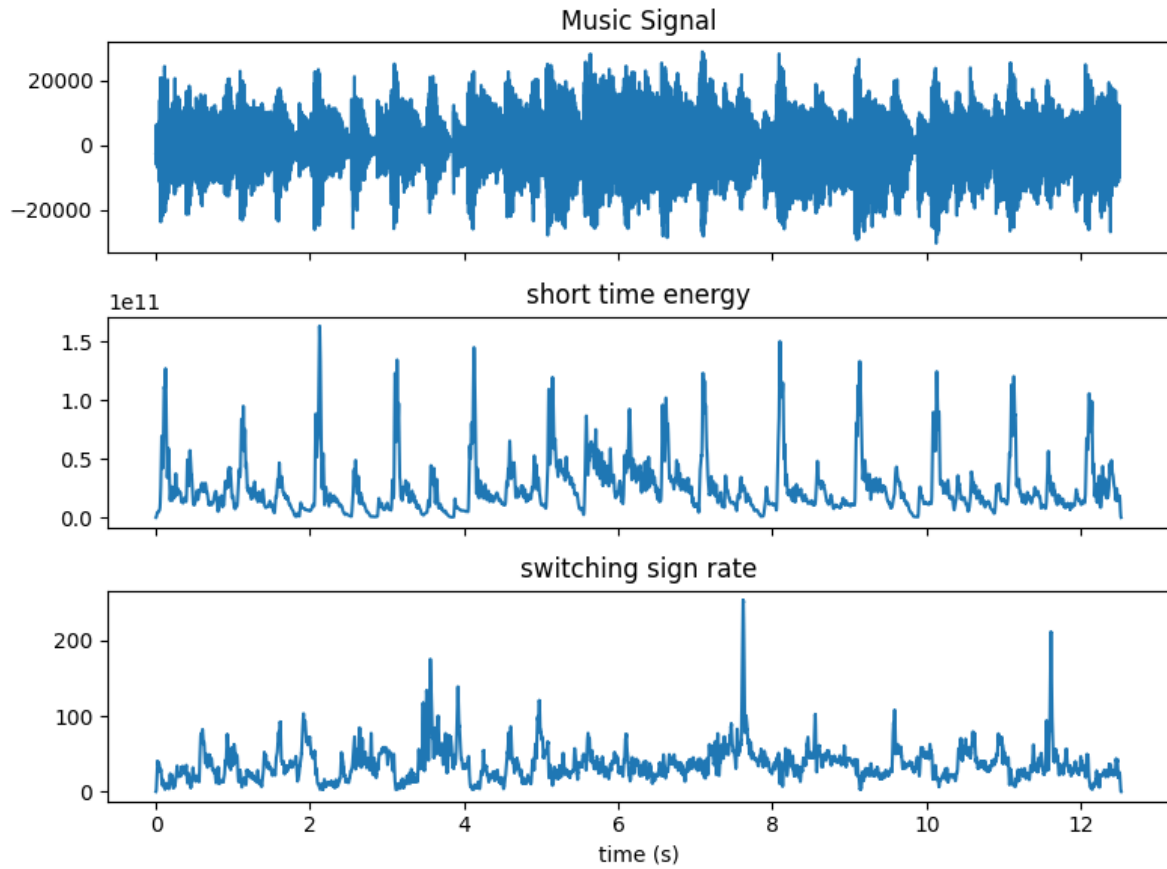


Διάγραμμα 2.5 – Φασματικό κέντρο σήματος / Διάγραμμα 2.6 – Φασματική ροή σήματος

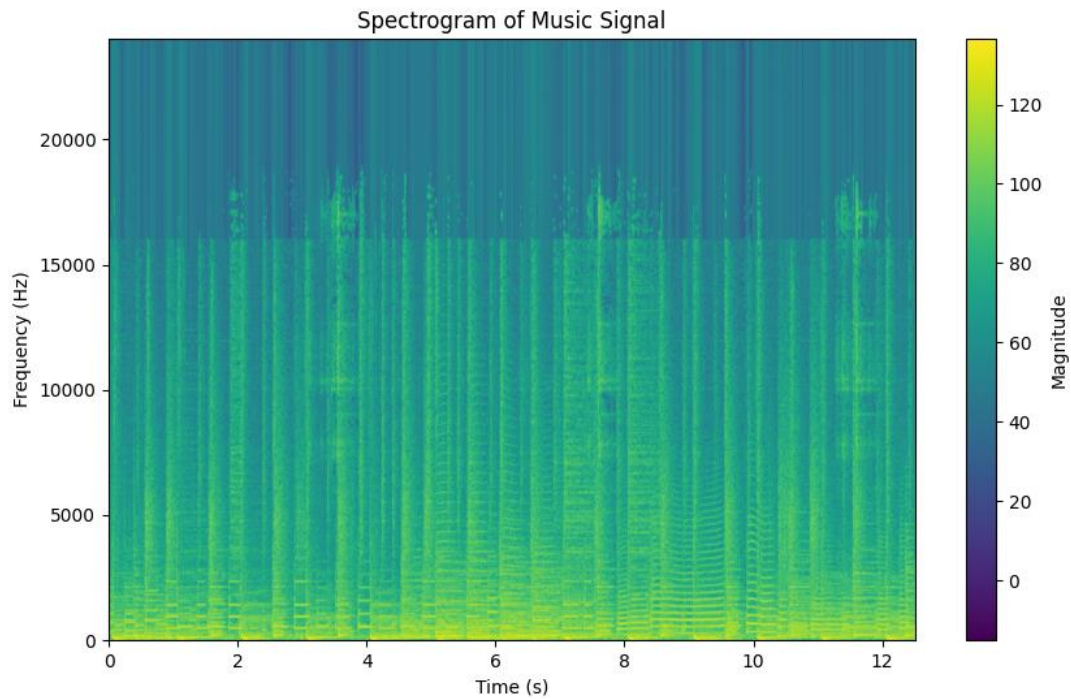
Από το φασματογράφημα παρατηρούμε πως το μεγαλύτερο μέρος της ενέργειας του σήματος συγκεντρώνεται από τα 0 – 1000 Hz. Μπορούμε να καταλάβουμε πότε μιλάει ο ομιλητής παρατηρώντας τις χρονικές στιγμές που έχουμε μεγάλα πλάτη και μπορούμε πάλι να διακρίνουμε πότε ο ομιλητής προφέρει το “s” επειδή στα 1.2 s έχουμε μεγάλο πλάτος στις υψηλές συχνότητες. Από το διάγραμμα του φασματικού κεντροειδούς παίρνουμε πληροφορίες ανάλογες με αυτές που πήραμε από το ρυθμό εναλλαγής προσήμου, μάλιστα τα δύο διαγράμματα μοιάζουν πολύ στη μορφή τους. Το διάγραμμα της φασματικής ροής είναι δύσκολο να ερμηνευτεί.

2.4 Ίδια διαδικασία για το σήμα μουσικής “music.wav”

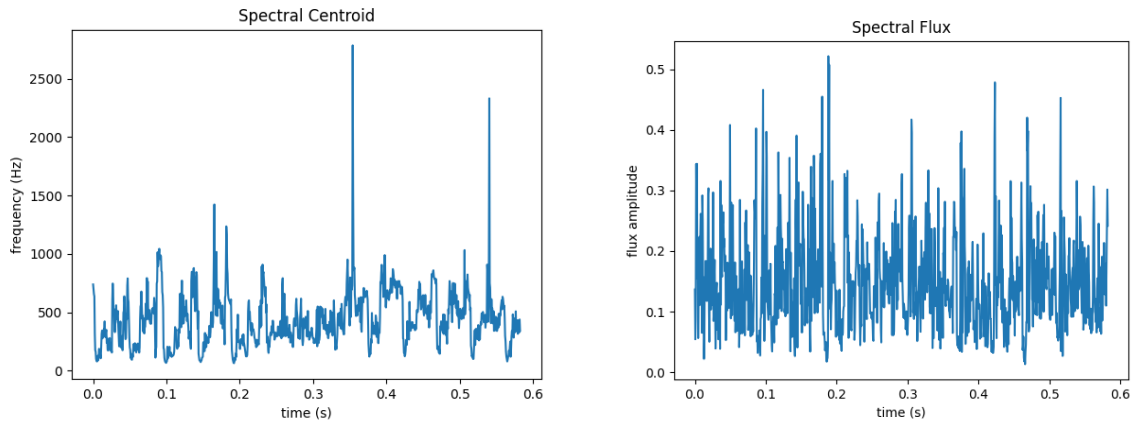
Χρησιμοποιώντας και πάλι παράθυρα μήκους 25 ms υπολογίζουμε όσα υπολογίσαμε πριν για το σήμα μουσικής. Όλες οι πληροφορίες φαίνονται στα διαγράμματα 2.7 – 2.10.



Διάγραμμα 2.7 – Χαρακτηριστικά του σήματος μουσικής στο πεδίο του χρόνου



Διάγραμμα 2.8 – Φασματογράφημα σήματος μουσικής (πλάτος σε dB)



Διάγραμμα 2.9 – Φασματικό κέντρο σήματος / Διάγραμμα 2.10 – Φασματική ροή σήματος

Από το διάγραμμα της ενέργειας βραχέος χρόνου του σήματος μουσικής, μπορούμε να καταλάβουμε πότε γίνονται πράξεις που αυξάνουν έντονα την ένταση του ήχου, όπως πάτημα πλήκτρου στο πιάνο κλπ. Από το διάγραμμα ρυθμού εναλλαγής προσήμου μπορούμε να καταλάβουμε μερικές φορές τα όργανα που παίζουν ή αν αρχίσει να παίζει κάποιο όργανο κλπ. Στη χρονική στιγμή 7.8 s για παράδειγμα που το πλάτος στο διάγραμμα ρυθμού εναλλαγής προσήμου αυξάνεται απότομα μπορούμε να ακούσουμε το χτύπημα από ένα πιατίνι στα drums. Στο διάγραμμα 2.8 (φασματογράφημα) παίρνουμε πολύτιμες πληροφορίες σχετικά με το σε ποιες συχνότητες είναι συγκεντρωμένη η ενέργεια του σήματος. Σε αντίθεση με το σήμα φωνής στο σήμα μουσικής έχουμε μεγάλα πλάτη σε συχνότητες 0 – 18 kHz αν και το πλάτος είναι συνήθως μεγαλύτερο στις χαμηλές συχνότητες. Παρόμοιες πληροφορίες με τα προηγούμενα διαγράμματα μπορούμε να πάρουμε και παρατηρώντας τα διαγράμματα 2.9 και 2.10, φασματικό κέντρο σήματος και φασματική ροή σήματος αντίστοιχα.

Μέρος 3ο - Ψηφιακή Επεξεργασία και Φασματική Ανάλυση Βιοσημάτων Κητωδών

Περιγραφή

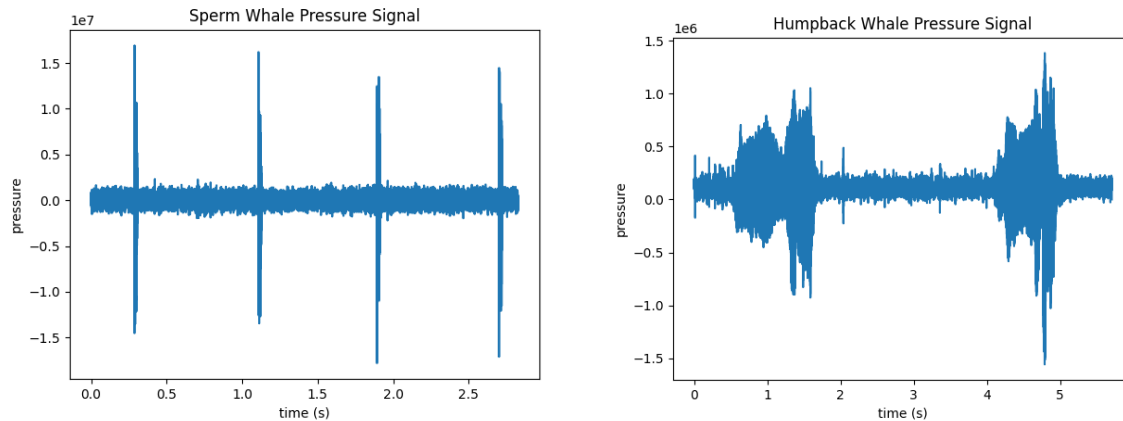
Στη συγκεκριμένη άσκηση θέλουμε να επεξεργαστούμε δύο σήματα από ήχους φαλαινών που έχουν συλλεχθεί με υδρόφωνο. Το ένα είναι από μια φάλαινα φυσητήρα (sperm whale) ενώ το άλλο από μια μεγάπτερη φάλαινα (humpback whale).

3.1. Μετατροπή από σήματα τάσης σε σήματα πίεσης

Τα δύο σήματα που έχουμε συλλέξει με το υδρόφωνο είναι σήματα τάσης. Με χρήση του τύπου:

$$pressure = \frac{3.5 * y * GF}{10^{\left(\frac{sensitivity}{20}\right)}}$$

Μπορούμε να μετατρέψουμε τα σήματα τάσης σε σήματα πίεσης. Οι παράμετροι GF και sensitivity δίνονται, ενώ y είναι το σήμα τάσης. Εφαρμόζουμε τη σχέση και απεικονίζουμε τα αποτελέσματα στα διαγράμματα 3.1, 3.2.



Διάγραμμα 3.1 – Σήμα πίεσης για τη φυσητήρα / Διάγραμμα 3.2 – Σήμα πίεσης για τη μεγάπτερη

3.2. Υπολογισμός της p_{rms} και SPL_{rms}

Σε αυτό το ερώτημα θέλουμε να υπολογίσουμε τη μέση τετραγωνική ρίζα της πίεσης καθώς και το επίπεδο ηχητικής πίεσης που δίνονται από τις σχέσεις:

$$p_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N pressure_i^2} \text{ και } SPL_{RMS} = 20 \log_{10} \frac{p_{rms}}{p_{ref}}$$

Αντίστοιχα. Τελικά βρίσκουμε τα αποτελέσματα:

sperm rms pressure = 749840.25

humpback rms pressure = 195774.66

sperm SPL rms = 237.49937497627803

humpback SPL rms = 225.8351294017253

3.3. Επαλήθευση του θεωρήματος Parseval

Θα επαληθεύσουμε το θεώρημα Parseval για το σήμα πίεσης της φάλαινας φυσητήρα. Το θεώρημα Parseval λέει πως η ενέργεια του σήματος στο πεδίο του χρόνου είναι ίση με την ενέργεια του σήματος στο πεδίο της συχνότητας. Υπολογίζουμε την ενέργεια και με τους δύο τρόπους και βρίσκουμε:

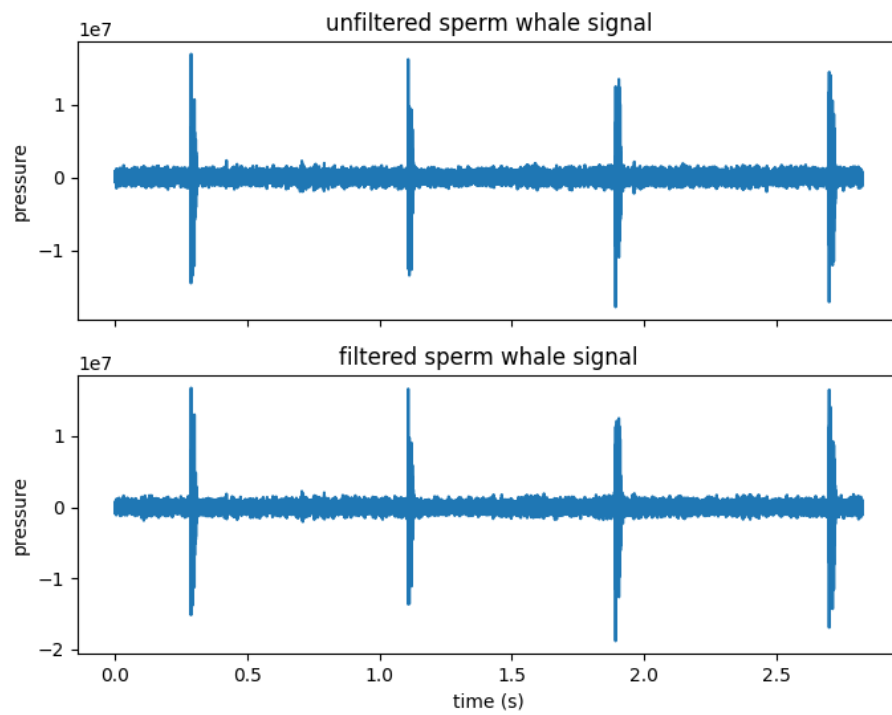
energy in the time domain = 3.50434506250816e+16

energy in the frequency domain = 3.5043447763147804e+16

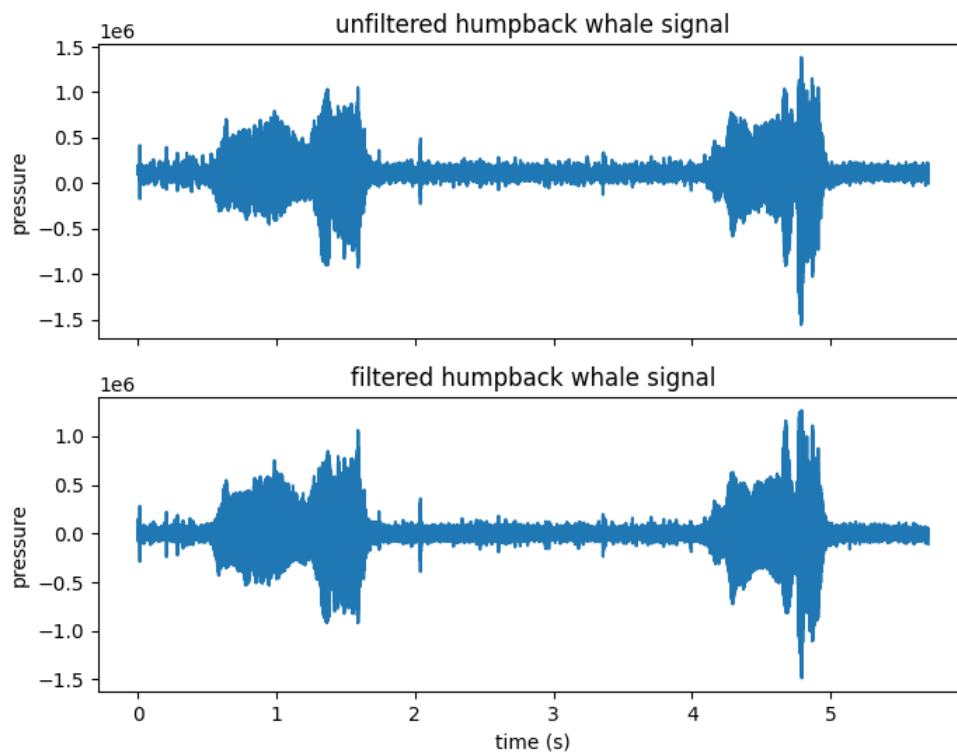
Πράγματι είναι περίπου ίσες.

3.4. Φτιάχνουμε φίλτρα για να απομονώσουμε τις υψηλές συχνότητες του σήματος πίεσης.

Επειδή ο ανθρωπογενής θόρυβος εντοπίζεται συνήθως στις χαμηλές συχνότητες, θα δημιουργήσουμε φίλτρα butterworth τάξης 3 για να αποκόψουμε τις συχνότητες που είναι μικρότερες από 200 Hz από τα σήματά μας. Τα αποτελέσματα φαίνονται στα διαγράμματα 3.3, 3.4.



Διάγραμμα 3.3 – Σήμα πίεσης για τη φάλαινα φυσητήρα πριν και μετά το φιλτράρισμα



Διάγραμμα 3.4 – Σήμα πίεσης για τη megάλπτερη φάλαινα πριν και μετά το φιλτράρισμα

3.5. Εξατάζουμε αν ο ανθρωπογενής υποθαλάσσιος ήχος επηρεάζει τη συμπεριφορά των φαλαινών

Κάνουμε την υπόθεση πως αν τα επίπεδα ηχητικής πίεσης SPL_{rms} ξεπερνούν τα 100 dB relative to 1μPa, η συμπεριφορά των φαλαινών επηρεάζεται. Μετρώντας την SPL_{rms} για το σήμα pile driving.wav, ένα σήμα τάσης που έχει μετρηθεί και πάλι με υδρόφωνο, βρίσκουμε:

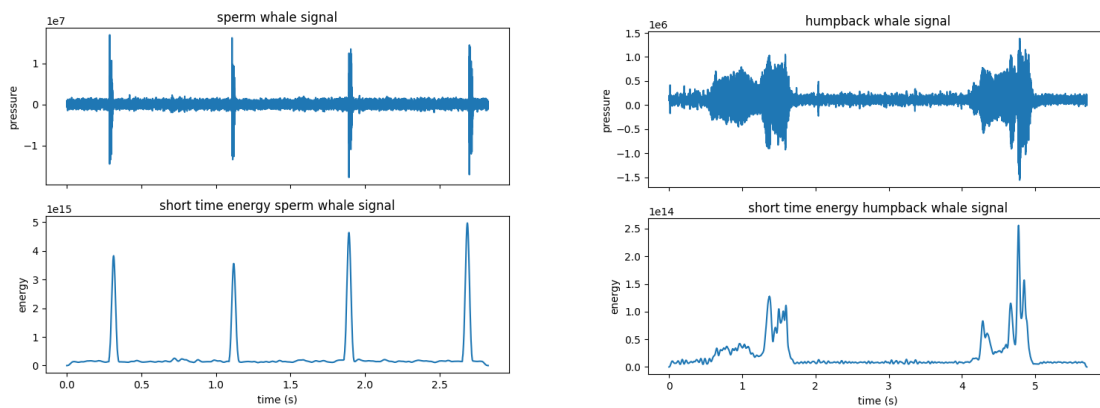
pile pile driving SPL rms = 285.1529718143794

pile driving SPL rms value exceeds 100 dB relative to 1uPa

Οπότε η συμπεριφορά των φαλαινών όντως επηρεάζεται από τον ανθρωπογενή υποθαλάσσιο ήχο.

3.6. Υπολογισμός ενέργειας βραχέος χρόνου

Υπολογίζουμε την ενέργεια βραχέος χρόνου για τα δύο σήματα πίεσης και απεικονίζουμε τα αποτελέσματα στα διαγράμματα 3.5, 3.6.



Διάγραμμα 3.5 – ενέργεια βραχέος χρόνου για τη φυσητήρα / Διάγραμμα 3.6 – ενέργεια βραχέος χρόνου για τη μεγάλη πτερη

Στα διαγράμματα φαίνεται πως η ενέργεια βραχέος χρόνου σχεδόν περιβάλλει το αρχικό σήμα με τη διαφορά ότι είναι κλιμακωμένη με μία σταθερά.

3.7. Υπολογισμός ενεργειακού τελεστή Teager-Kaiser για τα δύο σήματα

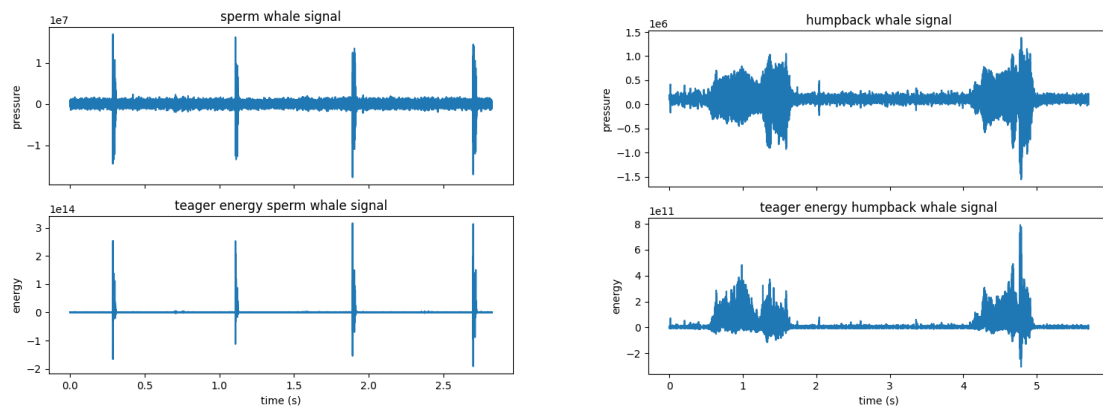
Ο υπολογισμός του ενεργειακού τελεστή Teager-Kaiser για διακριτά σήματα δίνεται από τον τύπο:

$$\Psi[p[n]] = p^2[n] - p[n-1]p[n+1]$$

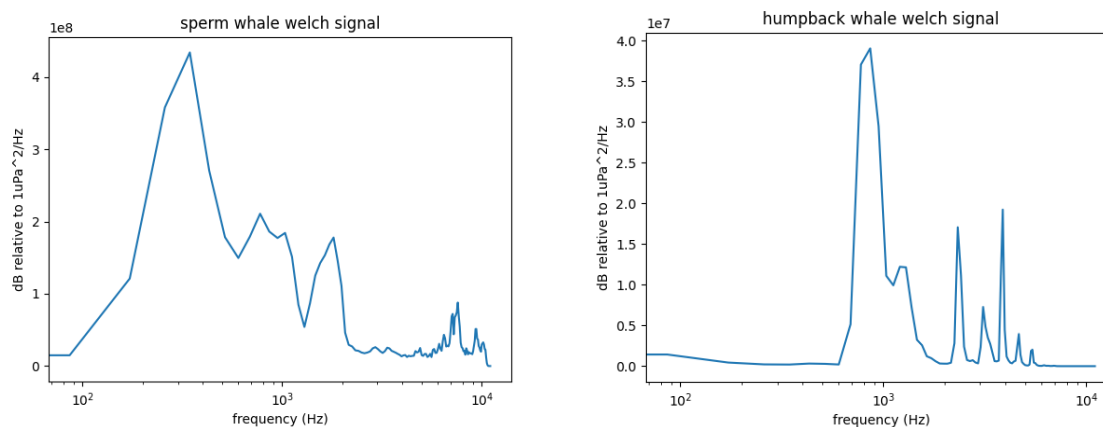
Εφαρμόζουμε τον τελεστή στα δύο σήματα πίεσης και απεικονίζουμε τα αποτελέσματα στα διαγράμματα 3.7, 3.8. Παρατηρούμε πως ο θόρυβος έχει μειωθεί με την εφαρμογή του τελεστή. Ακόμη, σε αντίθεση με την ενέργεια βραχέος χρόνου βλέπουμε πως το σήμα που προκύπτει με την εφαρμογή του τελεστή παίρνει και θετικές και αρνητικές τιμές.

3.8. Εκτιμητής φασματικής πυκνότητας για τα δύο σήματα πίεσης

Χρησιμοποιώντας την συνάρτηση `scipy.signal.welch` μπορούμε να υπολογίσουμε έναν εκτιμητή φασματικής πυκνότητας ενός σήματος από το μέσο τροποποιημένο περιοδόγραμμα Welch. Απεικονίζουμε τα αποτελέσματα για τα δύο σήματα στα διαγράμματα 3.9, 3.10. Από τα διαγράμματα συμπεραίνουμε πως η ενέργεια του σήματος της φάλαινας φυσητήρα συγκεντρώνεται στις χαμηλές συχνότητες ενώ της μεγάλης πτερης φάλαινας στις υψηλές.



Διάγραμμα 3.7 – Εφαρμογή τελεστή Teager-Kaiser στο σήμα της φάλαινας φυσητήρα / Διάγραμμα 3.8 - Εφαρμογή τελεστή Teager-Kaiser στο σήμα της megάλπτερης φάλαινας



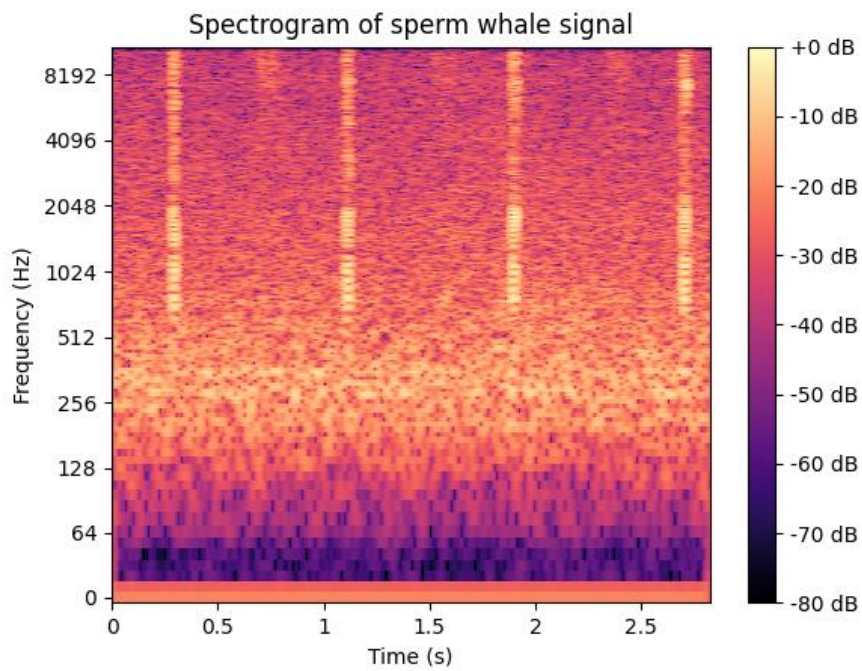
Διάγραμμα 3.9 – Περιοδόγραμμα Welch για το σήμα της φάλαινας φυσητήρα / Διάγραμμα 3.10 - Περιοδόγραμμα Welch για το σήμα της megάλπτερης φάλαινας

3.9. Σχεδίαση φασματογραφημάτων των δύο σημάτων

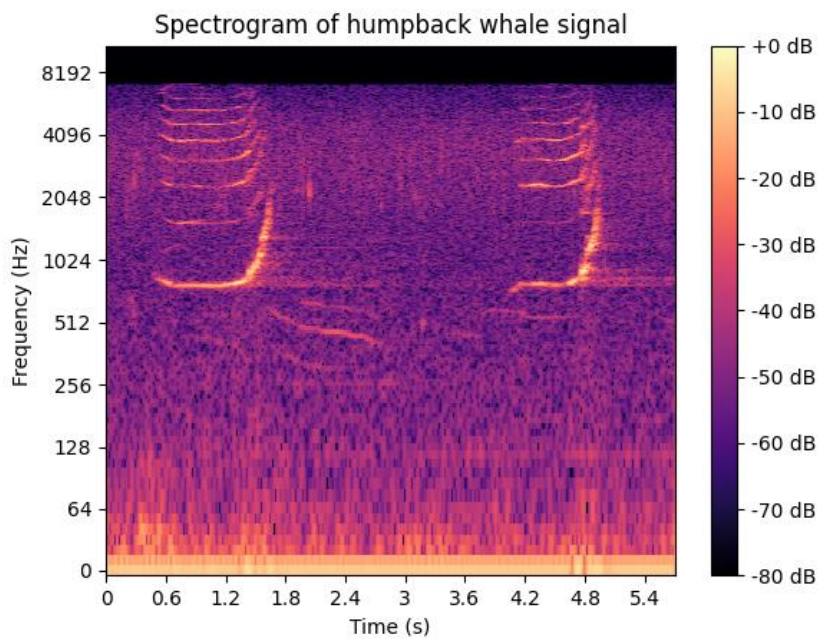
Υπολογίζουμε τον short time Fourier transform για τα δύο σήματα και απεικονίζουμε τα αποτελέσματα σε φασματογραφήματα στα διαγράμματα 3.11, 3.12. Τα συμπεράσματα που βγάλαμε στο προηγούμενο ερώτημα συμφωνούν με όσα βλέπουμε στα φασματογραφήματα εν μέρει. Όντως στο φασματογράφημα της megάλπτερης φάλαινας βλέπουμε περισσότερες κόκκινες περιοχές (μεγάλες τιμές) συγκεντρωμένες σε υψηλές συχνότητες, σε αντίθεση με το φασματογράφημα της φάλαινας φυσητήρα. Όμως αυτό πιθανόν να οφείλεται στον θόρυβο που υπάρχει στις χαμηλές συχνότητες στο σήμα της φάλαινας φυσητήρα.

3.10. Εφαρμόζουμε την τεχνική PCEN και σχεδιάζουμε τα mel-spectrograms των δύο σημάτων

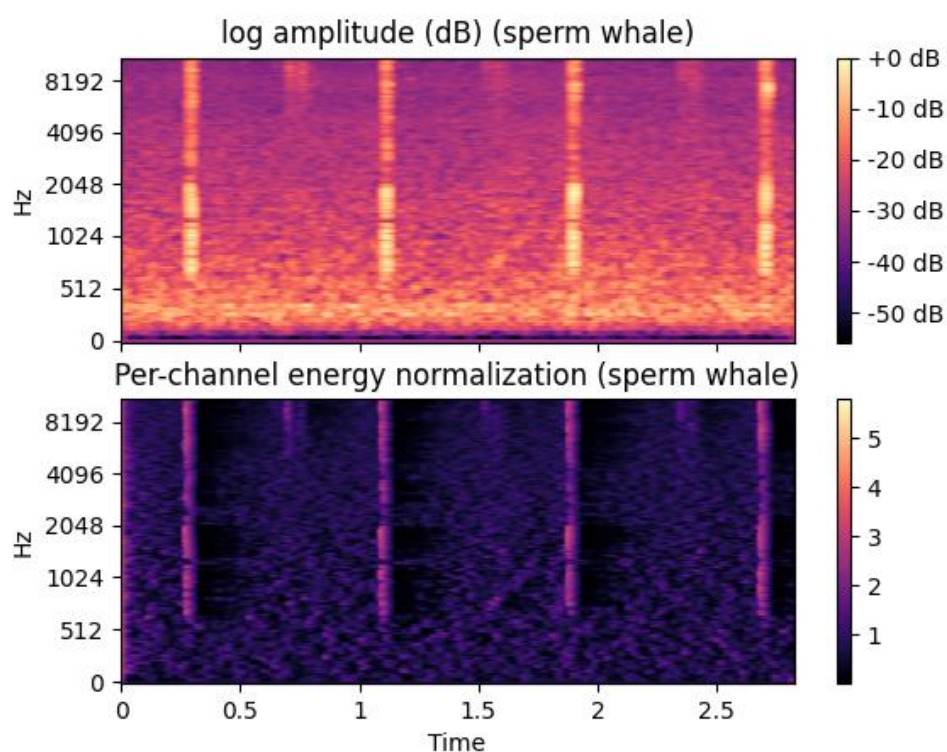
Για να ενισχύσουμε την αντίθεση ανάμεσα στον background θόρυβο και τα foreground μεταβλητά γεγονότα εφαρμόζουμε την τεχνική per channel energy normalization. Για να δούμε τα αποτελέσματα τα απεικονίζουμε σε mel-spectrograms στα διαγράμματα 3.13, 3.14. Όπως βλέπουμε το πρόβλημα που είχαμε στο προηγούμενο ερώτημα με τον θόρυβο στο σήμα της φάλαινας φυσητήρα έχει, σε μεγάλο βαθμό, αντιμετωπιστεί.



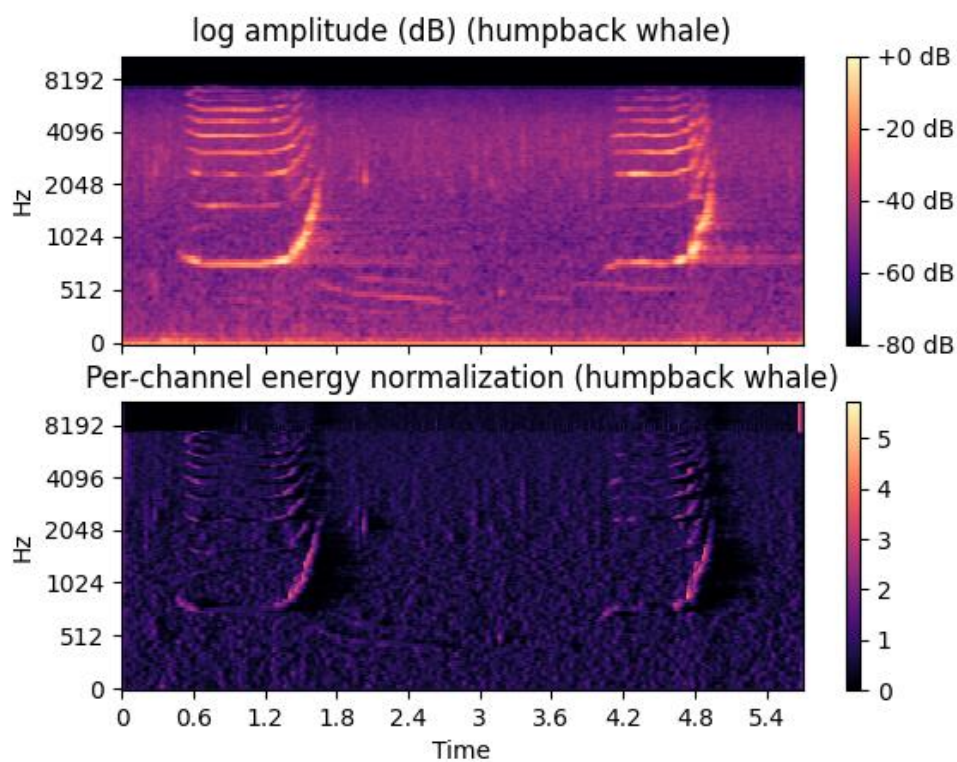
Διάγραμμα 3.11 – Φασματογράφημα του stft του σήματος της φάλαινας φυσητήρα



Διάγραμμα 3.12 – Φασματογράφημα του stft του σήματος της μεγάπτερης φάλαινας



Διάγραμμα 3.13 – mel-spectrogram του σήματος της φάλαινας φυσητήρα πριν και μετά την εφαρμογή PCEN



Διάγραμμα 3.14 – mel-spectrogram του σήματος της megάλπτερης φάλαινας πριν και μετά την εφαρμογή PCEN